



**UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE ESTUDOS DA LINGUAGEM**

**Lara Ribeiro da Silva**

**As diferenças entre o que se fala e o que se escreve no  
português do Brasil: a aquisição do clítico se indeterminador e  
apassivador**

**CAMPINAS,  
2017**

**Lara Ribeiro da Silva**

**As diferenças entre o que se fala e o que se escreve no português do Brasil: a aquisição do clítico se indeterminador e apassivador**

**Dissertação de mestrado apresentada ao Instituto de Estudos da Linguagem da Universidade Estadual de Campinas para obtenção do título de Mestra em Linguística.**

**Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Ruth Elisabeth Vasconcellos Lopes  
Co-Orientador (a): Prof.<sup>a</sup> Dr.<sup>a</sup> Mary Aizawa Kato**

**Este exemplar corresponde à versão final da Dissertação defendida pela aluna Lara Ribeiro da Silva e orientada pela Prof.<sup>a</sup> Dr.<sup>a</sup> Ruth Elisabeth Vasconcellos Lopes e pela Prof.<sup>a</sup> Dr.<sup>a</sup> Mary Aizawa Kato.**

**CAMPINAS,  
2017**

**Agência(s) de fomento e nº(s) de processo(s): CAPES**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Estudos da Linguagem  
Crisllene Queiroz Custódio - CRB 8/8624

Si38d Silva, Lara Ribeiro, 1993-  
As diferenças entre o que se fala e o que se escreve no português do Brasil : a aquisição do clítico se indeterminador e apassivador / Lara Ribeiro da Silva. – Campinas, SP : [s.n.], 2017.

Orientador: Ruth Elisabeth Vasconcellos Lopes.

Coorientador: Mary Aizawa Kato.

Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.

1. Língua portuguesa - Português escrito - Brasil. 2. Língua portuguesa - Português falado - Brasil. 3. Língua portuguesa - Gramática gerativa. 4. Aquisição da segunda linguagem. 5. Língua portuguesa - Clíticos. I. Lopes, Ruth Elisabeth Vasconcellos, 1960-. II. Kato, Mary Aizawa, 1934-. III. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. IV. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** The differences between what is spoken and what is written in Brazilian Portuguese : the acquisition of the se indeterminate and passive clitic

**Palavras-chave em inglês:**

Portuguese language - Written Portuguese - Brazil

Portuguese language - Spoken Portuguese - Brazil

Portuguese language - Generative grammar

Second language acquisition

Portuguese language - Clitics

**Área de concentração:** Linguística

**Titulação:** Mestra em Linguística

**Banca examinadora:**

Mary Aizawa Kato [Coorientador]

Maria Aparecida Corrêa Ribeiro Torres Morais

Sonia Maria Lazzarini Cyrino

**Data de defesa:** 19-12-2017

**Programa de Pós-Graduação:** Linguística



**BANCA EXAMINADORA:**

**Mary Aizawa Kato**

**Maria Aparecida Corrêa Ribeiro Torres Morais**

**Sonia Maria Lazzarini Cyrino**

**IEL/UNICAMP  
2017**

**Ata da defesa, com as respectivas assinaturas dos membros da banca, encontra-se no SIGA - Sistema de Gestão Acadêmica.**

*A todos que pensaram em desistir.*

## AGRADECIMENTOS

Primeiramente, à CAPES pelo financiamento parcial desta pesquisa, sem o qual não seria possível fazê-la.

À minha orientadora, professora Ruth, pelas orientações, correções e incentivos. Por ter acreditado em mim e em meu trabalho, e, principalmente, pelo acolhimento.

À minha co-orientadora, professora Mary Kato, por todo o aprendizado que me proporcionou.

À banca examinadora, tanto de defesa quanto da qualificação, professora Sônia e professora Cida, pelas considerações feitas e pela leitura atenta do trabalho.

Ao professor Emílio pelos momentos de ajuda com os dados, pelas conversas bem-humoradas e as aulas motivadoras.

À professora Charlotte, por ter acompanhado meu percurso pela pós-graduação desde às disciplinas, à orientação do PED e às apresentações de trabalhos.

Aos funcionários do IEL, que mantem o ambiente que se tornou minha segunda casa tão harmônico todos os dias. Gostaria de agradecer especialmente aos funcionários da Secretaria de Pós-Graduação pela solicitude ao atenderem minhas dúvidas e aos funcionários da Biblioteca Antônio Cândido pela receptividade.

Aos colegas do ForMa: Antônio, pelo carinho paternal e alto astral habitual; Willi, pelas conversas e ajudas estatísticas; Fernanda, pelos comentários ao trabalho no SETA e tudo o que veio depois disso; Carla, pelas orientações e fotos de gatos para animar, Harley, pelos grupos de estudos que tanto ajudaram em minha formação; Letícia, pelos momentos de aprendizado compartilhados e Ruan, pelas conversas divertidas. Todos os momentos que passamos juntos foram muito construtivos!

À minha amiga e colega de pós-graduação Aline pelas conversas, desabafos e, principalmente, pelo companheirismo nas viagens. Obrigada por trilhar esses sete anos de IEL comigo e por me mostrar que eu não estava sozinha!

A todos os colegas da pós-graduação que juntos, entre conversas no corredor e cafés, dividiram comigo a vivência acadêmica, em especial, Paulo, Jaque, Arnaldo, Helder, Julia e Ricardo!

À minha mãe por estar ao meu lado, mesmo à distância, nos momentos mais difíceis e nunca ter deixado de acreditar em mim e no meu potencial. Eu te amo, dona Daisy!

Ao meu irmão pelo bom humor e pela ajuda com as revisões. É muito bom contar com você!

À minha família pelo apoio e compreensão nos momentos de ausência: meu pai, minha avó Thereza, meu avô Zeca, minha tia Divone e tia Cida (IM). Sem vocês, eu não teria chegado até aqui!

À minha melhor amiga Verônica por simplesmente tudo, pois não é possível enumerar tudo que motiva minha gratidão por você.

Ao meu melhor amigo Vital por ter feito sempre o melhor que pôde para me ajudar e ter me ensinado que é isso que devemos sempre fazer.

Às minhas amigas Larissa Werneck, Sarah, Virginia, Lívia e Letícia pelo companheirismo inigualável.

Às minhas colegas de casa Larissa Cezarino, Hannah e Cris por terem construído um lar para onde eu podia voltar quando precisava de acolhimento.

Às minhas amigas de infância Camila e Carol por terem acreditado no meu futuro desde muito cedo.

Ao Guilherme (Garga) que, nos momentos finais, me deu a mão e me ajudou.

Às minhas psicólogas Márcia e Helô pela ajuda nessa travessia.

E, finalmente, a todos aqueles que trilharam esse percurso comigo.

*“O se ainda fará correr muita tinta.” (GALVES, 1986)*

## RESUMO

A fim de explorar o tema da distância entre o que se fala e o que se escreve no Português Brasileiro (PB), Kato, Cyrino e Correa (2009) propõem que a instituição escolar recupera determinados traços estilísticos da gramática portuguesa do século XIX e os postula como gramaticais, sendo um desses fósseis gramaticais o clítico *se* nas funções de indeterminador e de apassivador. Diante disso, o presente trabalho tem como objetivo investigar como se dá a aquisição desse clítico, principalmente por falantes letrados do PB. Pretende-se averiguar a hipótese de que a aquisição desse clítico para o PB se comporta tal como um item de uma língua estrangeira (L2), conforme proposto por Kato (2005). Para cumprir tal objetivo, foi traçado um breve panorama sobre os estudos realizados até o momento sobre a aquisição de L2 bem como sobre a natureza do *se* e seu percurso diacrônico. Esse último foi realizado com base nos estudos de Nunes (1990), no qual o autor estabelece o estatuto teórico das construções que envolvem o *se*, apassivador e indeterminador, dentro do quadro da Teoria da Regência e Ligação desenvolvida a partir de Chomsky (1981). Frente a dados diacrônicos do PB, o autor registrou alterações relativas principalmente à supressão do clítico em sentenças finitas e à sua inserção em sentenças infinitivas, além do processo de reanálise sintática que permitiu que o apassivador deixasse de receber a interpretação passiva para ser interpretado apenas como indeterminado. Com efeito, após a observação e a análise de um *corpora* composto por cinco bancos de dados, cada um representativo de uma fase da aquisição da escrita, viu-se que a função indeterminadora de *se* tomou espaço na língua e parece ser preferida em relação à forma apassivadora independente do material em que a sentença é produzida. Do mesmo modo, confirmou-se a hipótese de que a aquisição de *se* indeterminador e de *se* apassivador dependem da escolarização. Em bancos de dados representativos da fala em período de aquisição e dos primeiros anos escolares, sua presença é inexpressiva, levando em torno de 12 anos de escolarização para que se faça significativa. Atestou-se, portanto, que o conhecimento interno de um falante escolarizado do Português Brasileiro envolve duas gramáticas, tal qual Kato (2000) propôs: uma interna adquirida naturalmente, com inovações, e uma periférica adquirida na escola, conservadora e pautada nas normas da Gramática Tradicional, que ainda prevê o uso do *se* e condena outras formas de realizar a indeterminação. Essa última gramática é acionada em situações de fala monitorada ou na escrita que, por essa razão, não apresenta muitas inovações depois de adquirida a gramática escolar.

Palavras-chave: Aquisição de L2; Gramática Gerativa; clítico *se*.

## ABSTRACT

Kato, Cyrino & Correa (2009), discussing the distance between written and spoken Brazilian Portuguese (BP), show that the school system tends to recover certain stylistic features of the 19th Century Portuguese grammar. Those are treated as grammatical fossils by the authors. Amongst them the *se* indeterminate and passive clitic is to be found.

This thesis examines the acquisition of such clitic, mainly by highly educated speakers of BP. One of its goal is to see to what extent the *se* clitic is acquired as a second language (L2) grammar, as Kato (2005) has proposed. To achieve that, we examine different proposals for L2 acquisition and also the diachronic changes involving the clitic, focusing on its present syntactic status. For the diachronic picture we based our work on Nunes (1990), which examined the indeterminate and passive *se* constructions within the Government and Binding Theory framework (Chomsky, 1981). Nunes shows that the clitic disappears from finite sentences, while it is introduced in non-finite ones. He also shows that the passive *se* is reanalyzed, receiving an indeterminate interpretation.

Our results support his proposal. We have analyzed five different corpora, one for each stage in the acquisition of written BP. Our results show that the indeterminate interpretation took over the passive one, becoming the preferred form. They also show that the acquisition of the *se* clitic is dependant on schooling. It is basically nonexistent in data from young toddlers acquiring BP, as well as from texts produced by children in their first years at school. It takes around 12 years of schooling for it to become productive. These results show that a highly educated speaker of BP will have two internal grammars, as proposed by Kato (2000): One acquired as a first language – the internal core grammar of the speaker – and a marked one acquired at school, very conservative and based on grammatical fossils of the language. The latter is used in formal spoken BP and in written texts; which, thus, will not provide an environment for innovation.

Key-words: L2 Acquisition, Generative Grammar; clitic *se*.

## LISTA DE IMAGENS

Figura 1 – Esquema do processo das influências do processo de aprendizagem da escrita para a fala (retirado de Kato (1986)).....	17
Figura 2 – Estratégias de indeterminação em PE (Retirada de Duarte, Kato & Barbosa (2001)).....	76
Figura 3 – Estratégias de indeterminação em PB (Retirada de Duarte, Kato & Barbosa (2001)).....	77
Quadro 1 – Síntese das principais hipóteses sobre aquisição de L2.....	41
Quadro 2 – Síntese dos principais estudos sobre <i>code-switching</i> .....	42
Quadro 3 – Síntese das principais ideias de Gramática Tradicional sobre o <i>se</i> .....	49
Quadro 4 – Síntese das principais ideias de Gramática Funcionalista sobre o <i>se</i> .....	54
Quadro 5 – Síntese das principais ideias de Gramática Gerativa sobre o <i>se</i> .....	68

## LISTA DE GRÁFICOS

Gráfico 1 – Quantidade de <i>se Sarney</i> por banco de dados.....	86
Gráfico 2 – Frequência de cada tipo de <i>se</i> encontrado por ano.....	87
Gráfico 3 – Quantidade de dados coletados em cada banco de dados por ano.....	86
Gráfico 4 – Animacidade do argumento interno por tipo de <i>se</i> .....	91
Gráfico 5 – Posição do argumento interno em relação ao verbo por tipo de <i>se</i> .....	92
Gráfico 6 – Concordância do argumento interno em relação à posição do argumento interno.....	92
Gráfico 7 – Posição de argumento interno em sentenças com <i>se</i> indeterminador.....	95
Gráfico 8 – Transitividade do verbo por tipo de <i>se</i> Quantidade de ocorrência de cada tipo de <i>se</i> por Banco de Dados.....	96
Gráfico 9 – Quantidade de <i>se Sarney</i> em oração preposicionadas e não preposicionadas.....	98
Gráfico 10 – Quantidade de <i>se Sarney</i> por posição do argumento interno.....	99
Gráfico 11 – Quantidade de <i>se Sarney</i> por transitividade verbal.....	100
Gráfico 12 - Posição de <i>se</i> por tipo de <i>se</i> .....	101
Gráfico 13 – Quantidade de ocorrência de cada tipo de <i>se</i> por Banco de Dados.....	104

## LISTA DE TABELAS

Tabela 1 – Quantidade de <i>se</i> em cada banco de dados.....	81
Tabela 2 – Relação entre animacidade do argumento interno e tipo de <i>se</i> .....	89
Tabela 3 – Posição do argumento interno em relação ao verbo.....	90
Tabela 4 – Concordância em relação à posição do argumento interno.....	91
Tabela 5 – Concordância entre verbo e argumento interno por banco de dados.....	93
Tabela 6 – Posição do argumento interno por animacidade do argumento interno.....	94
Tabela 7 – Posição do argumento interno por tempo verbal.....	95
Tabela 8 - Transitividade do verbo em relação ao tipo de <i>se</i> .....	96
Tabela 9 – Tempo verbal por tipo de <i>se</i> .....	97
Tabela 10 - Posição de <i>se</i> por tipo de <i>se</i> .....	100
Tabela 11 - Posição de <i>se</i> por banco de dados.....	101

## LISTA DE ABREVIACOES

AGR	Sintagma de Concordncia
CS	Code-switching
DP	Sintagma Determinante
EPP	Princpio de Projeo Extendida (Extended Projected Principle)
GU	Gramtica Universal
IP	Sintagma de Flexo
L1	Lngua Materna
L2	Lngua Estrangeira
LF	Forma Lgica
MDG	Minimal Default Grammar
MLF	Matrix Language Frame
PB	Portugus Brasileiro
PC1	Portugus Clssico
PE	Portugus Europeu
PF	Forma Fonolgica
PS	Princpio do Subconjunto
SN	Sintagma Nominal
SPEC	Especificador
TP	Tense Phrase
vP	Voice Phrase
VP	Sintagma Verbal

## SUMÁRIO

Introdução.....	16
Objetivos.....	17
Metodologia.....	18
Hipóteses .....	19
<b>Capítulo 1 – O problema da aprendizagem do <i>se</i> por falantes do PB .....</b>	<b>20</b>
1.1 Um caso de aprendizagem em contexto de disparidade entre o que se fala e o que se escreve no PB .....	20
1.2 Variação e mudança.....	22
1.3 A aprendizagem do <i>se</i> via instrução como aquisição de L2.....	32
1.4. Concluindo o capítulo.....	41
<b>Capítulo 2 – O Estatuto do <i>se</i> indeterminador e apassivador .....</b>	<b>43</b>
2.1.1 O Índice de Indeterminação do Sujeito .....	43
2.1.2 A Partícula Apassivadora .....	46
2.2. Na Gramática Funcionalista .....	49
2.3. Na Gramática Gerativa .....	55
2.3.1 Teorias sobre outras línguas românicas.....	55
2.3.2 Teorias sobre o PB.....	62
2.3.4 Os dados de mudança .....	69
2.4. Concluindo o capítulo.....	79
<b>Capítulo 3 – Análise dos Dados .....</b>	<b>80</b>
3.1. Coleta e codificação dos dados.....	80
3.2. Distribuição dos dados .....	82
3.2.1 Variáveis Externas .....	82
3.2.2. Variáveis Internas .....	89
4. Discussão .....	104
<b>CONSIDERAÇÕES FINAIS.....</b>	<b>108</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>112</b>

## Introdução

Embora o mito da transparência entre a fala e a escrita ainda esteja presente nas escolas brasileiras, a distância entre o que se fala e o que se escreve no português do Brasil tem sido uma temática importante dentro dos estudos linguísticos contemporâneos. Muitos pesquisadores, não apenas brasileiros, têm considerado essa distância a fim de tentar determinar se as diferenças entre variantes linguísticas de prestígio – mais próximas à escrita e, portanto, mais conservadoras - e as vernaculares – mais distantes da escrita e mais inovadoras - pertencem a gramáticas diferentes ou a uma única gramática (CORREA, 1998).

Há quase duas décadas atrás, Kato (1999b), quando questionada pelo Banco Mundial sobre os problemas no ensino de língua materna no Brasil, afirmou que

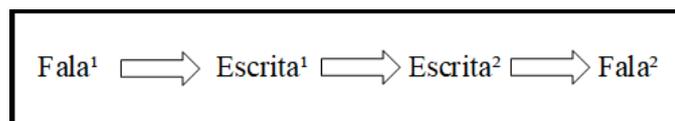
O problema principal no ensino de português é a distância entre a língua falada, vernacular, pré-letramento, e a língua escrita institucionalizada, que a escola visa como meta da aprendizagem. Em todas as línguas se observam diferenças entre a fala e a escrita, mas no português do Brasil (PB) existe um fosso de tal ordem que para a criança aprender a ler e a escrever, ela enfrenta a tarefa da aprendizagem da língua estrangeira. (KATO, 1999b, p. 1)

De acordo com Kato (2000), é na fala que se concentram as inovações linguísticas e a escrita daqueles que ainda estão aprendendo essa tecnologia é pouco marcada por essas inovações. A escrita daqueles que já estão familiarizados com essa tecnologia restringe suas inovações ao âmbito lexical e, dessa forma, a sua sintaxe se mantém conservadora.

Ao fim do processo de aprendizagem da escrita e já totalmente imersos no mundo das letras, os adultos letrados passam a apresentar também em suas falas marcas essencialmente da escrita. Uma dessas marcas é a presença do clítico *se* em sentenças impessoais, e sua localização enclítica, como nas sentenças abaixo:

- (1) Vive-se mal neste país.
- (2) Por meio do rádio difundiram-se as principais ideias sobre o assunto.

Para ilustrar as influências do processo de aprendizagem da escrita na fala, Kato (1986) propõe o seguinte esquema, apresentado na figura 1:



**Figura 1 – Esquema do processo das influências do processo de aprendizagem da escrita para a fala (retirado de Kato (1986))**

Nesse esquema, chama-se de Fala<sup>1</sup> àquela correspondente ao período pré-escrita, caracterizada como inconsciente e inovadora. A escrita<sup>1</sup>, consciente e inovadora, que viria logo após esse período, seria, então, muito próxima de uma transcrição da fala. O conservadorismo da escrita apareceria somente na Escrita<sup>2</sup> que, já sob o efeito da instrução escolar, é neutra em relação à fala. Chegando ao ponto que me interessa neste trabalho, esse esquema apresenta a Fala<sup>2</sup>. Conservadora e essencialmente filtrada pela tecnologia da escrita, essa modalidade de fala pode ser consciente ou não, deixando as inovações gramaticais à Fala<sup>1</sup> e à Escrita<sup>1</sup>. Vê-se, portanto, apenas as inovações lexicais na Escrita<sup>2</sup> e na Fala<sup>2</sup>, de forma que “a tecnologia da escrita reprime as inovações e faz o falante voltar a formas já eliminadas, ou no limiar do desaparecimento” (KATO, 1999a, p. 204).

Kato (2005a) afirma que a distância entre a Fala<sup>1</sup> e a Escrita<sup>2</sup> é ainda mais problemática pela ausência de estudos comparativos entre o conhecimento linguístico que a criança leva à escola e o conhecimento dos letrados contemporâneos para auxiliar a escola em sua tarefa de letramento. Logo, é nessa carência que se justifica esta pesquisa.

### **Objetivos**

O objetivo deste trabalho é, portanto, investigar as influências da aprendizagem da escrita na fala, ou seja, o percurso realizado pelos falantes do PB da Fala<sup>1</sup> à Fala<sup>2</sup>, portanto, analisar dados de um *corpus* que contemplará as quatro etapas do esquema acima, a fim de responder às seguintes perguntas:

- 1) Como se dá a aquisição do clítico *se* para falantes de PB?
- 2) Como se comporta o clítico *se* para as funções de apassivador e indeterminador antes, durante e depois da aprendizagem da escrita?

- 3) Qual a relação desses resultados com as ideias de Kato (1999a; 2005a) sobre a aquisição da gramática escolar?

### **Metodologia**

O *corpus* utilizado para as análises neste trabalho será composto por dados provenientes de quatro bases distintas em que cada qual corresponderá a um estágio do esquema.

Para ilustrar a Fala<sup>1</sup>, foram utilizados dados cedidos pelo CEDAE que fazem parte da coleção Projeto Aquisição da Linguagem Oral. Segundo consta no documento que introduz a finalidade do projeto, trata-se de “gravações de áudio e videoteipe com doze crianças, classe média, filhas de pais universitários”. Sob a coordenação da professora Cláudia Lemos, do IEL/UNICAMP, “o projeto teve como objetivo descrever e analisar o processo de aquisição do português como primeira língua em crianças de um a cinco anos”.

Os dados para a Escrita<sup>1</sup> serão retirados do projeto intitulado "A relevância teórica dos dados singulares no processo de aquisição da linguagem escrita", cujo desenvolvimento se deu no IEL/UNICAMP entre os anos de 1992 a 2004, como Projeto Integrado do CNPq, coordenado pela Prof<sup>a</sup>. Dra. Maria Bernadete Marques Abaurre e tendo como pesquisadoras principais a Prof<sup>a</sup>. Dra. Raquel Salek Fiad e a Prof<sup>a</sup>. Dra. Maria Laura Trindade Mayrink Sabinson. Destaco, porém, que mesmo extinta a modalidade Projeto Integrado, as pesquisadoras continuaram realizando projetos individuais utilizando o Banco de Dados reunido durante o Projeto Integrado.

A verificação da Escrita<sup>2</sup> será embasada no banco de dados da Comissão Permanente para Vestibulares sobre as provas de redação do vestibular da Unicamp. Serão utilizadas apenas aquelas consideradas pela banca corretora como acima da média uma vez que o vestibular preza pelo domínio da norma padrão. Quanto à Fala<sup>2</sup>, serão tomadas como *corpus* alguns diálogos pertencentes ao banco de dados do Projeto NURC/SP (Norma Urbana Culta – São Paulo).

A fim de representar o processo de aprendizagem da escrita com mais consistência, também foram coletados dados do caderno de política Folha de S. Paulo das décadas de 1980 e 2000 disponibilizados no arquivo digital do jornal.

### **Hipóteses**

A hipótese principal desta investigação é a de que o clítico *se* como indeterminador e apassivador se comporta tal qual uma perda linguística e sua aquisição natural pelas crianças falantes de PB é dificultada uma vez que outras formas de indeterminação do papel temático de sujeito foram incorporadas à gramática interna desses falantes. Dessa forma, a aquisição dessa marca só seria possível por meio do processo de escolarização. O *se* em estruturas passivas seria ainda mais difícil de ser adquirido já que, mesmo no ambiente escolar, sua aparição tem se restringido por processos internos da língua.

Essas hipóteses fundamentam-se também na proposta de Kato que afirma que a língua escrita ensinada na escola é tão distante daquela falada pelos falantes de PB que essa se assemelha ao resultado da aprendizagem de uma L2. Assim, nosso objeto de estudo faria parte da aprendizagem de uma L2.

Um dos alicerces desta pesquisa também é o trabalho de Nunes (1990), no qual o autor considera que houve no PB um processo de reanálise sintática em que o sujeito expletivo de construções com *se* apassivador foi reinterpretado como um pronome nulo referencial e que as construções com *se* indeterminador se tornaram mais frequentes que as construções com *se* apassivador, o que gerou outras mudanças linguísticas envolvendo *se* indeterminador no português brasileiro, como o seu apagamento em sentenças finitas e sua inserção em sentenças infinitas.

Essa dissertação está organizada da seguinte maneira: o primeiro capítulo será destinado à revisão bibliográfica dos trabalhos teóricos referentes aos processos envolvidos na aquisição L2; o segundo capítulo será destinado, por sua vez, à apresentação de pesquisas sobre o estatuto teórico e sintático do clítico *se*; o terceiro capítulo abrangerá o resultado da coleta de dados realizada e a discussão do que foi observado; e o quarto capítulo, finalmente, trará as considerações finais sobre o trabalho.

## **Capítulo 1 – O problema da aprendizagem do *se* por falantes do PB**

Este primeiro capítulo será dedicado à investigação da problemática da aquisição da gramática escolar como uma L2. Primeiramente, serão apresentados os estudos seminais de Kato, Cyrino e Correa (2009) e de Kato (1996, 2017) que colocam essa questão em relação à distância que se tem no PB entre língua falada e língua escrita. Em seguida, essa aquisição será analisada sob a perspectiva variacionista e, por fim, uma breve revisão das teorias sobre aquisição de L2 será traçada.

### **1.1 Um caso de aprendizagem em contexto de disparidade entre o que se fala e o que se escreve no PB**

Em trabalho mais recente, Kato, Cyrino e Correa (2009) revisitam o tema da disparidade entre o que se fala e o que se escreve no PB, dessa vez, olhando para a perda dos clíticos de terceira pessoa no PB. Após analisar dados de fala de estudantes de diversos anos escolares, elas verificaram que, tal qual defenderemos para o *se*, os clíticos acusativos de terceira pessoa não aparecem na fala até o último ano escolar. Esses dados constituem uma forte evidência, então, para a tese defendida pelas autoras: o uso de clíticos no PB depende principalmente do sucesso do sistema educacional, especialmente na escrita. Esta pesquisa corrobora essa ideia e, uma vez que as autoras se dedicam apenas ao estudo dos clíticos acusativos de terceira pessoa e não incluem o clítico *se* em sua pesquisa, também complementa esse trabalho em termos de objetivos e metodologia.

As autoras retomam a ideia de que há, na gramática de sujeitos brasileiros letrados, formas em competição e se perguntam qual a natureza do conhecimento linguístico para esses indivíduos. Uma vez que a língua-I de um falante, segundo a proposta Chomskyana, pode ser definida como “o conhecimento linguístico de uma pessoa, aquilo que está presente na sua mente e lhe permite usar uma língua-E para produzir e compreender palavras, sintagmas, frases e discursos” (KENEDY, 2013, p. 34), Kato, Cyrino e Correa (2009) se questionam como esse conhecimento é construído em falantes cultos já que o próprio conceito não dissocia o conhecimento linguístico de uma criança antes e depois da escolarização. Essa questão se evidencia quando se trata de adultos letrados cujo processo de escolarização durou mais de doze anos. Para além de entender questões relacionadas à aquisição de linguagem, as autoras também frisam a

importância dessa discussão para que se olhe mais criteriosamente para *corpora* escritos.

Kato (2017) sumariza as questões a respeito da aquisição do clítico de terceira pessoa no PB e suas hipóteses podem ser levantadas também para a aquisição do *se* indeterminador e *se* apassivador. Ela retoma, primeiramente, o lugar que a variação linguística ocupa dentro dos estudos gerativistas. Apesar de Chomsky (2010 apud KATO, 2017) limitar a variação ao componente morfofonológico, a partir da teoria de Princípios e Parâmetros, os pesquisadores passaram a se ocupar não só das propriedades que são comuns a todas as línguas – os Princípios – mas também dos elementos que as diferenciavam – os Parâmetros. Esse novo olhar abriu, então, as portas para a investigação da variação e da mudança dentro de uma língua, entendida como uma mudança no valor de um ou mais parâmetros.

Partindo, então, das afirmações de Duarte (1993), a autora prossegue nos estudos sobre a aquisição de clíticos acusativos e reafirma as mudanças no paradigma de clíticos do PB em relação ao Português Europeu e ao Português Clássico: o pronome *tu* é substituído pelo *você* e, conseqüentemente, os clíticos nulos e as formas não clíticas tomam o lugar dos clíticos de terceira pessoa. Ao passo que os clíticos de primeira e de segunda pessoa se mantêm, Kato, Cyrino e Correa (2009) observam uma queda acentuada no uso dos clíticos de terceira pessoa no fim do século XVIII.

Diante disso, Kato (2017) propõe que, se a gramática nuclear da criança falante do PB não apresenta os clíticos de terceira pessoa, essa aquisição só é possível por meio da escolarização e, portanto, ela não altera o valor do parâmetro até que a aquisição de sua gramática nuclear tenha se completado. Isso é validado, por exemplo, com base no estudo de Correa (1991), no qual foram observadas duas situações em que a gramática nuclear adquirida via assentamento de parâmetro não é afetada pelo input via instrução: em primeiro lugar, o fato de que esses clíticos aparecerem somente a partir da 5ª série do primeiro grau, doravante Ensino Fundamental II; em segundo lugar, a ocorrência de clíticos de terceira pessoa na fala de analfabetos ter uma frequência muito semelhante à ocorrência na fala de crianças no início da escolarização.

A forma como os escolares trazem esses clíticos na escrita também permite que se faça asserções a respeito dessa aquisição. Frequentemente, aparecem formas que, apesar de possíveis na diacronia do português, são consideradas agora como “erros” ou

“hipercorreções” por divergirem do input via instrução. Se isso é possível, talvez não seja incoerente afirmar que a GU da criança é ativada nesses casos (KATO, 2017).

Isto posto, a autora se questiona em que medida essas formas perdidas são recuperadas. Para responder a essa questão, ela resgata outro resultado interessante do estudo realizado por Kato, Cyrino e Correa (2009) ao compararem dados de falantes de nível universitário e dados coletados em textos escritos, especialmente peças teatrais, do século XIX. Em ambos os *corpora*, o uso de clíticos era muito semelhante. Cabe destacar, para compreender melhor esse resultado que, de acordo com dados diacrônicos do PB e do PE, até o século XIX, a gramática do PB apresentava um paradigma de clíticos idêntico ao do PE.

Diante disso, Kato, Cyrino & Correa concluem que a escola seleciona certos fósseis da linguagem, ou seja, certas marcas estilísticas que já não fazem parte da gramática internalizada de um sujeito falante de PB, e os postula como gramaticais. Nas palavras de Kato (1996)

Embora poucos brasileiros façam uso dos clíticos de terceira pessoa na fala espontânea, não monitorada, seu uso na escrita ainda é a norma. O leitor encontrará, em meu texto, tudo aquilo que venho afirmando estar em processo de perda. Não só usamos clíticos, mas usamos ênclise e mesóclise. (KATO, 1996, p. 15)

Desse modo, a criança pré-escolar apresenta apenas formas determinadas pela aprendizagem por seleção, ao passo que, após o início da escolarização, é possível notar em sua fala a variação entre formas adquiridas por seleção e formas aprendidas via instrução (KATO, 2017).

O conhecimento linguístico de um falante letrado do PB é, portanto, constituído por uma gramática nuclear, refletida na sua oralidade, e de uma periferia verificada por meio da presença dos tais fósseis da escrita na reprodução de regras estilísticas adquiridas durante o processo de escolarização. (KATO, CORREA E CYRINO, 2009).

## **1.2 Variação e mudança**

Como vimos, Kato, Correa e Cyrino (2009) propõem que essa alternância de estilo se dá na forma de *code-switching* (doravante CS). Esse fenômeno pode ser conceituado tal qual Poplack (2001) o faz:

Code-switching (CS) refers to the mixing, by bilinguals (or multilinguals), of two or more languages in discourse, often with no change of interlocutor or topic. Such mixing may take place at any level of linguistic structure, but its occurrence within the confines of a single sentence, constituent or even word, has attracted most linguistic attention. (POPLACK, 2001, p. 1)

Esse fenômeno foi, por muito tempo, tratado como aleatório até ser estudado sob a óptica da pragmática e da sociolinguística. Somente estudos recentes, como o de Poplack (2001), o de Toribio (2017) e o de Riehl (2005) o consideram como gramaticalmente motivado e passível de ser estudado à luz de teorias formais.

Toribio (2017) faz uma revisão dos estudos realizados sobre CS. Ela o define como um fenômeno bilingue em que um mesmo evento de fala apresenta duas línguas. Para a autora, há pares de línguas regulares que propiciam as ocorrências. É o caso do espanhol e do inglês, como no exemplo (13) abaixo:

(3) Mi amigo tiene una casa bien grande so nos gusta irnos ahí y sus papás están *cool* para que nosotros tomemos, *so* ahí nos juntamos *most of the time*... mi familia también estaba contenta de mí ‘*cause* no hice *drop out*... Mi mamá sí se graduó de *high school* y mi papá también, pero de ahí ya no fueron al colegio ni nada, y luego pues me hicieron *influence* a ir al colegio para, *I guess*, ‘cabar eso para ellos. ¿Así se dice? *Do you get me?* (BULLOCK & TORIBIO, 2012 apud TORIBIO, 2017)

‘Meu amigo tem uma casa bem grande então gostamos de ir lá e seus pais não se importam que nós bebamos, então nos juntamos lá a maior parte do tempo... minha família também estava feliz comigo porque não eu não abandonei... Minha mãe se formou no colégio e meu pai também, mas não foram à faculdade nem nada, e então me fizeram ir à faculdade para, eu acho, acabar isso para eles. Assim que se diz? Você me entendeu?’

Da mesma, o francês e o árabe:

(4) C’est le fer qui donne... illi yi-ha:rib l’anémie. (TORIBIO, 2017)

‘It’s iron that gives, that fights anemia’.

É o ferro que dá, que combate a anemia.

A autora olha para dados como (4) para buscar evidências de combinações de propriedades formais que não são visíveis quando se atenta para uma língua separadamente. Essas evidências estão principalmente nas reformulações feitas pelos falantes bilíngues. Diante desses dados, ela retoma as ideias de Poplack (1980 apud TORIBIO, 2017), que defende que o CS acontece em pontos da estrutura em que as duas línguas não divergem, como mostra o exemplo (5) em que o CS é feito em pontos de divergência e, por essa razão, a estrutura se torna agramatical.

(5) \*Luego lo have all perfect/ \*Then they it tienen todo perfecto.  
(TORIBIO, 2017)

‘\*Luego lo tienen todo perfecto/ Then they have it all perfect’.

\*Logo o tem todo perfeito.

Em contrapartida, Myers-Scotton (1993, 2012 apud TORIBIO, 2017) defende a existência do Matrix Language Frame (MLF), um modelo segundo o qual, para que o CS ocorra, uma língua deve prover o quadro morfossintático ao qual elementos de outras línguas são incorporados. Na sentença (6), por exemplo, o verbo *es* identifica o espanhol como MLF:

(6) La Lupita es una gifted student. (TORIBIO, 2017)

‘Lupita is gifted student’.

Lupita é uma estudante talentosa.

Gardner-Chloros (2009 apud TORIBIO, 2017) argumentam que esses padrões de CS são específicos das comunidades em que o falante está inserido. Logo, para eles, CS é um processo que envolve múltiplos fatores, até mesmo externos, como o quadro social do falante. Sebba (1999, 2009 apud TORIBIO, 2017) identifica três maneiras de criar a congruência entre sistemas gramaticais: neutralização dos elementos, em que a variável é incorporada ao verbo, como em (7); harmonização das categorias gramaticais, tratados como equivalentes nas línguas, como em (8); e o comprometimento, que aparece quando uma língua parece violar regras gramaticais das suas línguas como em (9):

(7) *fazer o save.* (TORIBIO, 2017)

Do the save.

‘To save’.

Salvar.

(8) *Had le pain.* (TORIBIO, 2017)

This the brad.

‘This bread’

Este pão

(9) *Dak le trajet kulha.* (TORIBIO, 2017)

‘This whole journey’

Toda essa jornada

Sebba (1999, 2009 apud TORIBIO, 2017) enfim concorda com o ponto de vista de Chan (2009 apud TORIBIO, 2017), o qual propõe que CS não depende apenas de propriedades formais, mas que se trata de um fenômeno resultante de padrões dependentes de sociolinguística e fatores cognitivos.

Diante dessa revisão teórica, Toribio (2017) ressalta a necessidade da criação de um banco de dados mais robusto para os estudos contemporâneos.

Mais recentemente, Riehl (2005) advoga em favor da ideia de “palavragatilho”. Com base nos estudos de Clyne (2003 apud RIEHL, 2005), Riehl (2005) defende que há três categorias de itens lexicais que, embora não promovam CS efetivamente, facilitam o processo. São elas: nomes próprios, palavras homófonas e palavras que permitem transferências lexicais que são fonologicamente não integradas (ou integradas em um nível pequeno), exemplificados respectivamente em (10), (11):

(10) *Es war Mr Fred Burger, der wohnte da in Gnadenthal and he went out there one day and Mrs Roehr said to him: Wer sind denn die Männer do her?* (Clyne, 2003 apud RIEHL, 2005)

‘It was Mr. Fred Burger, he lived at Gnadenthal and he went out there one day and Mrs Roehr said to him: Who are all these men around here?’

Era o Sr. Fred Burger, ele morava em Gnadenthal e ele saiu um dia e a Sra. Roehr disse para ele: quem são todos esses homens por aqui?

(11) Keine Apfelsinen. Wir haben se gehabt but oh großes Feuer come [kam] throughand killed all the trees. (Clyne, 2003 apud RIEHL, 2005)

‘No oranges. We had them but oh big fire come through and killed all the trees.’

Sem laranjas. Nós as tínhamos, mas um grande fogo veio e matou todas as árvores.

A autora conclui que essas três categorias fazem parte de uma linguagem compartilhada entre as línguas em que esses itens estão, de alguma forma, conectados e dão margem, portanto, ao processo de code-switching. A autora destaca também que a atenção é um fator importante para o fenômeno.

Kato (1994), por sua vez, defende que, quanto mais distante são as línguas envolvidas, mais restrito o CS se limitará a itens isolados, como é o caso de falante de português e de japonês. É importante também ressaltar a diferença entre CS e *borrowing* (REYES, 1982 apud KATO, 1994). Ambos são processos de mixagem de códigos, mas o segundo envolve apenas a inserção de uma única palavra, enquanto o outro requer a interação de duas gramáticas e suas sintaxes distintas. Os exemplos (12) e (13) exemplificam estruturas em que há CS e os exemplo (14) e (15), *borrowing*:

(12) Os dois são uns amores, *very caring about their students*. (KATO, 1994)

‘Os dois são uns amores, muito preocupados com os alunos.’

(13) *For now* ele tá *working well* mesmo. (KATO, 1994)

‘Por agora ele tá trabalhando bem mesmo.’

(14) Vou para casa comer porque estou *starving*. (KATO, 1994)

‘Vou para casa comer porque estou esfomeado.’

(15) A cafeteria é horrível e a comida *disgusting*. (KATO, 1994)

‘A cafeteria é horrível e a comida asqueMarcelino.’

Embora a autora defenda que não é apenas a sintaxe que condiciona o processo mas também a prosódia, ela caminha no mesmo sentido que Sankoff&Poplack (1981 apud KATO, 1994), ao afirmar que são os itens funcionais que determinam a estrutura sobre a qual inserções de uma outra língua são feitas. Essa hipótese sugere que a língua matriz não é dada apenas por INFL, mas também pelo Determinante, ambos sintagmas funcionais. Isso permite afirmar que o mesmo ocorre em Comp, como nos exemplos abaixo em que, toda a sentença está em inglês exceto Comp:

(16) Ele foi muito legal de novo e disse que *he'd be happy*. (KATO, 1994)

‘Ele foi muito legal de novo e disse que ele ficaria feliz.’

(17) Dizem aqui que April showers bring May flowers! (KATO, 1994)

‘Dizem aqui que as chuvas de abril trazem as flores de maio!’

O que se observa, então, é que existe uma forte restrição que decorre da distinção entre palavras lexicais e funcionais. A linguagem da língua matriz é identificada como a que fornece o esqueleto funcional do idioma que será integrado. O esqueleto funcional, visto na perspectiva da estrutura Princípios e Parâmetros, contém: Comp, Tense, Agreement, Determiners, Negation e Morfemas de casos. Já as categorias lexicais incluem: substantivos, verbos, adjetivos, advérbios. As preposições parecem constituir uma categoria mista (KATO, 1994).

Consequentemente, sempre que aparece uma palavra ou morfema funcional em uma das línguas, abre a possibilidade de CS. Assim, em (18) é o complemento que autoriza a frase seguinte a estar em inglês e, dessa forma, o CS possa ocorrer repetidamente:

(18) I think that Garfield está apaixonado and that Paquita também está. (KATO, 1994).

‘Eu acho que o Garfield está apaixonada e que Paquita também está.’

Retomando a situação da aquisição de clíticos no PB, apesar de a morfologia das partículas ser aprendida via instrução pelo escolar, a prosódia e a sintaxe não permitem que a aquisição se complete, propiciando o CS (KATO, 2017).

Outro fator que interfere no processo é o grau de fluência do falante nas línguas envolvidas. Quando esse grau é mais balanceado, o falante tende a realizar combinações no interior das sentenças, como em (19), mas quando esse grau é menos balanceado, observam-se mais combinações inter-sentenciais, como em (20) (POPLACK, 1982 apud KATO, 1994):

- (19) Nas outras áreas até pode fazer diferença, mas nossa, *I'm sorry*, mas não faz mesmo. (KATO, 1994).  
 ‘Nas outras áreas até pode fazer diferença, mas nossa, sinto muito, mas não faz mesmo.’
- (20) *Anyway* eu tô tocando com ele. (KATO, 1994).  
 ‘De qualquer forma eu tô tocando com ele.’

A permeabilidade das gramáticas, ou seja, o que é transferido de uma língua para outra e em qual medida se torna, então, um ponto de discussão entre autores que investigam principalmente o bilinguismo. Kato (2001) retoma essa questão e levanta duas dessas visões diametralmente opostas: Weinreich (1953 apud KATO, 2001) e Bickerton (1981 apud KATO, 2001) que defendem que elementos da gramática da L2 permeiam o discurso de um bilíngue, mas não são incorporadas na língua como um código; e Silva-Corvalan (1993 apud KATO, 2001) que afirma que qualquer traço linguístico pode ser transferido de uma língua para outra desde que eles sejam compatíveis em estrutura.

Diante disso, Kato (2001) assume a proposta de Silva-Corvalan (1993 apud KATO, 2001) e afirma que o CS encontra condições favoráveis para emergir nas seguintes situações:

- a) Aprendizagem formal de uma língua estrangeira por uma criança depois de sua alfabetização na língua materna ter sido concluída;
- b) Aquisição tardia de uma L2 por um adulto imigrante;
- c) Aquisição de uma segunda língua no início da escolarização;
- d) Aquisição simultânea de mais de uma língua.

Todas as situações, exceto a primeira, envolvem um contexto de imersão na L2. Isso indica que CS seria, de alguma forma, culturalmente motivada. Por outro lado, vê-se que a escolarização representa um fator determinante, já que em situações como (c), nota-se que primeira gramática fica fossilizada em detrimento da adquirida posteriormente pelo falante (KATO, 1996). Essa situação seria, então, a que mais se assemelharia à situação aqui estudada: a criança chega à escola com sua gramática nuclear e passa a conviver com outra gramática na escola, na qual será alfabetizada. No entanto, ao invés de se tornar mais fluente na gramática escolar, o falante alterna entre os dois códigos.

Nesse ponto, cabe destacar que o processo de CS visível em crianças que estão adquirindo uma língua em contexto de mudança linguística, como é o caso do PB, é diferente do CS que ocorre entre a gramática nuclear do PB e a gramática escolar ao fim do processo de escolarização. A aquisição da L1 em contextos de mudança linguística compreende situações em que valores distintos de um mesmo parâmetro parecem estar presentes no “input” recebido pela criança durante a aquisição e, portanto, trata-se de duas gramáticas distintas, mas internamente coerentes. Não se trata de bilinguismo, mas de um processo de mudança linguística em que o agente é a criança. Igualmente, a variação, primeira fase da mudança, propicia contextos de reanálise de determinadas formas, em que uma delas passa a ser agramatical e deixa de aparecer no “input” recebido pela criança (LIGHTFOOT, 1991; ROBERTS, 1993 apud KATO, 1996). Esse é o caso de sentenças com *se* passivador que será apresentado a partir dos estudos de Nunes (1990) no próximo capítulo.

Conquanto Lightfoot chame esse tipo de mudança de catastrófica, o que se observa em termos de dados escritos, em tempo real, é que ela leva muito tempo para ser totalmente implementada. Isso se deve, em grande parte, aos agentes conservadores, que impedem a língua de mostrar a sua verdadeira cara. Por agentes conservadores, entendo toda a classe dos letrados, cujos membros, com maior ou menor submissão, seguem as normas institucionalizadas. A criança, com pais incapazes ou sem tempo de fornecer o “input” da linguagem escrita, enfrenta, na escola, uma situação de aprendizagem de uma língua estrangeira (KATO, 1996).

Sob a perspectiva variacionista, Kroch (2003) se debruça sobre os estudos contemporâneos a respeito de mudanças sintáticas e defende que, uma vez que esses

processos não ocorrem sem uma fase em que formas diferentes apareçam em um mesmo contexto, o fenômeno da competição de gramáticas pode explicar tais fases.

De acordo com Kroch (2003), autor que argumenta contra uma visão de mudança linguística inerente à sintaxe, durante o período de mudança linguística em curso, seja ela por fatores linguísticos ou extralinguísticos, os falantes da língua que estão passando pelo processo tendem a apresentar as formas antigas concomitantemente às formas novas. Trata-se, para ele, de duas gramáticas diferentes competindo na fala daquele sujeito. A principal evidência para se considerar essa proposta como válida é o fato de que as mudanças, de fato, ocorrem gradualmente, o que representa um desafio para a teoria gerativa. No entanto, a teoria encontra uma solução. Nas palavras do autor:

Dados os pressupostos da gramática gerativa, a variação em sintaxe que corresponde à fixação de oposições para parâmetros sintáticos deve refletir a co-presença concomitante num falante ou numa comunidade de fala de gramáticas mutuamente incompatíveis. Isso não é uma questão empírica, mas uma matéria de definição do conceito de parâmetro dentro da teoria. Recentemente, Kroch e seus colaboradores (cf. referências acima) descobriram evidência para essa consequência de pressupostos gerativistas. Eles mostraram, em vários estudos de caso, que a razão de mudança em diferentes contextos superficiais refletindo uma única mudança paramétrica era a mesma. Esse resultado, conhecido como o Efeito de Taxa Constante (Constant Rate Effect), é o que se espera se um único parâmetro gramatical é envolvido na mudança e a mistura de dois parâmetros opostos está lentamente mudando através do tempo numa dada comunidade de fala. (KROCH, 2003, p. 30)

Como nem todos os falantes estarão expostos aos mesmos dados para que o parâmetro correspondente à determinada mudança sintática seja refixado de maneira homogênea, o que se observa, então, é a criação de uma população mista na qual alguns falantes têm a fixação paramétrica antiga e outros, a nova. A escolha da forma que corresponde à mais bem formada, no entanto, não é objeto de investigação da teoria gerativa uma vez que corresponde ao domínio da *performance*.

O autor também considera a problemática da diferença entre o que se fala e o que se escreve na escola ao fim de seu trabalho, afirmando que a competição em diglossia sintática representa, neste caso, uma oposição entre uma língua literária conservadora e um vernáculo inovador. Comparando, então, o processo de competição de gramáticas ao bilinguismo, Kroch (2003) faz considerações acerca da estabilidade nesses dois processos. Se, por um lado afirma-se que o bilinguismo é instável por uma língua se sobrepor a outra em determinados momentos, quando se trata de diglossia sintática essa ideia precisa ser revista. Caso houvesse uma gramática mais “interna” que

outra, seria possível observar, nas comunidades linguísticas em que esse processo se dá, uma convergência para o uso de determinadas formas ao longo do tempo. Essa assimetria só seria alcançada, de acordo com o autor, quando uma gramática é adquirida durante a infância no vernáculo e a outra, referente à uma língua de mais prestígio, mais tardiamente. Dessa forma, o vernáculo operaria com uma força mais intensa para o falante e, portanto, seria o dito vencedor na competição.

Para esclarecer a distinção entre a visão variacionista e visão dos gerativistas, Kato (1992) recupera a noção de estilo, utilizada por gramáticos do século passado para definir o lugar da variação linguística nos estudos formais. Labov (1972 apud KATO, 1992), por exemplo, defende que o conceito de variante está ligado a estilo uma vez que a diferença entre o discurso cuidado e o discurso informal é estilística. Assim sendo, a autora propõe que

Enquanto os variacionistas focalizam o estilo como um fenômeno externo, entre a forma produzida (Língua-E) e o contexto, para os gerativistas, estilo se define como opções internas paramétricas do falante. Variação sintática para os primeiros constitui um fenômeno intra Língua-E enquanto para os últimos trata-se de um processo intra Língua-I. Estes tentam desvendar o mecanismo que permite tal 'code-switch' enquanto aqueles procuram enxergar uma ordem e uma sistematicidade na heterogeneidade do produto. (KATO, 1992, p. 134)

Diante disso, esta pesquisa caminha no sentido de unir as duas análises para cercar o fenômeno da aquisição da gramática escolar em toda a sua dimensão e contribuir linguisticamente, como afirma Kato (2013), para determinar com precisão as diferenças entre gramática nuclear que a criança traz para a escola e a língua-I dos sujeitos letrados, que nem sempre estão em consonância com o que prescreve a Gramática Tradicional. Aliar a investigação formal às teorias variacionistas, ademais, é, de acordo com Kato&Tarallo (1989), uma forma de alcançar um poder explanatório satisfatório para o estudo de fenômenos linguísticos.

Por fim, Kato (2011, 2017) afirma que o falante letrado se comporta como um bilíngue, ao realizar CS entre sua gramática nuclear e sua gramática adquirida na escola, sendo a variação o resultado desse processo. Logo, o conhecimento linguístico de um falante letrado do PB se distingue daquele de um falante que não foi tão afetado pela escolarização e pelo letramento escolar (KATO, 2009). Esse conhecimento se assemelha, então, a uma aquisição de L2. Cabe ressaltar que Kato (2005a) pondera também a possibilidade de esse conhecimento estar associado à recuperação do

conhecimento gramatical de um falante de uma época passada do PB ou, devido às políticas linguísticas que privilegiam a unidade entre o PB e o PE, esse saber seria pautado no conhecimento de um falante de PE. Contudo, essas hipóteses não podem se confirmar por duas razões: apenas alguns fósseis de gramáticas passadas são recuperados e obras escritas em PE carecem, muitas vezes, de tradução para o PB. Quando à segunda hipótese, a autora também reafirma a diferença entre o uso de clíticos em PE e PB, já que em PE, os clíticos de terceira pessoa ainda são muitos frequentes na fala.

Para estudar a aquisição desse conhecimento e reafirmar suas hipóteses, Kato (2005a), enumera algumas semelhanças entre a aquisição de L2 e a aquisição da escrita. Primeiramente, as duas aquisições não são biologicamente determinadas, mas socialmente motivadas. De igual modo, ambas começam após o chamado período crítico para aquisição e, nas duas situações, o falante tem consciência de sua aquisição. Além disso, ambos os processos são vagarosos e envolvem diferenças individuais. Resta, então, o seguinte questionamento: como se dá o processo de aquisição dessa L2 e no que ele difere da aquisição de uma L1? No item a seguir, serão apresentadas teorias e estudos sobre essa aquisição.

### **1.3 A aprendizagem do *se* via instrução como aquisição de L2**

De acordo com Kato (2005b), até o advento do modelo de Princípios e Parâmetros, acreditava-se que a aquisição de L2 era um processo completamente distinto de da aquisição de L1. Enquanto os teóricos estruturalistas investigavam os erros na aprendizagem; outros como Krashen (1982), Bley-Vroman (1989) e Espstein (1996)<sup>1</sup> distinguiam aquisição de aprendizagem, sendo os dois últimos adeptos da ideia de que essa aprendizagem de L2 se basearia em operações cognitivas do tipo Piagetiano. Xavier (2006) e Marcelino (2007) traçam o percurso das teorias sobre aquisição de L2 e, a partir de seus estudos, serão apresentados a seguir os principais autores que discorreram sobre diferentes perspectivas dessas teorias.

Inicialmente, cabe elucidar os conceitos de Gramática Universal (doravante GU), Língua-E e Língua-I de Chomsky (1981,1986). Enquanto pilares para o

---

<sup>1</sup> (apud KATO, 2005b)

entendimento da teoria gerativista, esses conceitos recortam o estudo da língua para definir qual o objeto de estudo dos pesquisadores gerativistas. Tendo em vista que, ao fim da aquisição da língua materna, a criança faz uso de recursos linguísticos que não estavam disponíveis no estímulo dado, o chamado input, a teoria da Gramática Universal (doravante GU), surge em resposta afirmando que há um estágio inicial de aquisição, a própria GU, que contém todas as propriedades das línguas naturais e que permite que qualquer língua seja aprendida pela criança. Já a Língua-E pode ser definida como a língua como idioma, um objeto social, político e geográfico. O foco chomskyano é, pois, a Língua-I: individual, intensional e interna, ela representa o conhecimento linguística que um falante tem de sua língua. Assim, ela contém os chamados Princípios – conjunto de características que não variam de uma língua para outra – e os Parâmetros – propriedades marcadas com valores (+) ou (-) que definem a variação entre as línguas. Quando todos os valores dos Princípios são definidos a partir da GU, Chomsky (1986) afirma que a gramática nuclear do falante está definida. Essa gramática nuclear pode vir acompanhada de uma periferia marcada (CHOMSKY, 1981), em que os fenômenos de mudança e inovação são inseridos. Assim sendo, a criança chegaria à escola com sua gramática nuclear estabelecida, mas só aquelas que são expostas massivamente a dados da gramática da escrita, terão em sua periferia os chamados fósseis linguísticos (KATO, 2005a). Situações desse tipo serão analisadas no capítulo dedicado aos dados colhidos nesta pesquisa.

Se para a aquisição de L1 tem-se como estado inicial a GU, para a aquisição de L2 esse estado inicial ainda gera controvérsias entre os pesquisadores. White (2003) apresenta essas discussões e divide hipóteses em dois grandes grupos: aquelas que defendem que o estado inicial da aquisição de L2 é a L1 e aqueles que sustentam a ideia de que se adquire L2 partindo da Gramática Universal.

A primeira hipótese apresentada pela autora está inserida na Hipótese do Acesso Nulo: *The Full Transfer Full Access Hypothesis* de Schwartz & Sprouse<sup>2</sup>. Esses autores afirmam que um falante em processo de aquisição de L2 parte da gramática completa de sua L1. Dessa maneira, à exceção da matriz fonética dos itens morfolexicais, o estado inicial da aquisição de L2 é o estado final da L1 e, portanto, o

---

<sup>2</sup> (1994,1996 apud WHITE, 2003)

desenvolvimento da L2 é determinado pela conjugação dos seguintes fatores: estado inicial, *input*, aparato da GU e condições de aprendizibilidade.

De modo semelhante, Clahsen (1988) também afirma que as opções paramétricas disponíveis na GU para falantes em fase de aquisição de L1 deixam de existir para falantes de L2. Bley-Vroman (1989), por sua vez, parte dessa proposta para postular a Hipótese da Diferença Fundamental, segundo a qual a aquisição de L2 depende de fatores afetivos, motivação e atitudes advindas da instrução formal em L1.

Kato (2005b) rejeita essa hipótese na medida em que ela não leva em conta que os valores paramétricos da L2 não são os mesmos de sua L1. Para exemplificar essa problemática, a autora supõe que

Um americano aprendendo o português poderá começar preenchendo sempre o objeto, como em inglês, mas se ele aprende efetivamente a gramática do português, deverá começar a usar nulos em contextos relevantes. Poderíamos pensar que isso é relativamente fácil pois se trata de uma propriedade que pode ser o resultado apenas de um apagamento fonético. Lembremos-, porém, que o objeto nulo seria apenas uma das manifestações de uma propriedade mais abstrata, que licencia também outras construções. De fato, línguas de objeto nulo não licenciam apenas objetos nulos, mas exibem outras características (cf Raposo, 1998) (...). Só poderíamos afirmar que o americano remarcou apropriadamente o parâmetro do objeto nulo se apresentar o conjunto de propriedades que se aprende através da definição de um valor de parâmetro. Além disso, algumas dessas propriedades seriam dedutivamente aprendidas, não necessitando de exposição direta a elas. (KATO, 2005b, p. 196)

Além disso, para ir de encontro à Hipótese de Acesso Nulo, a autora também ressalta que a gramática intermediária do aprendiz de L2 (interlíngua) apresenta restrições da GU que não são encontradas no *input* fornecido ao falante.

Logo, o falante precisaria acessar a GU para mudar os valores desses parâmetros e, para dar conta dessa problemática, surge a Hipótese de Acesso Parcial, segundo a qual o acesso à GU só é permitido por meio da L1. Tem-se, por exemplo, *The Minimal Trees Hypothesis de Vanikka&Young-Scholten*<sup>3</sup> que, apesar de também assumir que o estado inicial da aquisição de L2 é baseado na L1, defende que a gramática desse estado inicial é apenas parcialmente baseada na L1. O que se utiliza da L1 seria somente o que concerne às categorias lexicais e à sua orientação linear. As categorias funcionais e gramaticais, por conseguinte, são adquiridas por acesso à Gramática Universal a partir do *input* da L2 em um esquema “bottom up”, ou seja,

---

<sup>3</sup> (1994, 1996 apud WHITE, 2003)

categorias mais altas como CP não podem ser adquiridas antes de categorias mais baixas, como DP. Assim, os falantes, ao adquirirem a L2, chegam, pelo menos parcialmente, ao mesmo estado final.

Eubank<sup>4</sup> traça um caminho semelhante ao proposto pela *The Minimal Trees Hypothesis* por meio de outra teoria: *The Valueless Features Hypothesis*. O autor afirma que os falantes alcançam o mesmo estado final, embora o estágio inicial da L1, do qual partiriam os falantes, contenha tanto categorias funcionais quanto lexicais. No entanto, a característica que dá nome à teoria e a distingue das demais é a seguinte: apenas os traços fracos da L1 são transferidos, ao passo que os traços fortes são desenvolvidos à partir do input da L2.

Também favorável à chamada Hipótese do Acesso Parcial, Meisel (2000) argumenta que os valores paramétricos assentados pelo falante ao fim da aquisição da L1 não podem ser alterados na aquisição de L2. Essa mesma teoria pode assumir outras duas formas: a) os valores paramétricos podem ser alterados desde que somente os princípios ativados durante a aquisição de L1 sejam acessados pela L2 e b) somente princípios não parametrizados podem ser acessados para que a L2 seja adquirida.

Já em defesa do acesso total à Gramática Universal pelos sujeitos em fase de aquisição de L2, alguns pesquisadores, como Selinker (1972) se apoiam na sistematicidade da interlíngua, na comparação entre falantes de L1 e falantes de L2 e na pobreza de estímulo. Uma vez que a gramática da interlíngua apresenta, muitas vezes, traços que não são provenientes da L1 nem da L2, mas que estão em consonância com restrições universais, vê-se que não há outra possibilidade senão esse acesso. Além disso, Birdsong (1992), outro estudioso dessa perspectiva, reafirma que os aprendizes de L2 adquirem valores paramétricos que são diferentes de sua L1 e que, portanto, se trata de um conhecimento que vai além do input de sua L1. Platzack<sup>5</sup>, autor da *The Initial Hypothesis of Syntax*, com base nessas evidências, propõe então que o estágio inicial para a aquisição de L2 é idêntico ao estágio inicial para a aquisição de L1.

Por outro lado, a hipótese *Full Access (Without Transfer)* de Flynn & Martohardjono<sup>6</sup>, Flynn<sup>7</sup> e Epstein et al<sup>8</sup> propõe que a Gramática Universal está

---

<sup>4</sup> (1993, 1994, 1996 apud WHITE, 2003)

<sup>5</sup> (1996 apud WHITE, 2003)

<sup>6</sup> (1994 apud WHITE, 2003)

disponível em todos os estágios da aquisição. Embora White (2003) afirme que a hipótese de *Full Access (Without Transfer)* careça de explicações sobre o que seria o estágio inicial, os estudos realizados por Epstein et al a fim de encontrar contra evidências para a *Minimal Trees Hypothesis* propõem que a gramática da qual partem os aprendizes de L2 contém um conjunto completo de categorias funcionais e, portanto, a interlíngua de falantes de diferentes L1 será a mesma, o que corrobora a Hipótese do Acesso Total.

Retornando, então, à questão da aquisição de clíticos no PB, Kato, Cyrino e Correa (2009) defendem que, uma vez que essa aquisição ocorre por meio da escrita como L2, o falante deve acessar à GU para encontrar uma possibilidade paramétrica que seja adequada a essa L2, como se vê no trecho a seguir:

Looking at the positions where the clitics were placed, further evidence of this accessibility can be found. In fact, during the early literacy period, students produce forms that belong neither to the L1 vernacular, nor to the “input” given at school, but which were possibilities available in earlier periods, namely clitic climbing, clitic doubling with the third person, and interpolation. It is important to point out that although 3<sup>rd</sup> person clitics are taught in school, a thorough instruction regarding their position is missing in the grammar syllabus. The learner has to search the possibilities in UG, until s/he finally sets the position partially according to his/her L1 namely obeying the rightward directionality of cliticization. (KATO, CORREA & CYRINO, 2009, p. 267)

Paralelamente, Roeper (1999) também defende que a aquisição de L2 só é possível pelo acesso à GU em sua teoria do Bilinguismo Universal. Para ele, todos os falantes são potencialmente bilíngues na medida em que tanto a opcionalidade quanto a variação lexical devem ser caracterizados como uma forma de bilinguismo. Isso é possível porque, em tese, todo falante possui um conjunto de minigramáticas para diferentes domínios sintáticos para dar conta de escolhas “contraditórias” existentes em todas as línguas. Mesmo que a diferença envolva apenas uma regra, Roper (1999) considera que se trata de duas gramáticas distintas. Dessa forma, a criança durante o período de aquisição da L1 alternaria entre duas gramáticas distintas, a GU e a gramática do input, assim como um falante adquirindo uma L2. A GU estaria disponível por meio de estruturas denominadas *Minimal Default Grammar (MDG)*, compostas por representações default baseadas nos princípios de economia de Chomsky (1995).

---

<sup>7</sup> (1996 apud WHITE, 2003)

<sup>8</sup> (1996, 1998 apud WHITE, 2003)

Essa teoria se relaciona, então, com o Princípio do Subconjunto (doravante PS) de Slabakova (2002). Essa proposta prevê que todo falante em fase de aquisição opte por uma gramática de abrangência menor, ou nos termos da teoria, que seja um subconjunto da outra. Isso só é possível quando há dois valores distintos para um mesmo parâmetro disponíveis de modo que são gerados dois sistemas gramaticais em uma relação de conjunto e subconjunto. Ao passo que pesquisadores como Manzini & Wexler (1987) atestaram que crianças parecem partir de uma gramática maximamente restritiva para adquirir sua L2, para a aquisição de L2 isso ainda não é uma afirmação incontestável. Diante disso, os autores se dividem entre postular que: a) o Princípio do Subconjunto não pode operar em aquisição de L2 para adultos<sup>9</sup>; b) não é possível testar PS, pois não há relação subconjunto-conjunto entre valores paramétricos existentes<sup>10</sup>; c) PS só é operante na aquisição adulta de L2<sup>11</sup>.

No entanto, Kato (2005a) ressalta que essa hipótese é interessante pela possibilidade de ser considerada como Acesso Total ou Acesso Parcial por meio da periferia marcada. Uma vez que a autora também se coloca em defesa do acesso à GU durante a aquisição da escrita. Ela assume que esse acesso não se dá diretamente, mas via gramática da língua falada, que seria a L1. Desse modo, a gramática da L1 apresentaria uma periferia marcada com valores paramétricos opostos aos da gramática nuclear e esses valores assumiriam um valor competitivo em relação a estes durante o período de escolarização. Para validar essa explicação, a autora se reporta aos estudos de Hershensohn (2000) e à teoria do construcionismo, a qual prevê que aprendizes de L2 têm acesso a restrições sobre a forma imposta pela GU, mas não à estratégia de GU, ou seja, à completude posterior à idade crítica na aquisição.

Hershensohn (2000), trabalhando a partir do quadro minimalista, parte da ideia de que a sintaxe não varia entre as línguas e apenas morfologia e léxico apresentam variação. Diante disso, a autora defende que o falante adquirindo L2 passa por três estágios de aquisição de construções morfolexicais. No primeiro, a L1 apresenta apenas alguns itens lexicais da L2; no segundo, os traços e os parâmetros da L2 começam a ser adquiridos; e no terceiro, os parâmetros são refixados para o valor da L2.

---

<sup>9</sup> (Cf HIRAKAWA, 1990; WHITE, 1989 apud MARCELINO, 2007)

<sup>10</sup> (Cf MACLAUGHLIN apud MARCELINO, 2007)

<sup>11</sup> (Cf BERENT, 1994 apud MARCELINO, 2007)

Cabe também destacar aqui que a hipótese defendida pela autora também prevê que não existe um período crítico para aquisição de linguagem, nos termos de Lenneberg (1967). Para ela, o período compreendido entre os dois e os doze anos de idade é o chamado período crítico para a aquisição de linguagem devido ao fato de a puberdade definir a especialização do hemisfério esquerdo para funções linguísticas. Já Kato (2003) defende que essa idade existe apenas nos níveis prosódico e fonológico, mas não no nível sintático. Isso acontece porque os parâmetros ligados à morfossintaxe são fixados depois dos relacionados à morfologia e à fonologia (KATO, 2003).

Marcelino (2007), também apoiado na visão defendida por Kato (2003), discute, então, qual seria o estágio em que a L2 é completamente adquirida pelo falante. A dificuldade para determinar esse estágio final também gera discussões entre os pesquisadores. Gass & Selinker (2001 apud MARCELINO, 2007), por exemplo, postulam três possibilidades de definição para esse estágio final: a) a primeira aparição de uma forma correta; b) certa porcentagem no uso de formas precisas ou; c) a primeira de três ocorrências consecutivas numa amostra de três semanas em que um morfema é visto em mais de 90% dos contextos obrigatórios. Outra proposta é a de Brown (2000 apud MARCELINO, 2007) em que o autor propõe o esquema de *backsliding*, no qual o falante, após o uso de uma forma gramatical corretamente, “retrocede” para um estágio anterior antes que a fase final se estabilize.

Marcelino (2007), contudo, questiona a visão de que a aquisição se dá apenas em termos de desempenho, ou seja, em relação ao uso de formas corretas. Para ele, é essencial levar em conta que, ao passo que a criança adquirindo uma L1 alcança estágio final em que se torna falante nativo de uma língua, o falante adquirindo L2 pode estacionar em um estágio cuja gramática não representa a da língua alvo. É nesse ponto que cabe discutir o papel da instrução e da aprendizagem formal para que ele saia desse estágio. Isto posto, Marcelino (2007) especula que, para prosseguir em direção ao estágio final dessa aquisição, alguns falantes necessitem de formalização, o que não é possível afirmar para falantes adquirindo L1. Por outro lado, há falantes capazes de sair desse estágio por si só e o que determina qual caminho cada falante segue ainda não é bem determinado pela teoria.

Diante disso, o autor defende que o estágio ideal, que se assemelha à competência do falante nativo, seria a somatória do estágio natural de desenvolvimento

da interlíngua acrescida dos ajustes para se adequar à gramática da língua alvo. Essa hipótese está de acordo com a hipótese do acesso indireto à GU e diminui o papel da GU enquanto fator limitante para o desenvolvimento de L2 já que reconhece que há outros fatores pré-existentes à aquisição e que não são de cunho estritamente linguístico. Logo, o insucesso na aquisição de L2 seria resultado da somatória de fatos linguísticos e extralinguísticos.

Convém discutir também aspectos apresentados por Pires (2015) na medida em que o estudo da autora partilha de objetivos semelhantes aos aqui propostos. A fim de investigar a realização do objetivo direto de 3ª pessoa anafórico no PB, ela defende que os clíticos acusativos se tornaram outro fóssil gramatical recuperado pela instituição escolar. Seguida da revisão bibliográfica dos autores que discutem a aquisição de L2, o estudo de Giusti (2004 apud PIRES, 2017) é exposto por Pires (2017) e, de igual modo, ele traz considerações significativas também para esta pesquisa.

Giusti (2004 apud PIRES, 2017) propõe que a língua escrita, salvos os fatores culturais independentes do módulo cognitivo da linguagem, tem relação estrita com a capacidade geral da mente humana. Para ela, o registro oral e o escrito não estão necessariamente ligados ao meio em que são produzidos, pois um texto cujo meio de transmissão é escrito pode apresentar características da fala da mesma forma que um texto que é veiculado oralmente pode apresentar características da escrita. Por essa razão, ela opta por se referir a registro oral ou escrito, ao invés de se referir à língua oral ou escrita.

Dando continuidade, ela argumenta que a alfabetização, por ocorrer em um momento em que o falante ainda pode adquirir uma L2 tal qual uma L1, funciona como aquisição de L2 e a exposição a dados de língua escrita seria o *input* para o processo, embora nem sempre ele seja suficiente. Ela defende que é possível adquirir a linguagem escrita de modo natural e que a exposição a dados não-selecionados facilita essa aquisição. Para ela, “a exposição a textos escritos por meio de leitura em voz alta por parte de adultos é o meio mais direto para transmitir características não explicitamente linguísticas do registro escrito” (GIUSTI, 2004, p. 78 apud PIRES, 2017)

Tal qual Kato (2005a), Giusti (2004 apud PIRES, 2017) propõe que os primeiros textos escritos em uma língua cuja distância entre o discurso falado e o discurso escrito é muito grande apresentam características da fala. No italiano, língua

na qual a autora se baseia para realizar seu estudo, ela verifica que a variação entre italiano escrito e o dialeto falado se dá também na forma de *code-switching*.

Giusti (2004 apud PIRES, 2017) passa, então, a refletir sobre o papel da instrução nessa aquisição específica e defende que, em relação à competência sintática, ela é pouco produtiva já que é justamente a exposição aos dados que desenvolvem a competência nessa L2. Portanto, essa aquisição só se completa quando o aprendiz é exposto a dados de registro escrito da mesma forma em que é exposto a dados de fala, ou seja, sem que isso se restrinja apenas à esfera escolar.

Para ela, a habilidade que parece ser tangente à competência sintática e que merece destaque na reflexão sobre a aquisição da escrita é a chamada *competência textual*. Essa competência seria a forma mental da competência pragmática, que ainda está em processo de aquisição em uma criança de oito anos de idade. Desse modo, as diferentes características das produções escritas de crianças que estão adquirindo esse registro estão relacionadas não apenas à sintaxe, mas também à pragmática já que segundo Giusti (2004 apud PIRES, 2017), a competência textual do registro escrito nunca é concluída ao contrário da competência sintática, cuja aquisição se dá por completa aos quatro anos de vida do falante. Por exemplo, em uma sentença como (21) abaixo, o uso generalizado do pronome clítico dativo de terceira pessoa *le*, que seria apenas feminino, não estaria relacionado à sintaxe apenas, pois a sintaxe não é violada, mas a competência textual não está totalmente adquirida.

(21) [Um bambino...] Quando arrivà la pagella a casa la mamma le diede una punizione.

[A boy...] When came the mother to home, mom him gave a punishment.

‘[A boy...] When the mother came home, she gave him a punishment.’

[Um menino...] Quando a mãe chegou em casa, ela lhe deu uma punição.

Ocorrências semelhantes a essa foram vistas pela autora tanto em produções escritas infantis quanto em textos literários do século XIII. Para ela, embora esses casos possam ser analisados como uma interferência do registro oral no registro escrito, esse

fato indica que a competência textual se desenvolve paralelamente à verificação dessa interferência.

#### 1.4. Concluindo o capítulo

Neste primeiro capítulo, foi apresentada e discutida a problemática da aquisição de clíticos no PB. A partir do trabalho de Kato, Cyrino & Correa (2009), viu-se que os clíticos acusativos de terceira pessoa são adquiridos apenas via escolarização e, propõe-se que essa seja também a situação da aquisição do clítico *se* indeterminador e apassivador. Com isso, a gramática aprendida na escola, que contém fósseis da linguagem que são postulados como gramaticais, seria aprendida tal qual uma L2 para os falantes do PB. Para discutir essa aquisição, foram apresentadas algumas das principais teorias acerca da aquisição de L2. O quadro 1 abaixo sintetiza as ideias principais de cada autor resenhado acima:

Autor	Hipóteses sobre o estágio inicial da GU	Contribuição
Schwartz & Sprouse (1994, 1996); Clahsen (1988).	Hipóteses do Acesso Nulo à GU	<ul style="list-style-type: none"> <li>• À exceção da matriz fonética dos itens morfolexicais, o estado inicial da aquisição de L2 é o estado final da L1;</li> <li>• As opções paramétricas disponíveis na GU para falantes em fase de aquisição de L1 deixam de existir para falantes de L2.</li> </ul>
Bley-Vroman (1989)	Hipótese da Diferença Fundamental	A aquisição de L2 depende de fatores afetivos, motivação e atitudes advindas da instrução formal em L1.
<ul style="list-style-type: none"> <li>• Vanikka&amp;Young-Scholten (1994, 1996);</li> <li>• Meisel (2000).</li> </ul>	Hipótese de Acesso Parcial à GU	<ul style="list-style-type: none"> <li>• Apenas os traços fracos da L1 são transferidos da GU, ao passo que os traços fortes são desenvolvidos à partir do input da L2;</li> <li>• Os valores paramétricos assentados pelo falante ao fim da aquisição da L1 não podem ser alterados na aquisição de L2</li> </ul>
<ul style="list-style-type: none"> <li>• Selinker (1972);</li> <li>• Platzack (1996);</li> <li>• Birdsong (1992);</li> <li>• Flynn &amp; Martohardjono (1994).</li> </ul>	Acesso Total à GU	<ul style="list-style-type: none"> <li>• Aprendizes de L2 adquirem valores paramétricos que são diferentes de sua L1;</li> <li>• A gramática da qual partem os aprendizes de L2 contém um conjunto completo de categorias funcionais e, portanto, a interlíngua de falantes de diferentes L1 será a mesma.</li> </ul>
Roeper (1999)	Bilinguismo Universal	Todos os falantes são potencialmente bilíngues na medida em que tanto a opcionalidade quanto a variação lexical devem ser caracterizados como uma forma de bilinguismo.
Slabakova (2002).	Princípio do Subconjunto	Todo falante em fase de aquisição opte por uma gramática de abrangência menor, ou nos termos da teoria, que seja um subconjunto da outra.
Hershensohn (2000)	Construcionismo	Aprendizes de L2 têm acesso a restrições sobre a forma imposta pela GU, mas não à estratégia de GU, ou seja, à completude posterior à idade crítica na aquisição.

Marcelino (2007)	Estágio final da aquisição de L2	O falante adquirindo L2 pode estacionar em um estágio cuja gramática não representa a da língua alvo
Giusti (2004)	Papel da instrução na aquisição de L2	A aquisição só se completa quando o aprendiz é exposto a dados de registro escrito da mesma forma em que é exposto a dados de fala.

### Quadro 1 – Síntese das principais hipóteses sobre aquisição de L2

Assume-se, então, a hipótese do Acesso Parcial tal qual defendida por Kato (2005, 2003) para a aquisição da gramática escolar como L2. Em seguida, viu-se que, adquirida a gramática ensinada na escola, o falante escolarizado alterna entre essa e a gramática adquirida naturalmente pelos falantes do PB na forma de *Code-switching* (CS). A fim de discutir esse fenômeno, também foram apontados os principais autores que tratam do assunto. Segue abaixo o quadro 2 que sumariza as discussões apresentadas:

### Quadro 2 – Síntese dos principais estudos sobre *code-switching*

Autor	Principal discussão	Referencial teórico	Contribuição
Poplack (2001)	Conceito de CS	Poplack (2001)	CS é um fenômeno observado na fala de bilíngues em que ocorre a mixagem de duas ou mais línguas em um mesmo discurso.
Toribio (2017)	Revisão bibliográfica dos estudos sobre CS	<ul style="list-style-type: none"> <li>• Poplack (1980);</li> <li>• Myers-Scotton (1993);</li> <li>• Gardner-Chloros (2009);</li> <li>• Sebba (1999, 2009);</li> <li>• Chan (2009).</li> </ul>	<ul style="list-style-type: none"> <li>• CS acontece em pontos da estrutura em que não divergem nas duas línguas;</li> <li>• Para que o CS ocorra, uma língua deve prover o quadro morfossintático ao qual elementos de outras línguas são incorporados (MLF);</li> <li>• CS é um processo que envolve múltiplos fatores, até mesmo externos, como o quadro social do falante.</li> </ul>
Riehl (2005)	Contextos favorecedores de CS	Clyne (2003)	“Palavras-gatilho”: nomes próprios, palavras homófonas e palavras que permitem transferências lexicais.
Kato (1994)	<ul style="list-style-type: none"> <li>• Conceito de CS</li> <li>• Contextos favorecedores de CS</li> </ul>	<ul style="list-style-type: none"> <li>• REYES (1982);</li> <li>• Sankoff&amp;Poplack (1981).</li> </ul>	<ul style="list-style-type: none"> <li>• Diferença entre CS e <i>borrowing</i>;</li> <li>• Não é apenas a sintaxe que condiciona o processo mas também a prosódia;</li> <li>• são os itens funcionais que determinam a estrutura sobre a qual inserções de uma outra língua são feitas.</li> </ul>
Kato (1996, 2001)	Contextos favorecedores de CS	<ul style="list-style-type: none"> <li>• Weinreich (1953);</li> <li>• Bickerton (1981);</li> <li>• Silva-Corvalan (1993);</li> <li>• Lightfoot (1991);</li> <li>• Roberts (1993).</li> </ul>	<ul style="list-style-type: none"> <li>• A escolarização representa um fator determinante;</li> <li>• A criança chega à escola com sua gramática nuclear e passa a conviver com outra gramática na escola, na qual será alfabetizada. No entanto, ao invés de se tornar mais fluente na gramática escolar, o falante alterna entre os dois códigos.</li> </ul>
Kroch (2003)	CS e mudança linguística	Kroch (2003)	Durante o período de mudança linguística em curso, seja ela por fatores linguísticos ou extralinguísticos, os falantes da língua que estão passando pelo processo tendem a apresentar as formas antigas concomitantemente às formas novas.

## Capítulo 2 – O Estatuto do *se* indeterminador e apassivador

Tal qual uma marca do conservadorismo escolar, o clítico *se* indeterminado e apassivador aparece nos livros e manuais da Gramática Tradicional como fator gerador de dúvidas. No entanto, seu destaque tem sido reduzido. As explanações acerca desse clítico aparecem nas gramáticas tradicionais nos capítulos dedicados aos termos essenciais da oração, especialmente o sujeito - para o *se* indeterminado - e no capítulo sobre voz ativa e voz passiva - para o *se* passivo.

### 2.1.1 O Índice de Indeterminação do Sujeito

Antes de resenhar o que é dito pelos gramáticos contemporâneos acerca do *se* como elemento de indeterminação, cabe ressaltar a preocupação expressa na maioria das obras em orientar o leitor a não confundir sujeito indeterminado com sujeito indefinido. Esse tipo de alerta permite que se diga que a nomenclatura e o funcionamento desse clítico não são de fácil compreensão para alunos em idade escolar, que consultam as gramáticas. Além disso, foi possível notar que esse elemento tem ganhado menos destaque nas obras. Aparentemente, embora a escola não dispense totalmente a construção de indeterminação com o *se* em uma boa estrutura, outras formas de indeterminação do sujeito têm ganhado espaço nos manuais. A exemplo disso, Rocha Lima (1994) afirma que a língua se vale de dois expedientes para realizar a indeterminação. A primeira seria o emprego da 3ª pessoa do plural, sem referência anterior ao pronome *eles* ou *elas* como em (22). O segundo seria usar o verbo na 3ª pessoa do singular acompanhado da partícula *se* desde que o verbo seja intransitivo ou traga complemento preposicional, como em (23).

(22) Levaram todo o dinheiro da loja. (ROCHA LIMA, 1994)

(23) Dorme-se cedo nesta casa. (ROCHA LIMA, 1994)

No capítulo dedicado à discriminação dos tipos de sujeito, Bechara (1992), por sua vez, define sujeito indeterminado como aquele que “não se nomeia ou por não se querer ou por não se saber fazê-lo”. Em seguida, apresenta outra forma de realizar a indeterminação do sujeito: o uso de expressões como *a gente* ou *alguém*, como em (24) e (25).

(24) Para o bolo ficar bonito, a gente tem que bater os ovos separadamente.

(BECHARA, 1992)

(25) Alguém abre a loja todos os dias. (BECHARA, 1992)

Dito isso, o autor considera o *se* como outra forma de indeterminação do sujeito, chamando a partícula de “índice de indeterminação do sujeito” e afirma que, para se utilizar dessa estratégia, deve-se empregar o pronome *se* junto a verbo de modo que a oração passe a equivaler a outra que tem por sujeito alguém, a gente ou expressão sinônima. Essa estrutura seria equivalente à colocação do verbo na 3ª pessoa do singular, como em (26), ou do plural, como em (27) sem referência a pessoa determinada.

(26) Diz que eles vão bem. (BECHARA, 1992)

(27) Estão chamando o vizinho. (BECHARA, 1992)

(28) Abre-se às dez. (BECHARA, 1992)

Mais recentemente, Bechara (2009) chama sentenças como (28) de *impessoais* por não haver substantivo claro, deixando a função de sujeito do conteúdo predicativo para o *se* e ainda acrescenta, em suas observações finais, que

o *se* como índice de indeterminação de sujeito - primitivamente exclusivo em combinação com verbos não acompanhados de objeto direto -, estendeu seu papel aos transitivos diretos (onde a interpretação passiva passa a ter uma interpretação impessoal. (...) A passagem deste emprego da passiva à indeterminação levou o falante a não mais fazer concordância, pois o que era sujeito passou a ser entendido como objeto direto, função que não leva a exigir o acordo do verbo. (BECHARA, 2009, p. 178)

Cegalla (2007) adiciona ainda uma terceira forma de realizar a indeterminação, deixando-se o verbo no infinitivo não-flexionado, como em (29) e (30):

(29) Era penoso carregar aqueles fardos enormes. (CEGALLA, 2007)

(30) É triste assistir a estas cenas repulsivas. (CEGALLA, 2007)

Ao encontro dessa ideia, Luft (1976) já excluía como estratégia de indeterminação o uso do pronome *se*. Para ele, há apenas os outros dois tipos já

mencionados: emprego da 3ª pessoa do plural sem sujeito explícito e infinitivo não flexionado. Ele acrescenta, no entanto, que

é má doutrina gramatical dar os pronomes indefinidos como sujeitos “indeterminados”: Alguém bateu à porta. Neste caso, o sujeito sintaticamente “determinado” - é alguém. Sujeito indeterminado é aquele que, logicamente existente, não existe na frase (sem ser oculto por elipse). (LUFT, 1976, p.130)

Contudo, antagonicamente a esses últimos autores, que ainda consideram outras formas de indeterminação, Cunha&Cintra (2001) apenas identificam o sujeito indeterminado em sentenças com verbo na 3ª pessoa do plural sem sujeito expresso ou na 3ª pessoa do singular com o pronome *se*. Cipro Neto & Infante (2003) também defendem essa proposta, mas esclarecem que construções com *se* são permitidas apenas com verbos que não apresentam complemento direto (verbos intransitivos, transitivos indiretos e de ligação), como em (31) e (32):

(31)Falam sobre a reforma agrária. (CUNHA&CINTRA, 2001)

(32)Fala-se sobre a reforma agrária. (CUNHA&CINTRA, 2001)

Outro adendo interessante feito pelos autores é que, quando se utiliza a indeterminação com verbo na terceira pessoa do plural, quem produz a oração não se inclui necessariamente como agente da ação, ao contrário do que ocorre em sentenças como (32), já que o enunciador pode ou não se incluir na ação de falar sobre a reforma agrária.

Nota-se, portanto, que as advertências sobre o que é ou não permitido em sentenças com *se* indeterminador variam de um autor para o outro. A fim de aclarar essas situações, o Manual de redação e estilo do Estado de São Paulo, publicado em 2010 pelo jornal, resume as situações em que não se deve utilizar o pronome *se* como índice de indeterminação do sujeito:

Não se usa o pronome *se*: 1 — Em expressões como difícil de, fácil de, bom de, ruim de, passível de, duro de, agradável de, etc., que já têm valor passivo (difícil de ser feito, fácil de ser entendido): fácil de entender, difíceis de vender, bom de fazer, ruim de tomar, passível de errar, duro de roer, agradável de recordar. 2 — Quando o conjunto pronome mais infinitivo equivale a um adjetivo (é de admirar = é admirável): Foi de espantar (espantosa) tamanha ousadia. / Eram de esperar (esperadas) melhores relações entre os dois países. / Foram de pasmar suas ofensas. / Serão de temer novos retrocessos/ É de entender que ele se comporte dessa forma. Igualmente: de notar, de impressionar, de compreender, de tolerar, de acreditar. No negativo, usa-se a mesma forma: Não serão de temer tais

retrocessos. / Não é de espantar que se comporte assim. / Não seria de esperar tamanha compreensão. 3 — Quando não tem função alguma na oração: É preciso pensar (e não “pensar-se”) nisso. / É preciso cogitar (e não “cogitar-se”) desse caso. / É hora de fazer o inventário. / É difícil conseguir a cura da aids. / É comum encontrar pessoas paradas nessa esquina. / Esta é uma forma de fortalecer o partido. / O contra-ataque é a maneira mais fácil de chegar ao gol. / Há planos para construir um centro de produção em São Paulo. “Se a”, “se o”, “se as”, “se os”. São erradas orações em que o se aparece na função de sujeito ao lado do pronome: Não “se o” sabe. / Sempre “se a” vê. / As páginas ficarão mais legíveis desenhando-“se-as” à tinta. / Fazendo “se-os” sair já, chegarão ao litoral antes de escurecer. Outros exemplos vetados: “dê-se-o”, “diz-se-a”, “acredita-se-o”, etc. Não confunda com o se conjunção: Se o homem sair... / Se as mulheres vierem...

Esse tipo de repreensão a formas, principalmente quando acompanhado de exemplos de sentenças que são de uso corrente em variáveis de menos prestígio, apenas reafirma a gramaticalidade dessas sentenças para os falantes. Inclusive, mais adiante nesse trabalho, os dados indicarão que essas formas são frequentes na fala de adultos letrados e, conseqüentemente, pouco estranhamento causariam se lidas em uma redação que circula numa esfera social em que se preza pela norma padrão culta.

### 2.1.2 A Partícula Apassivadora

As considerações sobre o uso do *se* apassivador nos manuais de Gramática Tradicional são levantadas, geralmente sem muito destaque, nas sessões dedicadas ao estudo das vozes verbais. Para Bechara (1997), o *se* surge como forma de passivizar sentenças em que o sujeito é indeterminado como em (33). A versão passiva, chamado então de passiva pronominal, ficaria como vemos em (34):

(33) Vendem casas. (BECHARA, 1997)

(34) Vendem-se casas. (BECHARA, 1997)

Em uma versão mais recente, Bechara (2009) dedica ao *se* apenas um item entre o uso dos pronomes, primeiramente falando de construções as quais ele chama de “voz média” ou “Passiva com se”. A exemplo dessas construções, o autor apresenta o seguinte exemplo:

(35) O banco só se abre. (BECHARA, 2009)

Para o autor, essa interpretação passiva depende apenas do significado lexical do verbo uma vez que essa interpretação prevalece ainda nas orações cujo sujeito não é inanimado e, portanto, passível de executar a ação verbal.

Cegalla (2007), por sua vez, acrescenta a essa mesma proposta a advertência de que, quando utilizada a voz passiva pronominal, não se deve declarar o agente do verbo: Por exemplo:

(36) Nas ruas assobiavam-se as canções dele pelos pedestres. (errado)  
(CEGALLA, 2007)

(37) Assobiavam-se as canções dele nas ruas. (certo) (CEGALLA, 2007)

Mais uma vez, esse tipo de censura gramatical revela que a forma circula no vernáculo mesmo que o grau de aceitabilidade da sentença varie entre os falantes.

Ao passo que esses últimos autores propõem uma distinção clara entre o que seria o *se* índice de indeterminação do sujeito e o *se* partícula apassivadora com base na forma assumida pelo verbo na sentença (singular e plural, respectivamente); isso não acontece em relação a outros autores. Cunha & Cintra (2001) exemplificam o *se* como elemento de passivização com dois exemplos:

(38) Aumentou-se o salário dos gráficos. (CUNHA&CINTRA, 2001)

(39) Conteve-se a inflação em níveis razoáveis. (CUNHA&CINTRA, 2001)

Cipro Neto & Infante (2003) explicam essa proposta com mais detalhes. Para eles, a chamada voz passiva sintética, é formada a partir de uma oração na voz ativa cujo sujeito está indeterminado. Dessa forma, uma sentença na voz ativa com sujeito indeterminado seria capaz de gerar as duas formas de sentença na voz passiva: a voz passiva analítica sem agente exposto, como em (40), e a voz passiva sintética com o *se*, como em (41) e (42). Os autores colocam ainda que, na voz passiva sintética, o verbo concorda em número e pessoa com o sujeito da oração e que, assim como a voz passiva analítica, sua formação depende de verbos transitivos diretos ou transitivos diretos e indiretos.

- (40) O apartamento foi alugado. (CIPRO NETO & INFANTE, 2003)  
 (41) Alugou-se o apartamento. (CIPRO NETO & INFANTE, 2003)  
 (42) Alugaram-se os apartamentos. (CIPRO NETO & INFANTE, 2003)

A fim de resolver o problema da concordância entre o verbo e seu argumento interno que se coloca, o próximo item desse manual se dedica à distinção entre as duas funções do *se* que ali acima se confundem: partícula apassivadora e índice de indeterminação do sujeito. De acordo com Cipro Neto & Infante (2003), para distinguir as duas funções do *se* basta analisar o verbo: quando se tem verbos transitivos diretos e verbos transitivos diretos e indiretos, forma-se a voz passiva sintética. Com verbos com as demais transitividades, tem-se *se* como índice de indeterminação.

Seguindo essa proposta, então, os autores afirmam que a interpretação das sentenças, (43) e (44) abaixo devem ser diferentes. Em (43), “uma solução para o problema” deve ser o sujeito do verbo, embora o agente do processo esteja indeterminado. Em (44), “em teses suspeitíssimas” é objeto indireto, pois o sujeito está indeterminado.

- (43) Imagina-se uma solução para o problema. (CIPRO NETO & INFANTE, 2003)  
 (44) Confia-se em teses suspeitíssimas. (CIPRO NETO & INFANTE, 2003)

Como nota de observação, os autores dizem que é possível realizar a indeterminação do sujeito com verbos transitivos diretos desde que esses objetos apareçam preposicionados, como em (45) e (46):

- (45) Estima-se aos bons amigos. (CIPRO NETO & INFANTE, 2003)  
 (46) Ama-se aos pais. (CIPRO NETO & INFANTE, 2003)

A partícula apassivadora *se* não é, em suma, um ponto pacífico de discussão entre os manuais de Gramática Tradicional. Alguns autores marcam sua diferença em relação ao *se* índice de indeterminação do sujeito por meio da forma singular ou plural

dos verbos. Outros defendem que é a transitividade do verbo que determina a função do *se* na sentença.

Para sumarizar as principais ideias dos autores apresentados acima, o quadro 3 apresenta, por autor, a classificação dada ao clítico *se* e suas características na Gramática Tradicional:

**Quadro 3 – Síntese das principais ideias de Gramática Tradicional sobre o *se***

<b>Autor</b>	<b>Forma de Indeterminação</b>	<b>Características</b>
Rocha Lima (1994)	Índice de Indeterminação do sujeito	Verbo na 3ª pessoa do singular acompanhado da partícula <i>se</i> desde que o verbo seja intransitivo ou traga complemento preposicional
Bechara (1992)	Índice de Indeterminação do sujeito	<i>Se</i> junto a verbo de modo que a oração passe a equivaler a outra que tem por sujeito <i>alguém, a gente</i> ou expressão sinônima.
Cipro Neto & Infante (2003)	Índice de Indeterminação do sujeito	Construções com <i>se</i> são permitidas apenas com verbos que não apresentam complemento direto (verbos intransitivos, transitivos indiretos e de ligação)
Bechara (2009)	Índice de Indeterminação do sujeito	<i>Se</i> estendeu seu papel aos transitivos diretos (onde a interpretação passiva passa a ter uma interpretação impessoal.
Bechara (2009)	Pronome Apassivador	Essa interpretação passiva depende apenas do significado lexical do verbo uma vez que essa interpretação prevalece ainda nas orações cujo sujeito não é inanimado e, portanto, passível de executar a ação verbal.
Cegalla (2007)	Pronome Apassivador	Mesma proposta de Bechara (2009)
Cunha & Cintra (2001)	Pronome Apassivador	O verbo (transitivos diretos ou transitivos indiretos) concorda em número e pessoa com o sujeito da oração.
Cipro Neto & Infante (2003)	Pronome Apassivador	Sentenças com verbos transitivos diretos e verbos transitivos indiretos.

## **2.2. Na Gramática Funcionalista**

Para verificar o estatuto ocupado pelo *se* no PB contemporâneo cabe, então, olhar também para as chamadas Gramáticas Descritivas, cuja fundamentação teórica está na Gramática Funcionalista. Diferentemente da Gramática Tradicional, que é normativa e prescritiva, essa não se preocupa em apresentar o “bom uso” da língua, mas se propõe a afirmar o que se vê na língua a partir de um corpus, falado ou escrito. Castilho (2010), ao traçar os postulados e discorrer sobre a metodologia empregada no processo de criação de uma Gramática Funcionalista, esclarece o que primeiramente as difere das gramáticas normativas: o conceito de língua.

Aqui, ver-se-á a língua tal qual um conjunto ordenado de itens que estabelecem entre si relações linguísticas e desempenham funções sintáticas. A realização desses itens está sujeita à variação de uso, que deve ser examinada em suas correlações com fatores linguísticos e extralinguísticos. (CASTILHO, 2010, p. 45)

Embora não haja nenhuma especificação em relação à frequência de uso, Castilho apresenta três estratégias de indeterminação do sujeito, sendo uma delas inédita na Gramática Tradicional. Além do uso de *se* em sentenças com verbo na 3ª pessoa do singular e da elipse de sujeito com verbo na terceira pessoa do plural, ele traz também o sujeito expresso por pronomes pessoais de “referenciação genérica”, rótulo dado por Neves (2000, p. 463 apud CASTILHO, 2010, p. 298), como em (47) e (48):

(47) Normalmente, quando você não sabe o que fazer, é melhor não fazer nada. (NEVES, 2000, apud CASTILHO, 2010)

(48) Depois da crise econômica, eles deram de dizer que as centrais de atendimento não podem passar de um minuto para te atender. (NEVES, 2000, apud CASTILHO, 2010)

Em (47), *você* não é a segunda pessoa do discurso, com quem se fala, mas um agente genérico para ação, um sujeito indeterminado. Da mesma forma, em (48), *eles* não é um sujeito identificado previamente no discurso, mas também um sujeito indeterminado.

Quanto ao *se* apassivador, o autor define as passivas pronominais como as sentenças em que o verbo é fronteado, o sujeito é posposto e o complemento paciente é omitido, como em (49):

(49) Cortou-se o menino. (CASTILHO, 2010, p. 437)

De acordo com o autor, a concordância do verbo com o sintagma nominal posposto era obrigatória [sic]<sup>12</sup> na língua durante muito tempo. Essa mudança, no entanto, está diretamente relacionada à reorganização do quadro dos pronomes pessoais do PB e à perda do traço de reflexividade, o que resultou na perda do traço de apassivação.

O autor discute, então, se, na voz passiva pronominal, o verbo concorda ou não concorda com o substantivo plural em sentenças como (50):

---

<sup>12</sup>Mesmo que considerada uma gramática funcionalista, ainda se vê marcas de um discurso normativo.

(50) Vende-se casas com uma boa vista, só que à vista. (CASTILHO, 2010, p.481)

De acordo com o autor, no latim vulgar, o *se* reflexivo ganhou novos papéis, como o de pronome apassivador já que a ideia vinha de passividade, derivada da noção de espontaneidade. Dessa forma, bastava adicionar o agente da ação que se comprovava o valor de passividade. Quando o *se* reflexivo passou a ser considerado de terceira pessoa, o sentido de indeterminação do sujeito substituiu o de passivização, desaparecendo com a concordância e fazendo com que o sujeito passivo fosse reanalisado como objeto direto. Como resultado dessa mudança, a interpretação passiva se tornou restrita a contextos muito formais de comunicação.

Dito isso, Bagno (2011) se debruça sobre o que está acontecendo no PB em relação às formas de indeterminação do sujeito e afirma, de antemão, que não existe a chamada voz passiva sintética. Para ele, *se* só pode ser pronome-sujeito ou pronome reflexivo e, diante disso, “as denominações tradicionais de “índice de indeterminação do sujeito” e de “partícula apassivadora” são termos ad hoc e são um exemplo do excesso de nomenclatura para uma única palavra”. (BAGNO, 2011, p. 797)

Ao olhar, então, para sentenças como (51) e (52), o autor acredita que é inegável a existência de uma incongruência na tradição gramatical que oferece duas análises distintas para um mesmo clítico que exerce a mesma função de sujeito-agente [+indeterminado] tanto na voz passiva quanto na voz ativa:

(51) O vaso se quebrou. (BAGNO, 2011)

(52) Aqui se vive feliz. (BAGNO, 2011)

Além da perda da concordância, o autor apresenta duas explicações para a perda da interpretação passiva de *se*. Por um lado, esse fenômeno deve-se a um aspecto semântico sistematicamente desprezado pela Gramática Tradicional: orações que permitem tanto a interpretação indeterminada quanto a passiva apresentam verbos que só podem ser praticados por um sujeito com traço semântico [+animado] e/ou [+humano]. Por essa razão, os falantes mantem o verbo no singular, fazendo com que

ele concorde com o sujeito indeterminado, expresso no enunciado pelo clítico *se* (BAGNO, 2011).

Por outro lado, tem-se as sentenças em que *se* é empregado entre uma preposição e um infinitivo, como em (53). Nessas construções, o falante não interpreta como um lugar bom *de ser morado* mas como um lugar bom *para alguém morar*:

(53) Lugar bom de se morar. (BAGNO, 2011)

Em suma, o autor propõe que a forma que está competindo com o uso do *se* em sentenças indeterminadas é o uso de expressões genéricas, que variam de acordo com o grau de monitoramento da fala. Bagno (2011) lista algumas das mais utilizadas: *a criatura, a pessoa, beltrano, fulano, neguinho, o camarada, o cara, o cidadão, o cristão, o maluco, o outro, o pessoal, o povo, o sujeito, sicrano, o zé, o ser*.

Cabe também apresentar aqui as considerações de Neves (2000) em sua Gramática de Usos do Português. A autora afirma que, embora não seja de uso normativo, seu trabalho invoca a norma a fim de comparar as restrições aos usos contemporâneos da língua, muitas vezes condenados na tradição gramatical. No entanto, é necessário fazer uma ressalva em relação ao que a autora postula uma vez que, para realizar o trabalho descritivo da língua, a autora se baseou em um corpus escrito.

É interessante notar que a autora afirma que o *se* faz parte das construções com voz passiva, apesar de não classificar essas sentenças, especificamente, como os autores da Gramática Tradicional que a chamam de “voz passiva pronominal”. Tampouco o *se* é nomeado, ele apenas aparece como um indicador de voz passiva, caso esteja ligado a um verbo transitivo, como em (54) e (55):

(54) Dá-se manteiga e leite, alguma carne, roupas necessárias e pronto!  
(NEVES, 2000)

(55) Na prática, porém, viram-se cenas como os dois rapazes palestinos amarrados sobre o capo dos jipes militares formando um escudo humano contra as pedradas dos manifestantes. (NEVES, 2000)

Já no capítulo dedicado aos pronomes, há apenas uma pequena menção ao *se*. A autora considera que, dentro do eixo da terceira pessoa há um eixo não subjetivo, que abarca tudo o que não está implicado na interação verbal ou as entidades a que se faz referência na sala e, por isso, a terceira pessoa também é chamada de não pessoa (BENVENISTE, 1966 apud NEVES, 2000). Essa terceira pessoa, então, pode ser indeterminada, como já foi visto neste trabalho anteriormente. Além dessa, a autora elenca também duas formas já preconizadas pela Gramática Tradicional: o uso de verbos da terceira pessoa no plural e o uso de verbos na terceira pessoa do plural sem um pronome sujeito.

Para enriquecer essa descrição, é importante notar também o que Neves (2003) pondera em trabalho mais recente. Em seu Guia de Usos do Português, a autora apresenta sob outro formato uma descrição da língua falada no Brasil, dessa vez, em verbetes. No verbete *se*, vê-se duas asserções: primeiramente, como índice de indeterminação do sujeito; em seguida, como pronome apassivador. Na primeira, a autora descreve de antemão o que prescreve a gramática tradicional: “o pronome *se*, quando usado junto de verbo que na voz ativa não tem objeto direto, forma uma oração de sujeito indeterminado, com verbo na terceira pessoa do singular. Esses verbos podem se construir sem complemento (56) ou com complemento iniciado por preposição (57). Também é permitido que o verbo demande predicativo de sujeito (58) (NEVES, 2003, p. 690)”:

(56) Por ideias vive-se ou morre-se. (NEVES, 2003)

(57) Precisa-se de porteiro. (NEVES, 2003)

(58) Não se é imortal, vai-se morrer. (NEVES, 2003)

Contrariamente a essas prescrições, a autora flagra também o uso do pronome com verbos que, na voz ativa, prescindem de objeto direto.

Quanto ao *se* como pronome apassivador, a autora diz que nas lições tradicionais, o pronome *se* quando usado junto de verbo que na voz ativa tem objeto direto, forma uma passiva na qual essa expressão funciona como sujeito. A partir daí, indica-se que, se a expressão que é sujeito dessa voz passiva for plural, o verbo vai para o plural para concordar com seu sujeito.

Finalmente, a Gramática Houaiss da Língua Portuguesa de Azeredo (2008) descreve os dois processos mais citados acima para a construção da indeterminação do sujeito no PB: o uso da terceira pessoa do plural e o uso da terceira pessoa singular acompanhado de *se*. Além disso, vê-se também nessa obra o uso de verbos no modo infinitivo para indeterminar o sujeito da ação. Contudo, o autor não define situações gramaticais em que cada uma dessas ocorrências aparece com mais intensidade. O autor apenas observa que no registro informal típico da conversação espontânea,

a indeterminação do sujeito faz raro uso do *se*, dando preferência ao emprego da terceira pessoa do plural ou mesmo do singular. Em qualquer variedade da língua, o uso do pronome *se* como recurso de indeterminação permite que o enunciador se inclua, não como indivíduo, mas como parte do grupo social, na referência genérica e indeterminada do sujeito. A verdade, porém é que, na interação espontânea, a omissão da identidade do sujeito se exprime correntemente mediante o emprego de sintagmas nominais de significação genérica ou indeterminadora como “a gente”, “muita gente”, “todo mundo” no papel de sujeito simples. Observa-se ainda, que só é possível a indeterminação do sujeito quando o predicado se refere a algum ser humano no papel de sujeito. (AZEREDO, 2008, p 226)

Para sumarizar as principais ideias dos autores apresentados acima, o quadro 4 apresenta, por autor, a classificação dada ao clítico *se* e suas características na Gramática Funcionalista:

**Quadro 4 – Síntese das principais ideias de Gramática Funcionalista sobre o *se***

<b>Autor</b>	<b>Forma de Indeterminação</b>	<b>Características</b>
Castilho (2010)	a. <i>Se</i> indeterminador; b. Elipse de sujeito; c. Pronomes pessoais de referência genérica; d. <i>Se</i> passivador.	a. Verbo na 3ª pessoa do singular; b. Verbo na 3ª pessoa do plural; c. <i>Você, eu</i> ; d. O verbo é frontado, o sujeito é posposto e o complemento paciente é omitido.
Bagno (2011)	a. Sentenças indeterminadas; b. Expressões genéricas; c. Não existe a chamada voz passiva sintética	a. Verbo na 3ª pessoa do singular; b. A criatura, a pessoa, beltrano, fulano, neguinho, o camarada, o cara.
Neves (2000)	a. <i>Se</i> indeterminador; b. <i>Se</i> passivador; c. Uso de verbos da terceira pessoa no plural; d. Uso de verbos na terceira pessoa do plural sem um pronome sujeito.	a. Quaisquer tipos de verbo; b. Junto de verbos que na voz ativa tem objeto direto.
Azeredo (2008)	a. Uso da terceira pessoa do plural; b. Uso da terceira pessoa singular acompanhado de <i>se</i> ; c. Emprego de sintagmas nominais de significação genérica ou indeterminadora	a. e b. No registro informal típico da conversação espontânea, a indeterminação do sujeito faz raro uso do <i>se</i> , dando preferência ao emprego da terceira pessoa do plural ou mesmo do singular; c. <i>a gente, muita gente, todo mundo</i>

## 2.3. Na Gramática Gerativa

### 2.3.1 Teorias sobre outras línguas românicas

Muito semelhante ao *si* e ao *se*, do italiano e do francês respectivamente, o clítico *se* do português e seus diferentes usos são alvo de diversas discussões dentro dos estudos em Gramática Gerativa, principalmente em relação à sua posição na estrutura e o estatuto desse clítico: trata-se de um mesmo elemento ou de elementos distintos, homônimos apenas? Propostas se sobrepõem, então, umas às outras na busca por explicar formalmente o que autoriza essa partícula a se comportar de modo a dar margem à interpretação impessoal ou à passiva, dependendo da construção em que está inserida.

Hyams (1986) analisa o clítico *si* do italiano em dois tipos de construções diferentes, muito semelhantes às encontradas no PB, exemplificadas abaixo:

(59) In questa pasticceria si mangia soltanto i dolci al cioccolato. (HYAMS, 1986)

‘In this pastry shop one eats (3rd per. Sing.) only chocolate cookies.’

Nesta doceria, se come apenas biscoitos de chocolate.

(60) I dolci al cioccolato si mangiano in questa pasticceria. (HYAMS, 1986)

Chocolate cookies SI eaten in this pastry shop.

‘Chocolate cookies are eaten in this pastry shop.’

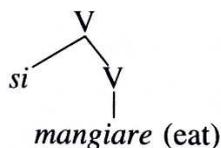
Biscoitos de chocolate se comem nesta doceria.

Nas construções tais quais (59), *si* é chamado pela autora de *si impessoal*, ela observa que há uma sentença impessoal ativa em que *si* funciona como sujeito com referência arbitrária, semelhante ao “one” do inglês. Nas sentenças semelhantes à (60), necessariamente com verbos transitivos, Hyams (1986) propõe que o objeto direto funciona como um sujeito ao passo que *si* não se comporta gramaticalmente como sujeito nem como objeto, mas como um elemento apassivador da sentença. Assim como Manzini (1868), a autora se refere a esse *si* como “si médio”.

Hyams (1986) propõe que o *si* impessoal seleciona uma raiz [+V,-N] e absorve o papel temático de sujeito, deixando um *pro* na posição de sujeito da sentença.

Além disso, para ela, para que *si* não seja motivo de fracasso de uma derivação, deve também absorver traços de Caso e traços nominais [+3ª pessoa, +plural, +masculino], por exemplo. Para que esses traços sejam satisfeitos, *si* deve ser afixado diretamente ao verbo ou como, tradicionalmente, Belletti (1982 apud HYAMS, 1986) os postula: em [Spec, INFL]. Essa posição, para a autora, é o que permite que o critério temático seja satisfeito já que VP lhe atribui papel temático de agente. A essa posição também se atribui Caso Nominativo por meio do nó de INFL (HYAMS, 1986).

Por outro lado, o tratamento dado ao “*si* médio” é bastante divergente: nessas construções *si* absorveria o papel temático de sujeito, mas receberia Caso Acusativo. Embora aparentemente o critério sintático que distinguiria as duas interpretações do *si* seja a atribuição de Caso, esse critério está intimamente ligado à posição de *si* na árvore. Para que os traços de Caso sejam satisfeitos, o “*si* médio” deve aparecer na posição de especificador de V, como no exemplo abaixo, e, portanto, absorve o papel temático de sujeito determinado por V e recebe Caso acusativo.



Contrariamente a essa ideia, Manzini e Savoia (2004), inseridos no modelo conhecido como Cartografia, defendem que os clíticos correspondem a uma categoria especializada e são inseridos diretamente em suas posições de superfície, ou seja, adjungidos a INFL. Essas categorias são ordenadas numa hierarquia universal. Argumentando a favor dessa proposta, eles refutam duas outras: a primeira, de Kayne (1989), em que o autor adota a ideia de que os clíticos são gerados como parte de um DP pleno na posição argumental e são movidos para uma posição adjunta de V ou para um núcleo funcional da sentença; e a segunda, de Sportiche (1996), em que o autor argumenta em favor de núcleos funcionais onde os clíticos seriam gerados de forma a não poderem mais ser reordenados por ali receberem Caso.

Manzini e Savoia (2004) desmontam ambas as asserções acima ao elegerem que, para a proposta de Kayne (1989) seria difícil prever ordem e quantidades fixas para os clíticos, ao passo que, para a proposta de Sportiche (1996), definir em termos de Caso seria inviável. Por se tratar de um traço não interpretável, os autores consideram

que a atribuição de caso deve ser considerada um mero diacrítico para o movimento. Essa ideia está alicerçada sobre a proposta de Brody (1997 apud MANZINI&SAVOIA, 1989) sobre a gramática estar restrita a traços interpretáveis.

Essas discussões, embora mais recentes, ainda deixam a desejar quanto a sua adequação explicativa quando comparadas a uma proposta mais abrangente como a de Uriagereka (2005). Com base nos trabalhos realizados por Uriagereka (1995), Raposo&Uriagereka (2005) propõem que existe uma propriedade da categoria F, a qual abrigaria os clíticos como solução para complexos paradigmas de colocação de clítico, como é o caso do Catalão e do PE. O que se observa nessas línguas é que não há distribuição complementar na colocação dos clíticos, ou seja, sem que haja prejuízos semânticos, um clítico pode se apresentar tanto em próclise como em ênclise em uma determinada sentença. A forma proclítica aparece nessas línguas ibéricas, em sentenças finitas, quando o verbo é precedido pela negação ou um elemento-WH. A ênclise, por sua vez, é verificada nos demais contextos e, principalmente, quando aparece um tópico entre C e V.

Os autores argumentam, então, que a única forma que explicaria essa última colocação especialmente seria a existência de uma categoria F alocada entre C e T, devido a propriedades morfofonológicas próprias dessa categoria, da seguinte maneira:

$$(61) [_{CP} \dots C [_{FP} \dots F [_{TP} \dots T [_{VP} \dots V [_{VP} \dots V \dots ]]]]]$$

Essa categoria, de acordo com Raposo&Uriagereka (2005), serviria de interface entre sintaxe e discurso uma vez que abrigaria os ditos “operadores afetivos”, ou seja, aqueles operadores que envolvem julgamentos de valor dos falantes, como alguns quantificadores, o morfema negativo “não” e advérbios aspectuais, por exemplo.

Se os clíticos não se encaixam nessa categoria de “operadores afetivos”, qual seria, então, seu lugar na estrutura? A resposta dada pelos autores é simples: os clíticos devem se mover para a Spec de F. Nas palavras dos autores:

It is not immediately obvious whether there are phrases which, not being (i.e., not involving value judgements), nonetheless target the F position – for example, a class of topics and left-dislocated elements (...) Semantically, these elements, too, are point-of-view dependent, in that, for instance, they set up a context for given speaker to comment on. Nonetheless, it is very clear that, with regard to associated clitic placement, elements of this sort behave very differently from “affective” elements. Thus, while displaced

affective phrases “trigger” proclisis, topics and “left-dislocated” phrases “trigger” enclisis. (RAPOSO&URIAGEREKA, 2005, p. 643)

Para fundamentar essa ideia, Raposo&Uriagereka (2005) ainda justificam que não é válido considerar essa categoria como uma parte de F, principalmente, pela presença de morfemas de concordância em F. Conforme foi adotada a Teoria de Verificação de Traços (CHOMSKY, 1995), essa ideia ficou ainda mais forte.

De acordo com os autores, o surgimento da teoria de “Bare” phrase-structure de Chomsky (1994) em conjunto com a ideia dos múltiplos especificadores embora tenha permitido a existência de uma quantidade indeterminada de sintagmas na periferia esquerda, não é evidência suficiente para se pensar em alocar uma quantidade indeterminada de categorias funcionais nessa periferia da sentença. Apesar de não ter a mesma fundamentação teórica, a criação de uma projeção para uma determinada função se assemelharia à proposta da cartografia de Cinque, Rizzi et al (1999, 1997).

Além disso, os autores propõem uma tipologia para selecionar e classificar as línguas de acordo com o diferente comportamento de F na derivação. Línguas do tipo A seriam, então, as mais conservadoras, em que F não seria projetado na componente LF. Já as línguas do tipo C seriam as mais radicais e permitiram os chamados afetivos na periferia esquerda da sentença, ao passo que as línguas do tipo B permitiriam as duas formas. Em línguas classificadas por eles como tipo C (RAPOSO&URIAGEREKA, 2005), como o francês, em que F não é projetado em LF, perdeu-se o movimento de sintagmas afetivos. Essa mudança diacrônica, em línguas do tipo A, cujo núcleo do foco é visível, para língua como C, deve ser analisada de forma que se postule uma marca morfológica forte, que veio a enfraquecer, em algum elemento funcional da periferia esquerda. Se cada categoria semântica observada ou propriedade semântica dos dialetos conservadores corresponde a uma categoria funcional diferente, não fica claro como ocorre a mudança. Se diferentes núcleos estão envolvidos, espera-se uma fenomenologia diferente em cada caso.

Essa ideia também é observada em DRACHMAN (1997):

Clitics move to spec-Agrs, then a non-finite verb raises over Agr to Mood/Comp. This gives enclisis for non-finite constructions in Italian and Greek. Alternatively, Galician shows clitics right-adjunction to an F (Focus-like) projection in tensed clauses, thus producing enclisis. But French has only proclisis, so in this analysis, French must lack an F-projection for the verb. (DRACHMAN, 1997, p. 235)

Em suma, o debate sobre a posição do clítico *se* na derivação de sentenças impessoais e passivas médias encontra uma bifurcação: por um lado, alguns autores defendem que *se* está adjungido a um núcleo; para outros, *se* ocupa posição de especificador. A escolha de qualquer um desses caminhos acarreta consequências para um fenômeno também muito discutido sobre essa partícula: a atribuição de Caso.

Cinque (1988), ao olhar para *si* do italiano, distingue os dois usos do *si* com base em suas características argumentais. Tem-se, então, um *si* argumental e um não argumental, indo de encontro a autores como Everett (1985 apud NUNES, 1990) que postulam uma única entrada para todas as funções desse morfema.

Para ele, uma vez que o clítico *si* é um clítico e, portanto, é um elemento lexical nominal, deve fazer parte de uma cadeia em que o caso é atribuído para satisfazer o Filtro do Caso<sup>13</sup>. Verifica-se essa ideia quando o clítico aparece em posição pós-verbal de sentenças não-finitas como em (62):

(62)\*Sarebbe meglio scoprirsi ir colpevole. (CINQUE, 1988)

\*It would be better to discover SI the culprit.

‘It would be better one to discover the culprit.’

Seria melhor se descobrir o culpado.

Pensando, por outro lado, nas características gerais do *si* impessoal, Cinque (1988) afirma que, sintaticamente, esse clítico apresenta traços [ $\pm$ argumento] e é coindexado com [NP, IP]. Em relação à atribuição de papel temático, o autor compromete-se com a seguinte asserção: Se *si* é um argumento, ele deve estar associado em todos os níveis de representação, incluindo a estrutura profunda e, portanto, como um clítico [NP, IP], ele deve estar particularmente associado com um papel temático de argumento externo atribuído pelo verbo ao [NP, IP]. Apesar de ser possível transpor essa mesma análise para o português europeu, o autor propõe que há diferenças quanto à atribuição de caso nas duas línguas, como é possível notar em (63):

(63) Seria melhor descobrir-se o culpado.

---

<sup>13</sup> O autor entende Filtro do Caso aqui tal qual proposto por Vernaud (2006), cuja proposta afirma que a todo argumento NP realizado deve ser atribuído um caso.

Diferentemente do italiano, a existência de um infinito acompanhado de *se* em PB corrobora a hipótese de que esse tipo de estrutura só é possível quando a língua possui infinitivo com marcação de pessoa em Agr. Diante disso, Cinque (1988) postula que o uso [-arg] do *si*, visto como uma forma supletiva da marcação pessoal em Agr, além de proibir o *si* em sentenças infinitivas, descarta hipóteses como as de Rizzi (1976), em que o *si* seria afetado pelo movimento do NP para dar lugar à ideia de que esse clítico é gerado em Infl ou em VP como todos os outros clíticos. Nessas sentenças, em que *si* é [-arg], o autor propõe, então, que existe um *pro* arb (arbitrário) que está em conjunção com a marca pessoal de Agr.

Já na forma como aparece no italiano, Cinque (1988) assume, pois, que *si* é um elemento nominal e, para que haja interpretação impessoal, deve fazer parte de uma cadeia com [NP, IP] à qual é atribuído caso nominativo na estrutura superficial. Essas considerações, para ele, demandam que o *si* seja, então, [+arg]. Constata-se essa ideia em sentenças como:

- (64) *Sembra non essersi lavorato a sufficienza.* (CINQUE, 1988)  
 It seems not to have SI worked sufficiently.  
 ‘It seems one not to have worked sufficiently.’  
 Parece não se ter trabalhado suficientemente.

Já em relação ao *se* apassivador, em sentenças construídas com a chamada voz média<sup>14</sup>, Cinque (1988) postula que se trata de um clítico [-arg] que suspende seu papel temático de argumento externo e o caso acusativo.

Voltando, então, ao trabalho de Raposo&Uriagereka (1996) em que os autores desenvolvem uma análise sobre construções em que o *se* impessoal aparece no PE. Com base nas ideias de Cinque (1988), os autores se apropriam da ideia de que o *se*

---

<sup>14</sup> Entendo voz média tal qual postulada por ALEXIADOU&DORON (2012, p.2): sentenças que não exigem, embora permitem, a participação de um argumento externo selecionado pela raiz (diferente do argumento interno). Quando o argumento externo é selecionado na derivação média, ele deve ser existencialmente ligado, uma vez que a voz média normalmente só deriva verbos intransitivos. No entanto, com raízes que não selecionam um argumento externo e, assim descrevendo os eventos sem um argumento externo na voz ativa, a voz média distingue-se da voz ativa por exigir um papel temático externo e atribuir a ele um papel temático padrão, de agente.

e o *si*, quando indefinidos são argumentais ao passo que o *se* genérico tem características não argumentais.

Já inseridas na proposta minimalista, as asserções de Raposo&Uriagereka (1996) se iniciam por definir a posição desses *se* nas orações. Ao passo que nos trabalhos anteriores, coloca-se o clítico como gerado em [NP, IP], para essa proposta, os autores negam que esse *se* sujeito esteja em [Spec, Infl] ou que ele seja ligado a uma categoria vazia nessa posição. Quando em posição pré-verbal, *se* é tampouco um tópico em [Spec, Top] ou uma sentença Wh em [Spec, F].

A atribuição de caso nos trabalhos de Raposo&Uriagereka (1996) já é pensada, portanto, dentro da ideia de verificação de traços de Chomsky (1995). Eles defendem que a atribuição de caso estrutural está intimamente ligado aos traços- $\Phi$ , formando um único traço formal, o que corrobora a antiga ideia de que Caso e *agreement* são diferentes manifestações de uma única relação básica.

Por fim, a análise de Dobrovie-Sorin (1998), cuja base também é o trabalho de Cinque (1988), abraça a ideia de que *se* como indeterminador e como apassivador são apenas homônimos, pois apresentam comportamentos deliberadamente distintos. Sua análise está bastante sustentada pela atribuição de caso que diferenciaria cada uso de *se* uma vez que, para ela, enquanto *se* reflexivo (não tratado aqui) recebe caso acusativo, os demais recebem caso nominativo. Segundo a autora, que também utiliza dados do italiano, *si* nominativo, liga uma categoria vazia em posição-A (sujeito). Isso acontece porque *si* acusativo é uma anáfora.

O outro *si* que recebe caso nominativo sob essa perspectiva é o das sentenças em voz média. A análise produzida por Dobrovie-Sorin (1998) caminha no sentido de estabelecer uma equivalência entre os processos de passivização de sentenças inergativas e transitivas. Isso obriga o *si* passivo a projetar um objeto direto. Assim, como qualquer outro clítico de objeto, *se* liga uma categoria vazia na posição de objeto direto. Uma consequência disso é que esse tipo de clítico nominativo não conta como argumento, mas como expletivo ou argumento referencial, a depender do verbo com o qual se combina. Outra possibilidade é que o papel temático atribuído seja o da posição-A ao qual o clítico está associado. Com isso, o traço [arg] de *si*, que não está presente em nenhum clítico nominativo, se caracteriza como um mero rótulo para o *si* passivo.

### 2.3.2 Teorias sobre o PB

Para o PB, Nunes (1990), em seu célebre trabalho sobre o “Famigerado Se”, analisa construções com *se* indeterminador e *se* apassivador sincrônica e diacronicamente a fim de definir um estatuto teórico para essas construções e verificar os processos de mudança pelos quais essas construções passaram. As sentenças (65) e (66) exemplificam, respectivamente, essas construções:

(65) Alugam-se casas. (NUNES, 1990)

(66) Aluga-se casas. (NUNES, 1990)

(67) A casa foi alugada. (NUNES, 1990)

Em oposição às passivas perifrásticas, como (67), ele (1990) as agrupa sob a designação de passivas pronominais. Antes de dar início à sua análise, o autor considera o que foi dito por Galves (1987) e Everett (1985) anteriormente, indicando seus pontos de discordância.

A problemática em torno das ideias de Everett (1985) gira em torno de sua pressuposição de que há apenas uma entrada lexical para os vários usos do *se* (reflexivo, ergativo, indeterminador, apassivador e inerente) em todas as línguas românicas. Para sustentar essa proposta, o autor afirma que *se* pode ser gerado em qualquer lugar e movido para qualquer lugar. *Se* indeterminador seria, então, gerado sob INFL, onde recebe papel temático do argumento externo, caso nominativo e é ligado pelo expletivo que ocupa posição de sujeito. Sua cliticização aconteceria em Forma Fonética e seus traços [+Caso] e [+Argumento] seriam selecionados. Por outro lado, *se* apassivador por estar associado a sujeitos derivados, se cliticizaria na sintaxe e perderia seu caráter anafórico. O clítico receberia papel temático de argumento externo em INFL e caso acusativo, deixando o caso nominativo para o argumento interno que concorda com o verbo. Galves (1986) já julgava muito altos os custos teóricos dessa análise principalmente pela dissociação de caso e Argumento. Nunes (1990), assim como Galves (1986), discorda dessa análise, mas tampouco reforça a proposta da autora, embora lhe pareça mais razoável.

Baseada na “hipótese da indexação livre” de Borer (1985 apud NUNES, 1990), Galves (1986) considera que *se*, por ser um clítico lexical que, como os demais,

se associa a posições argumentais, mas jamais pode ser argumento ou desempenhar a função de identificar categorias vazias nessas posições. Assim, em sentenças com *se* apassivador, o clítico realizaria o papel temático de argumento externo e o SN posposto ao verbo, de argumento interno. Contudo, somente o SN receberia Caso (nominativo) por estar coindexado com AGR; *se*, por ser a realização da função temática atribuída pelo verbo ao argumento externo, não teria necessidade de receber Caso. Sentenças com *se* indeterminador, em contrapartida, apresentariam uma cadeia que receberia papel temático de argumento interno e caso nominativo, e *se* seria coindexado com AGR, tornando *PRO* sujeito de uma anáfora pronominal. Para Nunes (1990), essa proposta é discutível tendo em vista que sentenças com *se* apassivador seriam geradas sem sujeito e que ela desconsidera, assim como Everett (1985), a relação anafórica do clítico também nessas sentenças.

O autor reafirma, então, a necessidade de uma análise que resolva essas questões teóricas. Sua proposta se origina, assim como as dos outros autores, do trabalho de Chomsky (1981), em que ele postula que *se* absorve caso acusativo, o SN posposto ao verbo recebe caso nominativo por constituir um elo de CADEIA e a posição do sujeito é detematizada.

O primeiro questionamento levantado por Nunes (1990) diz respeito à atribuição de caso ao clítico *se*. Em contraposição aos trabalhos de Galves (1986) e de Everett (1985), o autor afirma que trabalhar com a ambivalência entre nominativo e acusativo, nesse caso, é pouco produtivo. Entretanto, ele segue a análise, negando a possibilidade de *se* receber caso nominativo por se tratar de um clítico morfologicamente marcado como forma acusativa. Tampouco seria interessante aos olhos do autor admitir que *se* recebe caso acusativo por default já que isso não seria verdade em sentenças como (68), em que há uma locução verbal:

(68) Pode-se alugá-las. (NUNES, 1990)

Diante disso, para Nunes (1990) só seria possível aceitar que *se* absorve caso acusativo em sentenças como (68) em que há concordância entre o verbo e o argumento interno. Esse contexto seria também o único em que a absorção de papel temático de argumento externo pelo *se* não esbarraria em problemas teóricos. De acordo

com Jaeggli (1986 apud NUNES, 1990), passivas perifrásticas admitem que o afixo de particípio “\_do” absorva o papel temático de argumento externo. Analogamente, *se* também o faria devido às semelhanças de comportamento entre sentenças passivas pronominais e passivas perifrásticas. Nunes (1990) questiona, então, a possibilidade de dois morfemas tematicamente diferentes absorverem um mesmo papel temático já que isso violaria o Critério Temático<sup>15</sup>. Para esclarecer essa questão, o autor defende que não é necessário que clíticos anafóricos estejam vinculados à realização de funções temáticas uma vez que em sentenças com *se* inerente como (69), isso não é observado:

(69) João *se* arrependeu do crime. (NUNES, 1990)

Logo, obedecendo ao Princípio de Atribuição de Papel Temático<sup>16</sup> (CHOMSKY, 1981), tanto *se* quanto o afixo de particípio podem absorver papel temático apenas quando absorvem caso. Nunes (1990) conclui, então, que:

Assumindo que são gerados em INFL, quando *se* e “\_do” absorvem o papel temático reservado ao argumento externo, o verbo torna-se incapaz de atribuir caso acusativo, pois não marca tematicamente a posição de sujeito. Dessa forma, o argumento interno de passivas perifrásticas ou de construções com *se* apassivador deve estar associado ao caso nominativo. Se *se* e “\_do” absorvessem nominativo, não haveria em estrutura-S, de acordo com a Generalização de Burzio, caso disponível para o argumento interno. (NUNES, 1990, p. 39)

Paralelamente, Nunes (1990) observa sentenças como (70) e (71) para verificar o que acontece com *se* indeterminador:

(70) Tinham-se examinado os papeis. (NUNES, 1990)

(71) Tinha-se examinado os papeis. (NUNES, 1990)

Para ele, por ser [+V, +N], o afixo de particípio recebe caso acusativo do verbo *examinar* em sentenças como (70). Com isso, o SN *os papeis* recebe caso nominativo devido à formação de uma CADEIA com o expletivo, o que confirma a concordância. O verbo auxiliar *ter*, então, atribui caso acusativo a *se*, que absorveu

<sup>15</sup> (cf. CHOMSKY, 1981)

<sup>16</sup> “Um papel temático é atribuído a um elemento de uma cadeia que ocupe uma posição teta e” (CHOMSKY, 1981, p. 37 apud NUNES, 1990, p. 37)

papel temático do argumento externo. Já em (71), há um pronome nulo na posição de sujeito que absorve papel temático de argumento externo e caso nominativo. Consequentemente, o verbo *ter* atribui caso acusativo ao argumento interno e o clítico *se* não absorve caso pois se o auxiliar lhe atribuísse, haveria uma CADEIA sem caso. Portanto, *se* indeterminador deve apenas satisfazer seus traços anafóricos. Diante disso, o autor assume que, por ser um morfema clítico com traços [-V, +N], *se* não precisa satisfazer o Filtro dos Casos, deixando que a atribuição de Caso seja regida pelo Princípio de Atribuição de Papel Temático. Ademais, se *se* é gerado sob INFL, a atribuição de papel temático se dá em estrutura-D e o clítico enquanto apassivador só pode absorver o papel temático do argumento externo.

Diante disso, tem-se então:

(72) (Expl) Alugam-se<sub>pe</sub> casas<sub>pi</sub>. (NUNES, 1990)

(73) pro<sub>pe</sub> Aluga-se<sub>p0</sub> casas<sub>pi</sub>. (NUNES, 1990)

Finalizando sua proposição, Nunes (1990) explora as propriedades formais que indeterminam a referência do argumento externo nas sentenças passivas pronominais. Para tanto, ele se embasa nas seguintes afirmações:

- a) anáforas lexicais detentoras do papel temático de argumento interno, quando ligadas por uma categoria detentora do papel temático de argumento externo, assumem a referência dessa categoria;
- b) anáfora lexicais detentoras do papel temático do argumento externo, quando ligadas por uma categoria detentora do papel temático do argumento interno, assumem referência indeterminada. (NUNES, 1990, p. 62 e p. 63)

A generalização (b) explica o que ocorre em sentenças com *se* apassivador por explicarem o que acontece em passivas perifrásticas: enquanto a relação de ligação entre duas categorias se dá por meio de traços de concordância de gênero e número com o afixo de particípio, com o *se*, são os traços de concordância de número e pessoa.

Essas generalizações não contribuem para o entendimento do processo de indeterminação com *se* indeterminador, o qual é resultado da ligação entre o pronome expletivo *pro* que ocupa a posição de sujeito e o clítico. Nunes (1990) defende ainda que, assim como um elemento PRO, o pronome nulo precisa absorver papel temático de argumento interno para que essa referência seja arbitrária.

Apesar das diferenças formais, ele esclarece que a interpretação da referência indeterminada é a mesma tanto para sentenças com *se* indeterminador como em sentenças com *se* apassivador. Essa é a razão pela qual sentenças como (74) frequentemente são vistas como ambíguas:

(74) Tinha-se alugado uma bela casa. (NUNES, 1990)

Ele, então, analisa o *se* da seguinte maneira: o clítico *se* na função de apassivador absorve o papel temático e caso acusativo do argumento externo, e é ligado por um elo de cadeia temática que envolve o argumento interno, indeterminando a referência do argumento externo; quanto ao *se* indeterminador, o autor propõe que esse clítico só tem necessidade de satisfazer seus traços [+an, -pro] e, portanto, é ligado por um pronome referencial nulo, portador do papel temático do argumento externo e de caso nominativo, configurando uma relação anafórico-pronominal de referência indeterminada.

Melo (2012) também defende, a partir do trabalho de Naro (1976 apud MELO, 2012) que o *se* indeterminador surgiu a partir da reanálise das construções com *se* passivo. No entanto, a autora, fundamentada na proposta de Raposo & Uriagereka (1996), também toma como objeto de estudo o *se* e se propõe a analisar a ordem do DP-argumento interno nas sentenças com *se*. A autora observa que sentenças em o *se* passivo é utilizado, de acordo com Duarte (2003 apud MELO, 2012), apresentam concordância entre verbo e argumento interno plural, como em (75):

(75) Os artigos publicaram-se propositadamente no último número da revista. (DUARTE, 2003 apud MELO, 2012)

Nessas sentenças, a concordância se explica pelo de fato de o argumento interno exercer a função sintática de sujeito. Além disso, a autora também afirma que o argumento externo com papel temático de agente não pode ser expresso por meio de um sintagma preposicionado, como em (76):

(76) \*Os artigos publicaram-se propositadamente no último número da revista pelo editor. (DUARTE, 2003 apud MELO, 2012)

Dessa forma, ao clítico é atribuído o papel temático externo, ficando em [SPEC-VP] e ele recebe caso acusativo, o que reserva ao argumento interno o nominativo e permite que seja tratado como sujeito da sentença. Duarte (2003 apud MELO, 2012) defende ainda a existência de um *se* ambíguo, o qual estaria presente em sentenças com verbo transitivo direto no singular e argumento interno também no singular. Estruturas como essa, exemplificadas em (77) e (78), podem receber interpretação tanto passiva quanto indeterminada:

(77) Descobriu-se uma fuga no reactor nuclear. (DUARTE, 2003 apud MELO, 2012)

(78) Assaltou-se uma carrinha de transporte de valores. (DUARTE, 2003 apud MELO, 2012)

Contudo, em todas as sentenças o argumento externo é o sujeito: quando há *se* indeterminador, ele é um argumento, recebe papel temático e ocupa a posição de sujeito, com traços [+humano], [3ª pessoa], [+clítico], [+nominativo] e [-definido]; e quando há *se* passivo, ele funciona tal qual um morfema capaz de mudar a diátese da voz verbal, mas não é um argumento. (MELO, 2012)

Depois de observar os dados diacrônicos de jornais cariocas do século XIX, Melo (2012) observa em seus resultado que, em sentenças indefinidas, genéricas ou ambíguas, a preferência é pela posposição do DP, enquanto que em sentenças passivas e reflexivas, a preferência é pela anteposição. Além dessa constatação principal, a autora também observou as seguintes características nas sentenças coletadas:

- i. A anteposição do DP em sentenças com *se* indefinido só é possível quando o traço de animacidade desse argumento for [-animado]
- ii. O tempo verbal é determinante para a interpretação genérica quando houver anteposição.
- iii. Há aumento no uso de argumento nulo em sentenças com *se* indefinido.

- iv. Informações novas tendem a estar no fim da sentença, logo, sintagmas pesados são DPs-argumentos internos pospostos.

No capítulo seguinte, essas hipóteses serão testadas a fim de verificar se esse comportamento se mantém constante nos dados coletados nesta pesquisa.

Por fim, para sumarizar as principais ideias dos autores apresentados acima, o quadro 5 apresenta, por autor, a classificação dada ao clítico *se* e suas características na Gramática Gerativa:

**Quadro 5 – Síntese das principais ideias de Gramática Gerativa sobre o *se***

<b>Autor</b>	<b>Classificação</b>	<b>Características</b>
Hyams (1986)	e. <i>Si</i> impessoal; f. <i>Si</i> médio.	e. <i>Si</i> funciona como sujeito com referência arbitrária, semelhante ao “one” do inglês; f. <i>Si</i> não se comporta gramaticalmente como sujeito nem como objeto, mas como um elemento apassivador da sentença.
Cinque (1988)	d. <i>Si</i> argumental; e. <i>Si</i> não-argumental.	a. Apresenta traços [ $\pm$ argumento] e é coindexado com [NP, IP]. Como um clítico [NP, IP], ele deve estar associado com um papel temático de argumento externo atribuído pelo verbo. b. Clítico [-arg] que suspende seu papel temático de argumento externo e o caso acusativo.
Raposo&Uriagereka (1996)	a. <i>Si</i> indefinido; b. <i>Si</i> genérico.	a. Argumental; b. Não argumental. (CINQUE, 1988)
Dobrovie-Sorin (1998)	a. <i>Si</i> nominativo; b. <i>Si</i> acusativo.	a. Liga uma categoria vazia em posição-A; b. Anáfora.
Everett (1985)	Apenas uma entrada lexical para os vários usos do <i>se</i> (reflexivo, ergativo, indeterminador, apassivador e inerente) em todas as línguas românicas.	<i>Se</i> indeterminador: gerado sob INFL, onde recebe papel temático do argumento externo, caso nominativo e é ligado pelo expletivo que ocupa posição de sujeito. Sua cliticização aconteceria em Forma Fonética e seus traços [+Caso] e [+Argumento] seriam selecionados. <i>Se</i> apassivador: <i>se</i> cliticizaria na sintaxe e perderia seu caráter anáforico.
Galves (1986)	a. <i>Se</i> indeterminador; b. <i>Se</i> apassivador.	a. As sentenças apresentariam uma cadeia que receberia papel temático de argumento interno e caso nominativo, e <i>se</i> seria coindexado com AGR, tornando <i>PRO</i> sujeito de uma anáfora pronominal; b. O clítico realizaria o papel temático de argumento externo e o SN posposto ao verbo, de argumento interno.
Nunes (1990)	a. <i>Se</i> indeterminador; b. <i>Se</i> apassivador.	a. Esse clítico só tem necessidade de satisfazer seus traços [+an, -pro] e, portanto, é ligado por um pronome referencial nulo, portador do papel temático do argumento externo e de caso

		<p>nominativo, configurando uma relação anafórico-pronominal de referência indeterminada;</p> <p>b. Absorve o papel temático e caso acusativo do argumento externo, e é ligado por um elo de cadeia temática que envolve o argumento interno, indeterminando a referência do argumento externo.</p>
Melo (2012)	<p>a. <i>Se</i> indeterminador;</p> <p>b. <i>Se</i> apassivador.</p>	<p>a. É um argumento, recebe papel temático e ocupa a posição de sujeito, com traços [+humano], [3ª pessoa], [+clítico], [+nominativo] e [-definido];</p> <p>b. Funciona tal qual um morfema capaz de mudar a diátese da voz verbal, mas não é um argumento.</p>

### 2.3.4 Os dados de mudança

A essa análise formal, Nunes (1990) soma uma análise variacionista, tal qual propõem Kato e Tarallo (1989), a partir de 1374 dados coletados agrupados em quatro corpora. O objetivo desse outro olhar vai além do rastreamento do percurso diacrônico das sentenças envolvendo *se* indeterminador e *se* apassivador: o autor deseja consolidar que uma análise puramente formal não dá conta de todos os aspectos envolvidos nas mudanças pelas quais passaram essas sentenças.

Composto por diários, documentos e cartas do período compreendido entre 1555 e 1989, o corpus diacrônico verificou que o português antigo sofreu um processo de reanálise sintática e, conseqüentemente, o autor observou uma das hipóteses deste trabalho: no PB, a antiga construção com *se* apassivador se restringe apenas à escrita. Essa conclusão foi possível na medida em que o autor comparou os dados diacrônicos a dados sincrônicos de entrevistas do Banco de Dados da PUCSP, de entrevistas do português europeu e de reportagens da revista *Veja* dos anos 1988 e 1989.

A primeira constatação feita por Nunes (1990) ao olhar para seu conjunto de dados se refere ao surgimento da passiva pronominal sem concordância. Até o século XVIII, a discordância entre verbo e argumento interno não ultrapassou 20% dos dados. Contudo, essa porcentagem subiu para 62% no século XIX e chegou a 84% no século XIX, tornando-se a forma canônica. De acordo com o autor, a explicação para esse

fenômeno reside em um processo de reanálise sintática<sup>17</sup>: a categoria vazia ilustrada em (84) deixou de ser interpretada como um expletivo, como acontece em (83), para ser vista como um pronome referencial. Em decorrência disso, o SN posposto, portando caso acusativo, é agora interpretado como objeto direto. Ele também verificou que a discordância ocorre em frequência muito maior quando o argumento interno está posposto ao verbo.

Retomando a ideia de que a interpretação da referência indeterminada em *se* indeterminador e *se* apassivador têm diferenças puramente formais, é possível afirmar que:

A reinterpretção da categoria vazia na posição de sujeito não provoca alteração na referência de CADEIA temática envolvendo o papel temático de argumento externo: pelo contrário, mantém-se até uma das propriedades que diferencia construções com *se* apassivador de passivas perifrásticas: o traço [+hum] (mais humano) (cf. NARO, op. cit., p. 781); PIMENTA-BUENO (1979, P. 120); IKEDA (1977, p. 21)). Construções com *se* apassivador só tomam como referência para o argumento externo um elemento com traço [+hum], ao contrário de passivas perifrásticas. (NUNES, 1990, p. 94)

A consequência primeira dessa reanálise, de acordo com o autor, que permitiu construções com pronome referencial nulo na posição de sujeito indeterminado pelo clítico *se*, foi a expansão dessas construções para que fossem admitidos quaisquer tipos de verbos. Logo, verbos transitivos diretos usados intransitivamente (79), verbos intransitivos (80), verbos transitivos preposicionados (81), verbos de ligação (82), verbos ergativos (83) e verbos em construções passivas perifrásticas (84) passaram a ter a referência do sujeito indeterminada por *se*, nessa respectiva ordem. Seguem abaixo alguns exemplos dessas construções:

(79) Come-se muito no inverno. (NUNES, 1990)

(80) Trabalha-se bastante neste lugar. (NUNES, 1990)

(81) Precisa-se de empregados. (NUNES, 1990)

(82) É-se feliz quando se é jovem. (NUNES, 1990)

(83) Chegou-se tarde à reunião. (NUNES, 1990)

---

<sup>17</sup> O autor entende reanálise sintática tal qual propõe Langacker (1977, p. 58 apud NUNES, 1990): “a change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation. Reanalysis may lead to changes at the surface level (...), but these surface changes can be viewed as the natural and expected result of the functionally prior modifications in rules and underlying representatios.”

(84) Aqui se é visto por todos. (NUNES, 1990)

Com isso, duas outras mudanças linguísticas antagônicas encontraram espaço na língua: o apagamento de *se* indeterminador em sentenças finitas e sua inserção em sentenças infinitas (NUNES, 1990). Quanto ao primeiro fenômeno, Galves (1987) já chamava a atenção para sentenças como (85), provenientes de (86) e (87), que apesar de ter referência determinada no português europeu, mantinham a interpretação indeterminada mesmo sem o *se* em PB:

(85) Nos nossos dias, não usa mais saia. (NUNES, 1990)

(86) Nos nossos dias não usam mais saia. (NUNES, 1990)

(87) Nos nossos dias não se usa mais saia. (NUNES, 1990)

Para explicar formalmente o que permite a realização dessas construções, Nunes (1990) se vê diante de duas hipóteses: uma justificada pelo paradigma de concordância do PB que permite “ele/eles fala” e outra, pela supressão de clíticos em geral no PB.

Já nos séculos XVII e XVIII, o autor observou algumas ocorrências em que o clítico sofre elisão em sentenças em que há coordenação, como (88):

(88) Não se comprou nem vendeu um só carro. (NUNES, 1990)

No entanto, no século XIX, quando as estruturas com o expletivo foram substituídas pelas que continham *pro* referencial, a supressão do clítico se dissemina para estruturas sem coordenação. Pouco mais de um século depois, o autor encontrou ocorrências desse tipo em 79% de suas entrevistas.

Apesar de rápida e intensa essa mudança, Nunes (1990) verificou que se trata de uma estrutura sensível à influência da escolaridade uma vez que, enquanto 93% dos informantes que haviam completado o 1º grau produziam esse tipo de sentença sem *se*, apenas 47% dos que tinham 3º grau completo apagavam o clítico.

Por outro lado, a segunda mudança linguística mencionada não ficou estigmatizada pelos falantes escolarizados, apesar de frequentes censuras da Gramática Tradicional.

A partir do século XIX, em que as construções com *se* indeterminador se tornaram majoritárias, esse clítico passou a ser inserido em sentenças com verbos no infinitivo e no gerúndio como (89) e (90), respectivamente:

(89) É impossível se achar lugar aqui. (NUNES, 1990)

(90) Está se sujando muito a rua. (NUNES, 1990)

Detendo-se especificamente em sentenças do tipo (62), Nunes (1990) explica:

Até o século XVIII estruturas infinitivas, como as finitas, privilegiavam as construções com *se* apassivador (com concordância). Entretanto, quando no século XIX, as passivas pronominais com discordância se tornam majoritárias, as construções infinitivas dão um salto considerável, ultrapassando as finitas na preferência pela construção com *se* indeterminador. Isso pode constituir um indício de que é nesse momento a se dar o fenômeno da inserção de *se* junto a infinitivos, pois a inserção deve ter fornecido combustível adicional para a mudança que substituíra construções infinitivas com concordância por construções sem concordância. Dito de outro modo, no século XIX o português brasileiro, ao invés de PRO, começa a preferir a categoria *pro* para ocupar a posição de sujeito de infinitivas. (NUNES, 1990, p. 115)

O autor chama, portanto, esse *se* de *Se Sarney* fazendo referência à publicação de Josué Machado na seção “Língua” da revista *Imprensa* de 1989 (nº 25), do qual o trecho abaixo faz parte:

“ ‘Como se evitar a decomposição do país.’ Ganha um prêmio quem descobrir a função do ‘se’ desse título publicado no Caderno de Sábado do Jornal da Tarde de algumas semanas atrás. O prêmio será uma página inteira da obra *Marimbondos de Fogo*, do dr. José Sarney, em francês. Por que esse prêmio? Porque esse ‘se’ pode ser batizado com o nome de Sarney. Crueldade, talvez, mas é um *se* muito em evidência, sem função nem utilidade. Para não dizer ridículo e errado. Mas isso não fica bem dizer. O ‘se’ Sarney é aquele que vem grudado à maioria dos verbos no infinitivo, sem que se saiba para que serve.’

De encontro às ideias de Nunes (1990), Cavalcante (2006), em seu estudo diacrônico sobre o uso do *se* infinitivo, notou expressiva presença dessas construções já no século XVI e uma queda, até mesmo nos contextos sintáticos favorecedores da estrutura, no século XIX. A autora se apoiou nos dados encontrados para defender que

as diferenças encontradas entre Português Clássico (doravante PCI), Português Europeu e Português Brasileiro confirmam que se trata de três gramáticas distintas.

Dentre esses contextos favorecedores, destacaram-se principalmente a presença de um agente da passiva, a posição do argumento interno em relação ao verbo infinitivo e a oração preposicionada. Essa última se justificaria principalmente nas propostas de Raposo (1987 apud CAVALCANTE, 2006) e Cinque (1988). Seguem abaixo exemplos de sentenças em que esses contextos favorecedores aparecem:

A posição sujeito de infinitivo, de acordo com esses autores, é licenciada por Agr não finito, que deve ser marcado com Caso devido à presença de um atribuidor de Caso ao Comp da oração infinitiva. Dessa forma, a preposição seria o elemento atribuidor de Caso. Além disso, tanto no PCI quanto no PE, *se* acompanha exclusivamente verbos transitivos no infinitivo, o que não parece ser uma regra para o PB. O tipo de material analisado também parece influenciar de alguma forma para o PB: aproximadamente 50% dos dados com *se* acompanhando verbos infinitivos estava em dados escritos contra 20% em dados de fala.

Em relação à variação, Cavalcante (2006) notou que nos dados do PCI, a sentença de *se* com infinitivo variam com as sentenças passivas, pois se trata de um *se* passivo, enquanto que no PE, a variação se dá entre *se* e *PRO* na posição de sujeito infinitivo. Já no PB, há variação entre *se*, *a gente*, *você* e *pro* na posição de sujeito de infinitivo. Isso acontece porque o Agr defectivo no PB, embora licencie uma posição sujeito, não é capaz de interpretar o referente dessa posição e, com isso, é possível ter tanto *pro* como *se*, *a gente* ou *você* no caso de sujeitos arbitrários (CAVALCANTE, 2006). O PE, por sua vez, possui dois tipos de infinitivo: o Infinitivo Flexionado e o Infinitivo Não-Flexionado, sendo que *se* está associado ao primeiro. Por essa razão *se* surge como uma estratégia para desambiguar o referencial do sujeito, tal como os pronomes com referência indeterminada.

Negrão & Viotti (2015), recentemente, retomaram as investigações sobre o *se* a fim de mostrar que as mudanças no sentido tanto de perda quanto de inserção do clítico precisam ser vistas como um fenômeno também de natureza semântica uma vez que a interpretação da eventualidade das sentenças é alterada. Para elas, o processo linguístico que impulsionou as mudanças relativas à indeterminação do sujeito no PB é a expansão da categoria de verbos que permitem alternância causativa, processo em que

ocorre uma alteração na diátese verbal de modo que o número de argumentos e de papéis temáticos atribuídos pelo verbo é reduzido, como nas sentenças abaixo:

(91) João quebrou o vidro. (NEGRÃO&VIOTTI, 2015)

(92) O vidro se quebrou. (NEGRÃO&VIOTTI, 2015)

(93) Ana abriu a porta. (NEGRÃO&VIOTTI, 2015)

(94) A porta (se) abriu. (NEGRÃO&VIOTTI, 2015)

(95) A esquadra inimiga afundou o navio. (NEGRÃO&VIOTTI, 2015)

(96) O navio (se) afundou. (NEGRÃO&VIOTTI, 2015)

Nas sentenças ergativas, como (94) e (96) em que o verbo é monoargumental e pede apenas um objeto direto, os eventos ocorreram espontaneamente, sem que um agente estivesse envolvido na mudança de estado da qual se fala. A possibilidade de alternância para uma forma como (93) e (95) biargumental, está associada à não conceitualização da causa que originou o evento, o que permite que ele seja interpretado como um evento espontâneo. A mudança que vem ocorrendo, portanto, expande essa possibilidade para verbos que não designam uma mudança de estado como (97), (98) e (99) (NEGRÃO&VIOTTI, 2015):

(97) Meu jardim destruiu todo com a reforma. (NEGRÃO&VIOTTI, 2015)

(98) Eu só vou trocar o carpete depois que a casa acabar de pintar.  
(NEGRÃO&VIOTTI, 2015)

(99) Não tem nenhum concurso que anulou por causa de mérito.  
(NEGRÃO&VIOTTI, 2015)

A diferença desse novo grupo de verbos está na sua impossibilidade de se associar ao clítico *se*. Diante disso, as autoras defendem que

A presença do clítico *se* indica que o evento ao qual o verbo se refere está sendo conceitualizado como um evento energético, ou seja, como um evento causado por alguma fonte de energia, ainda que essa fonte seja indeterminada, arbitrária ou vaga. A ausência do clítico, por outro lado, significa que o que está sendo conceitualizado é apenas o resultado de um evento, não sua origem ou causa. (NEGRÃO&VIOTTI, 2015, p. 51)

Assumir essa hipótese significa afirmar que o uso do clítico não indica uma marcação morfológica da forma anticausativa da alternância verbal, mas uma indicação da supressão do participante relativo à causa do evento. Logo, para as autoras, a diminuição do uso do clítico *se* está associada a essa mudança semântica.

Por outro lado, essa expansão, de acordo com Negrão&Viotti (2008), pode ter suas origens no contato do PB com o quimbundo e, à medida que essa mudança se consolida na língua, afasta-se de línguas românicas como o PE e aproxima-se de línguas africanas.

Além dessas mudanças, é importante apontar que Kato&Tarallo (1986) já indicaram outras formas de indeterminação que vêm emergindo no PB. Com base nos estudos de Laberge & Sankoff sobre o inglês canadense falado na região de Montreal, os autores apontaram o uso do pronome de tratamento *você*, sem menção à pessoa com que se fala, como forma de indeterminar o agente da ação. Nas sentenças abaixo, vê-se exemplos desse recurso no francês – que provavelmente motivou a mudança no inglês canadense –, no inglês e no PB:

(100) Puis ça, moi j'ai pour mon dire, on peut pas se mettre dans la peau.

(KATO&TARALO, 1986)

Then that, I have to tell you, one can not put ourselves in the skin.

'Then that, I have to tell you, we can not put ourselves in the skin.'

Então, eu tenho que dizer, não se pode entrar na pele.

(101) Anything you can do. (KATO&TARALO, 1986)

Qualquer coisa você pode fazer.

'Qualquer coisa que você pode fazer.'

(102) Depois que você termina o comércio, você vai na área residencial.

(KATO&TARALO, 1986)

Os estudos preliminares dos autores mostraram que essa forma é preferida na fala, seguida pelo uso da categoria vazia, como em (103), pela forma *a gente*, como em (104), pelo pronome *nós*, como em (105), e, por último, pelo pronome *eu*, como em (106):

(103) No Brasil, fala português. (KATO&TARALO, 1986)

(104) E se a gente falar que não tem? (KATO&TARALO, 1986)

(105) Em primeiro lugar, nós temos identificado claramente uma nova consciência crítica da classe média. (KATO&TARALO, 1986)

(106) Se eu pego aquela rua ali, eu chego mais rápido. (KATO&TARALO, 1986)

O surgimento dessas formas, segundo Kato&Tarallo (1986) é mais uma das evidências na mudança no paradigma de clíticos e de pronomes do PB que contribuem para a tese de Pereira (1981 apud KATO&TARALLO, 1986) de que os clíticos estão sendo substituídos por suas formas tônicas, por DPs plenos e pela ausência de anáfora. Esse processo, no entanto, se distingue daquele verificado em Montreal na medida em que o francês mantém a forma indefinida *on* na língua falada e essa forma está competindo apenas com formas pronominais. No PB, a situação é oposta: o clítico foi perdido e as formas pronominais estão competindo com *pro*, como em (103).

A partir das constatações dessa pesquisa, Duarte, Kato & Barbosa (2001) investigaram, então, quais recursos são utilizados para a indeterminação do sujeito em PE e PB. Os resultados encontrados estão representados abaixo, nas figuras 1 e 2:

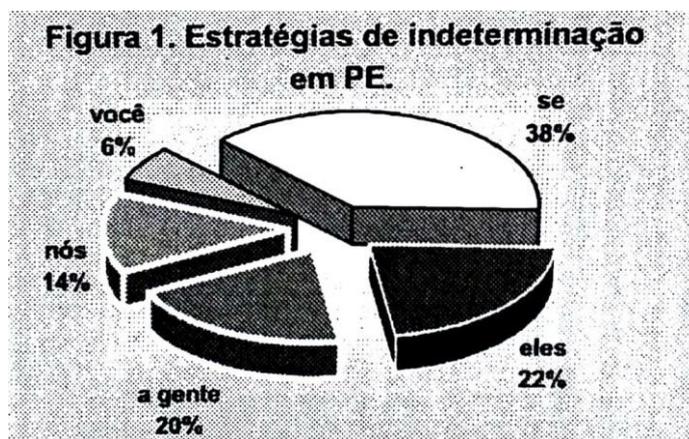


Figura 2 – Estratégias de indeterminação em PE (Retirada de Duarte, Kato & Barbosa (2001))

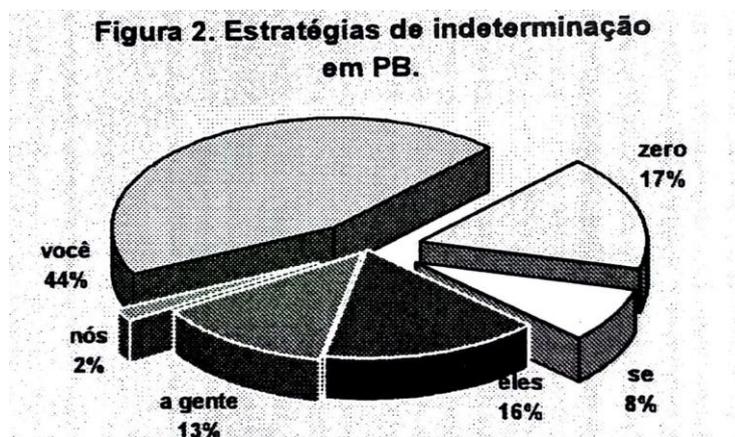


Figura 3 – Estratégias de indeterminação em PB (Retirada de Duarte, Kato & Barbosa (2001))

Excluídas as construções com verbos na terceira pessoa do plural, por sua baixa frequência e pela sua especificidade, as autoras notam que o *se* ainda é a estratégia de indeterminação mais utilizada em PE, ao passo que no PB é a forma *a gente* a mais utilizada. Duarte (1995 apud DUARTE, KATO & BARBOSA, 2001) relaciona o aparecimento dessas construções pronominais nominativas à perda do sujeito nulo no PB já que também no contexto de sujeitos arbitrários o PB preenche mais o sujeito por meio de construções pessoais, como *você* e *a gente*.

Outro dado interessante trazido pelas autoras foi o estudo de formas de indeterminação em gêneros textuais específicos. Por exemplo, ao olharem para receitas culinárias, elas perceberam que o uso de *se* é bastante produtivo e sentenças como (107) são muito comuns:

(107) Limpam-se as cebolas e cortam-se às rodelas não muito finas, limpam-se os pimentões das sementes e cortam-se também os tomates às rodelas não muito finas sem tirar as sementes e põe-se tudo num tacho ficando a cebolas no fundo. (DUARTE, KATO & BARBOSA, 2001)

Dentre as 60 receitas escritas em PE analisadas pelas autoras, 16 usavam o *se* passivo como estratégia de indeterminação, como em (107), 36 usavam o imperativo, como em (108), e apenas uma, o infinitivo, como em (109):

(108) Coza os aspargos em água temperada com sal. (DUARTE, KATO & BARBOSA, 2001)

(109) Aquecer o óleo num tacho. (DUARTE, KATO & BARBOSA, 2001)

Em contrapartida, nas 58 receitas escritas em PB estudadas, 57 utilizaram o imperativo para indeterminar o sujeito e apenas uma utilizava indicativo com modal como (110):

(110) Servidas como aperitivo, as ostras devem ser mergulhadas no molho na hora de comer. Acompanha vinho branco e pão de centeio com manteiga. (DUARTE, KATO & BARBOSA, 2001)

No entanto, ao buscarem receitas em livros brasileiros antigos republicados sem editoração, o uso do *se* como estratégia de indeterminação aparece igualmente produtivo como em PE. Nas palavras das autoras:

podemos dizer que o PB dos livros de cozinha, mais conservador, espelha a norma portuguesa, de uso produtivo de construções passivas com *se*, em geral com concordância. Já na imprensa, mais inovadora, a estratégia com *se* é evitada e substituída pelo imperativo, evitando-se as formas de indeterminação com sujeito pronominal nominativo, que marca os gêneros mais próximos da oralidade. (DUARTE, KATO & BARBOSA, 2001)

Portanto, elas defendem que há uma variação em função da modalidade e do gênero do texto e a ausência de *se* nos gêneros mais próximos da oralidade sugerem que o PB parece ter perdido a forma passiva de *se*, mantendo apenas residualmente, o *se* impessoal. “No estilo formal, contudo, o falante brasileiro vai buscar a construção passiva, mais conservadora. Tais formas podem ser consideradas fora do domínio da gramática nuclear, com estatuto de morfologia estilística, marcadora de formalidade” (DUARTE, KATO & BARBOSA, 2001).

A explicação dada por elas, diante dessas constatações, está apoiada na Teoria da Reflexibilidade de Reinhardt and Reuland (1993 apud DUARTE, KATO & BARBOSA, 2001) e Dobrovie-Sorin (1998 apud DUARTE, KATO & BARBOSA, 2001). A hipótese central é que o *se* passivo com concordância em PB deve-se à perda do parâmetro do sujeito nulo já que somente línguas de SN evidenciam essa partícula e à perda em curso dos clíticos no PB, que possibilita construções como (111):

(111) Conserta sapato. (DUARTE, KATO & BARBOSA, 2001).

Dessa forma, *se* estaria em FL, mas é apagado em FF. Além disso, para elas, as construções com *se* sem concordância em PB são ambíguas entre um *se* nominativo e argumental (semelhante ao *on* do francês) e o *se* passivo do francês em que o sujeito é um expletivo de terceira pessoa. Apesar de o PB não permitir inversão livre, o sujeito nulo de 3ª pessoa ainda é retido, sendo esse expletivo, quasi-argumental e, portanto, um nulo referencial anafórico. Na tipologia de Reinhardt and Reuland (1993 apud DUARTE, KATO & BARBOSA, 2001), então, o sujeito nulo de terceira pessoa é uma anáfora. Consequentemente, no PB, o DP sujeito nulo é referencialmente dependente e isso pode ser visto na orientação para o tópico e para o sujeito.

#### **2.4. Concluindo o capítulo**

Portanto, diante do exposto, assume-se neste trabalho a hipótese da reanálise sintática de Nunes (1990) e a hipótese de Kato (2005a) para a aquisição de dados escritos. Com isso em mente, defende-se neste trabalho que o clítico *se* nas funções de indeterminador e de apassivador é um fósil linguístico recuperado pela instituição escolar e sua aquisição só é possível quando o processo de aquisição da gramática escolar, enquanto L2, se completa. Seu uso, portanto, se dá na forma de *code-switching* entre a língua adquirida naturalmente como L1 pela criança e a gramática escolar aprendida via instrução como L2, como propõe Kato, Cyrino e Correa (2009).

## Capítulo 3 – Análise dos Dados

A seguir, serão apresentados, descritos e discutidos, os dados coletados nos cinco *corpora* descritos anteriormente. A primeira seção descreverá esses *corpora* e as variáveis controladas, exemplificando as ocorrências e os tipos de variáveis relevantes para a análise. Em seguida, a segunda seção trará a estatística descritiva dos resultados encontrados seguida pela discussão e análise desses resultados.

### 3.1. Coleta e codificação dos dados

Como dito anteriormente, foram analisados cinco bancos de dados para compor esta pesquisa, sendo eles compostos por:

- Aproximadamente 200 gravações transcritas de duas crianças entre 2 e 5 anos pertencentes à coleção “Projeto Aquisição da Linguagem Oral” do Centro de Documentação Alexandre Eulalio (CEDAE - IEL/UNICAMP);
- Aproximadamente<sup>18</sup> 100 redações do banco de redações infantis de 1º a 4º ano do projeto intitulado "A relevância teórica dos dados singulares no processo de aquisição da linguagem escrita desenvolvido pelas professoras Prof<sup>ª</sup>. Dra. Maria Bernadete Marques Abaurre, Prof<sup>ª</sup>. Dra. Raquel Salek Fiad e Prof<sup>ª</sup>. Dra. Maria Laura Trindade Mayrink Sabinson no IEL/UNICAMP;
- 120 redações consideradas pela banca como acima da média do vestibular da Unicamp disponibilizadas pela Comissão Permanente para Vestibulares COMVEST);
- 60 gravações transcritas pertencentes ao Projeto NURC/SP (Norma Urbana Culta – São Paulo), dentre essas, 40 correspondem à conversa entre dois informantes e 20 a elocuições formais realizadas individualmente pelos informantes.
- 480 notícias do jornal Folha de S. Paulo retiradas do caderno de política nos seguintes intervalos: de 1980 a 1990 e de 2000 a 2010.

Ao todo, foram coletadas 1897 ocorrências de *se*, as quais estão divididas entre: *se* indeterminado, exemplificado em (112); *se* passivador, exemplificado em

---

<sup>18</sup> Não foi possível contabilizar o número exato de redações lidas pois o banco de dados é formado por cadernos escolares, os quais foram analisados em sua totalidade e não apenas produções isoladas.

(113); *se* pronominal, exemplificado em (114); *se* reflexivo, exemplificado em (115); *se* Sarney, exemplificado em (116); e *se* recíproco, exemplificado em (117).

(112) Espera-se que as futuras gerações saibam dar o devido valor a elas. (COMVEST)

(113) Só assim, através da conscientização, que se podem alcançar mudanças importantes. (COMVEST)

(114) A dengue que se prolifera em um determinado local. (Redações infantis)

(115) ...pa se enfeitá... (CEDAE)

(116) Está se usando esse critério. (NURC)

(117) Dois amigos que se amavam muito. (Redações Infantis)

Para realizar a análise dos dados, primeiramente, eles foram sistematizados de acordo com variáveis externas e internas. Foram consideradas as seguintes variáveis externas: banco de dados proveniente, tipo de gravação (para dados do NURC) e data de realização.

Já as variáveis internas foram selecionadas a partir das características consideradas determinantes para o uso de cada tipo *se* pelos autores que compõem o referencial teórico desta pesquisa. Foram controladas:

- O tipo de *se*: indeterminador, apassivador, Sarney, pronominal, reflexivo e recíproco;
- O tipo de verbo: transitivo direto, transitivo indireto, bitransitivo, inacusativo, inergativo e verbo de estado;
- O tempo verbal: presente, pretérito perfeito, pretérito imperfeito, futuro do pretérito, infinitivo, gerúndio, presente do subjuntivo, pretérito do subjuntivo, futuro do subjuntivo;
- A concordância entre o verbo e o argumento interno explícito;
- O traço de animacidade do argumento interno explícito;
- A posição do argumento interno explícito;
- A posição do *se* em relação ao verbo.

As estatísticas descritivas apresentadas a seguir foram obtidos a partir do programa Microsoft Excel 2013. Diante das distribuições encontradas, foi realizada uma análise de regressão logística multifatorial para verificar se a interação encontrada entre as variáveis é estatisticamente relevante. Para tanto, foi utilizado o programa *R MARKDOWN*.

A título de esclarecimento, uma vez que o objetivo do trabalho é observar como se dá a aquisição do clítico *se* apassivador e indeterminador, não foram coletados dados em que o clítico foi suprimido.

### **3.2. Distribuição dos dados**

#### **3.2.1 Variáveis Externas**

O que primeiro se pode observar em relação aos dados foi sua distribuição entre os bancos de dados. A tabela 1 representa essa distribuição. Como esperado, tanto o *se* indeterminador quanto o *se* apassivador tem presença inexpressiva em dados obtidos do CEDAE e das redações infantis. Por essa razão, nesses bancos de dados, foi controlada também a presença do *se* reflexivo, do *se* pronominal e do *se* recíproco para salientar que a aquisição desses outros tipos de *se* parece se comportar de forma mais natural e mais prematura, ao passo que a aquisição de *se* indeterminador e de *se* apassivador parece depender exclusivamente do processo de escolarização, afirmação permitida pelo significativo aumento dessas marcas em redações de vestibular, no NURC e na Folha de S. Paulo. Vê-se que, embora as partículas sejam homófonas, são adquiridas apenas aquelas cujas pistas estão presentes no *input* fornecido à crianças em fase de aquisição. Desse modo, crianças que possuem em seu *input* marcas sintáticas típicas da escrita poderão adquirir esses elementos com mais facilidade.

**Tabela 1 - Quantidade de *se* em cada banco de dados**

Tipo de SE		CEDAE	Redações Infantis	COMVEST	NURC	Folha de São Paulo	Total	%
Indeterminador	N	2	16	114	533	205	870	45.86%
	%	0.23%	2%	13%	61%	24%		
Apassivador	N	0	4	52	44	60	160	8.43%
	%	0.00%	2.50%	32.50%	27.50%	37.50%		
Pronominal	N	15	33	0	0	0	48	2.53%
	%	31.25%	68.75%	0.00%	0.00%	0.00%		
Reflexivo	N	52	68	0	0	0	120	6.33%
	%	43.33%	56.67%	0.00%	0.00%	0.00%		
Sarney	N	0	0	100	425	150	675	35.58%
	%	0.00%	0.00%	14.81%	62.96%	22.22%		
Recíproco	N	12	12	0	0	0	24	1.27%
	%	50.00%	50.00%	0.00%	0.00%	0.00%		
Total	N	81	133	266	1002	415	1897	100.00%
	%	4.27%	7.01%	14.02%	52.82%	21.88%		

Embora haja esse padrão na distribuição dos dados, os resultados da regressão logística mostraram que não há interação entre as duas variáveis, tipo de *se* e banco de dados (p.valor = 0.9977). Contudo, quando analisada a interação entre tipo de registro (oral ou escrito) e tipo de *se*, verificou-se que as diferenças são estatisticamente relevantes para sentenças com *se Sarney*, *se apassivador* e *se indeterminador* (p.valor < 2.2e-16), ou seja, a quantidade de ocorrências é sensível ao material em que o discurso é veiculado.

Notou-se também interação significativa entre faixa etária (criança, adolescente e adulto) e tipo de *se* (p.valor = 0.00010068). Isso se deve ao fato de que, uma vez que a aquisição do clítico *se* depende da escolarização, seu aparecimento depende da idade do falante.

Além disso, cabe destacar que as duas ocorrências de *se* indeterminador nos dados coletados do CEDAE, não corroboram a ideia de que essa aquisição é natural. Uma vez que todas as crianças que participaram do projeto provem de famílias com elevado grau de escolaridade, a língua que lhes serviu de input está muito mais próxima da norma, o que parece ter contribuído para a produção desses dados. Além disso, essas duas ocorrências, apresentadas em (118) e (119), foram observadas na fala da mesma criança. A primeira, aos 3 anos e 9 meses; a segunda, aos 4 anos e 2 meses:

(118) Cabou-se a história. (CEDAE)

(119) É aqui que se guarda. (CEDAE)

A sentença (118) contribui também para que se reconheça que a exposição a dados de registro escrito provenientes de livros infantis e de desenhos animados, por exemplo, é outra forma de contato com uma língua que está mais próxima da norma. Esse tipo de sentença é típico de contos infantis lidos para crianças e, por essa razão, talvez seja cristalizada para a criança. O vernáculo no PB pode, portanto, apresentar esse tipo de estrutura.

Essa afirmação também é válida para os dados colhidos das redações infantis, em que foram registradas 16 sentenças com *se* indeterminador. Também filha de pais muito escolarizados, a criança que produziu a maior parte dos dados provavelmente também teve em seu input dados de uma fala mais filtrada pela escrita, o que motiva a presença de dados como (120), (121) e (122) em seus cadernos do primeiro ciclo do Ensino Fundamental I:

(120) Aquilo que se fez anteriormente. (REDAÇÕES INFANTIS)

(121) Quando ele terminou de dizer se ouviu uma voz esquisita dizendo...  
(REDAÇÕES INFANTIS)

(122) Ai a gente recorta a ponta e se faz a mesma coisa com outro papel colorido. (REDAÇÕES INFANTIS)

A afirmação de Kato (2017) para a aquisição de clíticos de terceira pessoa cabe também nessas situações:

a variação na criança começa, em geral, com o início da escolarização com formas adquiridas em sua gramática nuclear e variantes fornecidas pela escola. Há casos, porém, de crianças expostas a ambientes altamente letrados, com pais que regularmente leem para os filhos, sem haver instrução de regras, situação em que a criança já pode exibir variação. (KATO, 2017, p. 8)

Também é interessante retomar Kato (2005a) neste ponto para exemplificar o caso de crianças expostas desde muito cedo a uma fala que apresenta o registro escrito, nos temos de Giusti (2004 apud PIRES, 2017), e que, por essa razão, incorporam marcas dessa gramática em seu conhecimento periférico.

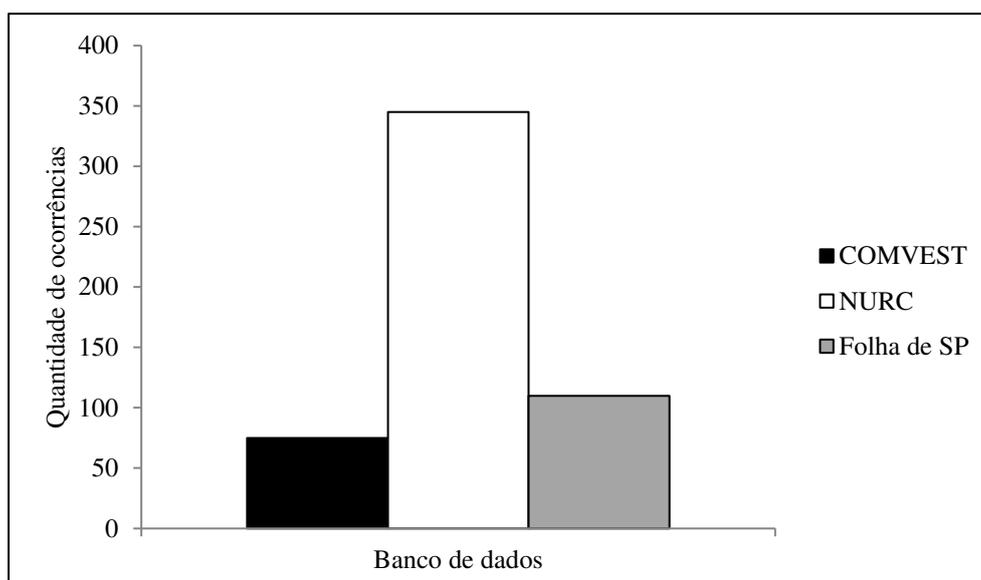
Por outro lado, dados como (123) e (124), embora não marcados como tal, parecem a tradicional atividade escolar de cópia:

(123) No dia 1º de maio comemora-se o Dia do Trabalho. (REDAÇÕES INFANTIS)

(124) Naquela fazenda, não se produz açúcar. (REDAÇÕES INFANTIS)

Com isso, é possível reafirmar que a aquisição de *se* apassivador e *se* indeterminador está, de fato, ligada à escolarização. Após 12 anos de Ensino Fundamental e Ensino Médio, o que é observado nas redações da COMVEST é um aumento para 13% das sentenças com *se* indeterminador.

Outro ponto a ser destacado na tabela 1 é o uso considerável de estruturas com o *se* Sarney. O primeiro dado com esse tipo de estrutura aparece em redações de vestibular para, em seguida, ter 63% de suas ocorrências registradas no NURC, ultrapassando as formas apassivadora e indeterminadora do clítico. Essa situação poderia ser explicada pelo grau de formalidade e de pressão do uso da modalidade de prestígio em cada uma das situações já que, como afirma Nunes (1990), é a forma inovadora mais bem aceita pelos falantes. Os dados extraídos da Folha de S. Paulo objetivam representar, assim como os dados da Revista Veja foram para o trabalho de Nunes (1990), uma amostragem da língua com poucas inovações ou apenas aquelas admitidas pelas variedades de prestígio. O que se observa, então, é a aceitação sem repulsa dessa forma condenada pela Gramática Tradicional. Cavalcante (2006) também encontrou uma correlação entre o tipo de material em que o dado foi produzido e o aparecimento de *se* com verbos infinitivos: apenas 10% das ocorrências foram vistas em dados de fala contra 50% em dados de escrita. No entanto, não é essa relação que o gráfico 1 abaixo apresenta. O que se vê é uma frequência muito superior de *se* com infinitivo em dados de fala do NURC em detrimento dos dados de escrita da COMVEST e da Folha de São Paulo:



**Gráfico 1 - Quantidade de *se* Sarney por banco de dados**

Logo, vê-se que sua aquisição depende da escolarização, mas a forma não depende necessariamente do registro escrito para ocorrer.

Da mesma forma, a considerável presença de *se* apassivador em um texto de concurso, como a COMVEST, lhe confere um grau de adequação à modalidade escrita bastante satisfatório. No entanto, ao contrário do que se esperava, a porcentagem de ocorrências de *se* apassivador em dados de fala chega próximo de 28%. Mesmo que as porcentagens mais altas estejam nos materiais escritos, vê-se que a forma ainda não está completamente extinta da fala<sup>19</sup>.

A discrepância entre as porcentagens de *se* indeterminador e *se* apassivador também chamam a atenção. Tendo em vista o que Nunes (1990) discutiu em seu trabalho a respeito da substituição de sentenças com *se* apassivador por sentenças com *se* indeterminador, vê-se que a porcentagem de uso de *se* indeterminador ultrapassa 77% no total dos dados. Até mesmo nos materiais escritos, há uma preferência pelo uso da norma indeterminadora em detrimento da apassivadora. Isso corrobora a ideia de que

No português brasileiro, essa mudança linguística praticamente já está em seus momentos finais: a construção com *se* indeterminador se tornou a forma canônica e a antiga construção com *se* apassivador ficou relegada à escrita. O dialeto brasileiro registra ainda duas mudanças mais recentes: o apagamento de *se* indeterminador em construções finitas e sua inserção em construções infinitas. (NUNES, 1990, p. iv)

<sup>19</sup> Posto que o NURC é composto por dados dos anos 1970 e 1980, e o processo estudado aqui ainda é uma mudança em curso, cabe verificar se essas porcentagens se verificam com dados mais recentes em uma pesquisa futura.

Confirma-se, então, o processo de reanálise sintática que tem feito com que o *se* indeterminador substitua o *se* apassivador no PB. Contudo, o desaparecimento do *se* apassivador e a expressiva queda do *se* indeterminador é uma mudança linguística que está em curso. Em vista disso, se faz necessário investigar a frequência desses dados ao longo dos anos em que foram colhidos. O gráfico 2 mostra essa frequência:

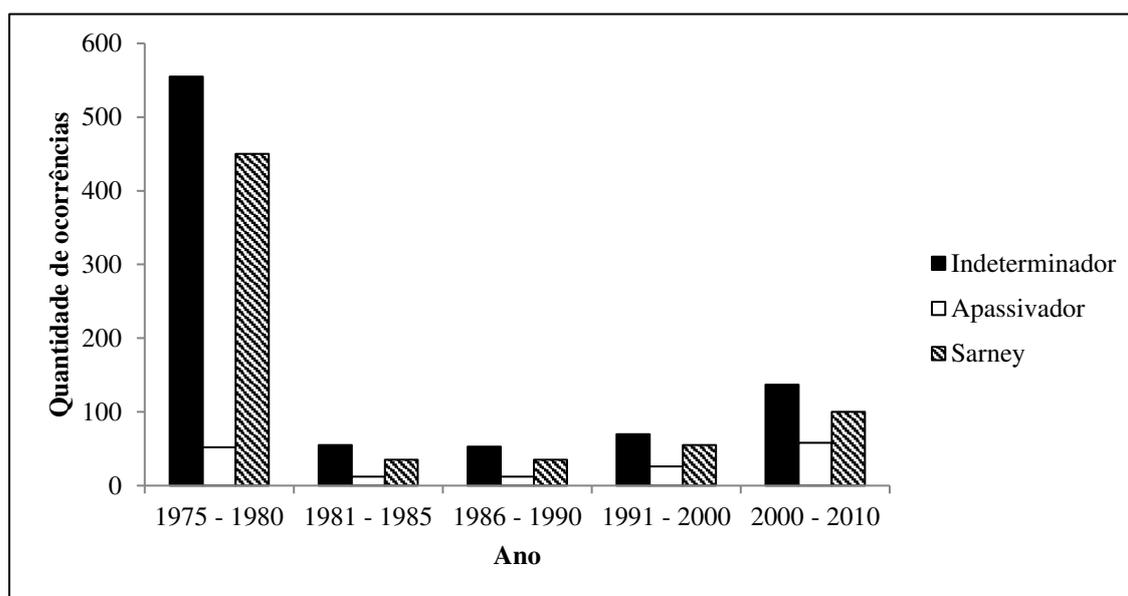


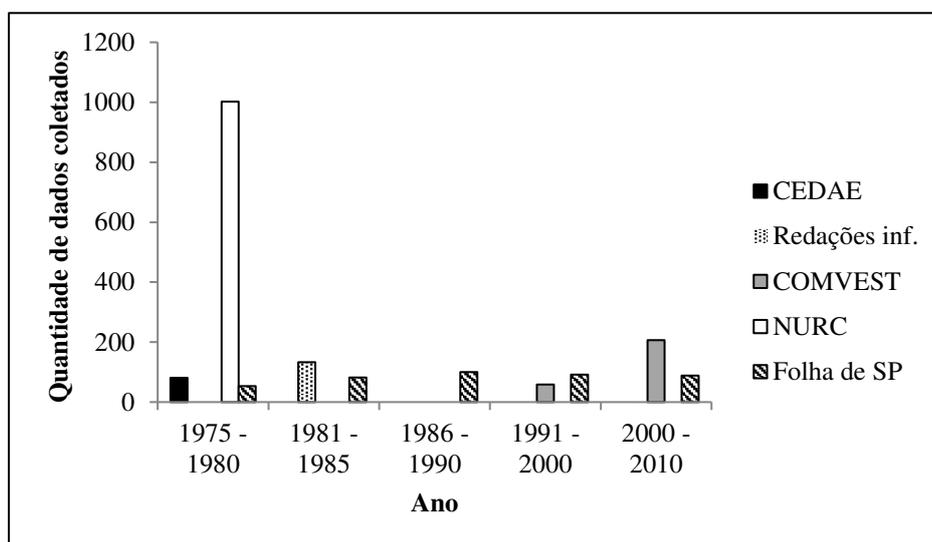
Gráfico 2 - Frequência de cada tipo de *se* encontrado por ano

O gráfico 2 reafirma a maior frequência de *se* indeterminador em relação ao dois outros que estão em questão. Dessa vez, não se observou o comportamento de *se* reflexivo, *se* recíproco e *se* pronominal pois o objetivo desta pesquisa é verificar o que fato de que embora sejam elementos homófonos, seus comportamentos são muito diferentes. Vimos que a criança não tem dificuldade para adquirir os outros três tipos de *se* ao passo que *se* indeterminador. *Se* apassivador e *se* Sarney dependem da escolarização e, por isso, apenas esses últimos serão destacados a partir deste momento.

Os testes de regressão logística, entretanto, não mostraram interação entre as variáveis Tipo de *se* por ano (p.valor = 0.9963).

De volta ao gráfico 1, *se* Sarney é o segundo mais utilizado, restando apenas menos de 50 ocorrências de *se* apassivador em todos os períodos analisados. Por outro lado, ele revela que a maior quantidade de *se* indeterminador e de *se* Sarney foi realizada no período compreendido entre os anos 1975-1980. Entre os anos 1981 e 1990, as frequências encontradas foram muito parecidas, não ultrapassando 100

ocorrências de cada tipo de *se*. A partir de 1991, a frequência de uso dessas marcas voltou a ter crescimento, principalmente após os anos 2000. Possivelmente, esse crescimento se deve ao fato de que a maior parte dos dados dos anos 2000 são escritos e são redações de vestibular, momento em que a pressão do uso de formas tradicionais é mais intensa. Esse fato também pode explicar o aumento de *se* apassivador nesses anos, de forma a alcançar a maior quantidade de ocorrências. Uma vez que a situação exige formalidade, a gramática escolar do falante é acionada e o uso de sentenças passivas surge como marca dessa gramática. Por essa razão, o gráfico 2, para ser analisado, precisa ser comparado aos dados do gráfico 3 seguinte. Nele, a questão da irregularidade da quantidade de dados por ano é apresentada:



**Gráfico 3 – Quantidade de dados coletados em cada banco de dados por ano**

Observa-se que a quantidade de dados não é regular entre os bancos de dados. A maior quantidade de sentenças coletadas está no NURC, o qual apresenta dados apenas do período compreendido entre 1975 e 1980, assim como o CEDAE. Da mesma forma, as redações infantis são todas do período compreendido entre 1981-1985. Os dados da COMVEST, por sua vez, representam as décadas de 1990 e 2000, ao passo que apenas os dados da Folha de SP puderam ser coletados em todos os períodos analisados.

Essa questão metodológica indica que, embora o gráfico 2 mostre queda no uso de uso dos três tipos de *se*, essa queda também pode ser explicada pela menor quantidade de dados, principalmente de fala, a partir dos anos 1990.

### 3.2.2. Variáveis Internas

Para dar início à exposição das variáveis internas, serão apresentadas primeiramente as que estão relacionadas ao argumento interno por ser um elemento muito discutido na literatura sobre o *se*. Em 126 sentenças, o argumento interno não estava explícito, para a verificação das variáveis estudadas a seguir, essas sentenças foram descartadas.

Um dado curioso encontrado em relação aos argumentos internos é a preferência por argumentos inanimados. A tabela 2 apresenta essa distribuição tomando como conjunto as sentenças em que o argumento interno foi explicitado pelo falante:

**Tabela 2 – Relação entre animacidade do argumento interno e tipo de *se***

Tipo do SE		- Animado	+ Animado	Total	%
Indeterminador	N	777	12	789	81.85%
	%	98.48%	1.52%		
Apassivador	N	68	3	71	7.37%
	%	95.77%	4.23%		
Sarney	N	103	1	104	10.79%
	%	99.04%	0.96%		
Total	N	948	16	964	100.00%
	%	98.34%	1.66%		

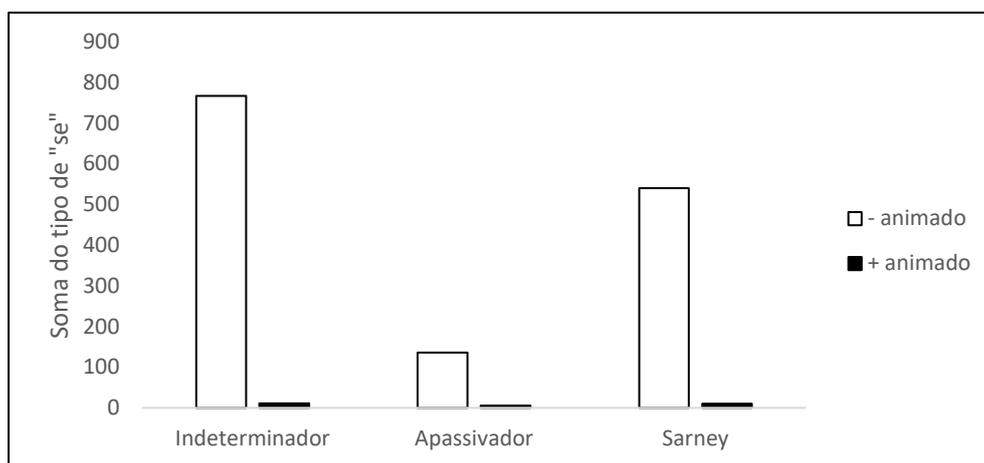
A interação entre essas variáveis também foi comprovada pela regressão logística (p.valor < 0.04607)

Embora termos não sejam sempre animados, apenas 1,67% do total de sentenças apresentaram o argumento interno [+animado]. Então, sentenças como (125) são muito mais frequentes que (126):

(125) Se cortava o bambu, tirava fatias de bambu. (NURC)

(126) Se contrata mulheres. (NURC)

Essa preferência também pode ser vista no gráfico 4 abaixo:



**Gráfico 4 – Animacidade do argumento interno por tipo de se**

Ainda em relação ao argumento interno, verificou-se recorrente preferência em expressá-lo posposto ao verbo, como mostra a Tabela 3:

**Tabela 3 - Posição do argumento interno em relação ao verbo**

Tipo do SE		Anteposto ao verbo	Posposto ao verbo	Total	%
Indeterminador	N	245	538	783	81.13%
	%	31.49%	68.51%		
Apassivador	N	30	41	71	7.40%
	%	42.25%	57.75%		
Sarney	N	42	68	110	11.47%
	%	38.18%	61.82%		
Total	N	317	647	964	100.00%
	%	33.06%	66.94%		

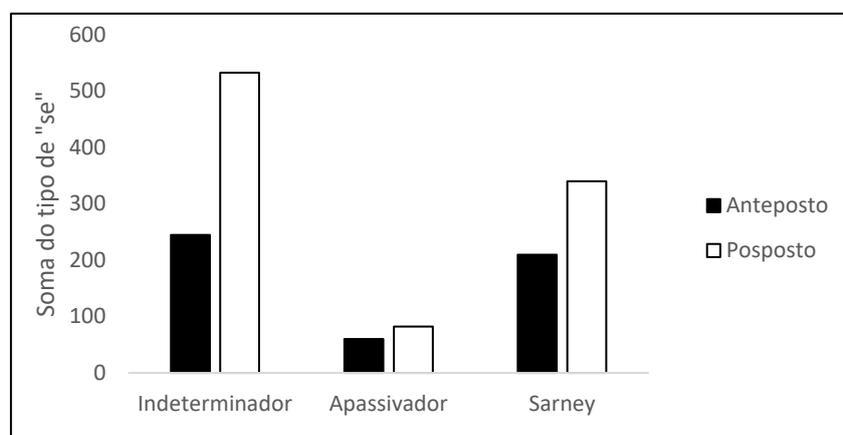
Para os três tipos de *se*, a porcentagem de sentenças como (127), em que o argumento está posposto ao verbo, ultrapassa 50% enquanto sentenças em que o argumento interno está posposto ao verbo, como em (128), são aproximadamente 40% do total. Esses resultados confirmam as afirmações de Melo (2012) sobre a preferência pela posposição. Esse critério não foi aplicado a sentenças em que o argumento interno não era explícito ou a sentenças com verbos inergativos como (129):

(127) Já nem se fala mais na simples falsificação de cheques preenchidos.  
(FOLHA DE S. PAULO)

(128) Piada de Presidente da República não se faz. (FOLHA DE S. PAULO)

(129) Do jeito que se vive, a gente vive correndo. (NURC)

Esses resultados também podem ser verificados o gráfico 5 a seguir:



**Gráfico 5 – Posição do argumento interno em relação ao verbo por tipo de se**

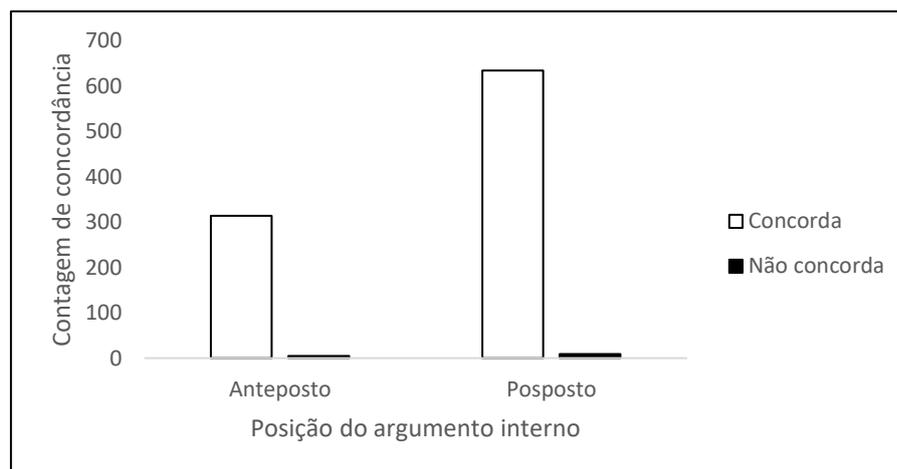
A interação entre as variáveis também foi comprovada pela regressão logística (p.valor = 0.025055).

Esses dados se aproximam muito do que foi encontrado por Nunes (1990). Em seu conjunto de dados referentes ao século XX, o autor verificou um índice de 84% de posposição do argumento interno em relação ao verbo. Tal porcentagem lhe foi útil para entender os contextos em que há falta de concordância entre verbo e argumento interno já que em 92% dos casos de discordância vistos por ele na revista *Veja*, as sentenças exibiam o argumento interno posposto ao verbo. Para investigar se a posição do argumento interno interfere na concordância, a tabela 4 cruza esses dois fatores para sentenças em que o argumento interno foi explicitado pelo falante:

**Tabela 4 – Concordância em relação à posição do argumento interno**

Argumento Interno		Há concordância	Não há concordância	Total	%
Anteposto	N	315	5	320	33.16%
	%	98.43%	1.57%		
Posposto	N	635	9	644	66.84%
	%	98.60%	1.40%		
Total	N	950	14	964	100.00%
	%	98.54%	1.46%		

Há casos isolados de falta de concordância e, por isso, não é possível dizer seguramente se a posição do argumento interno é um fator determinante, embora a quantidade de ocorrências sem concordância em sentenças em que o argumento é posposto ao verbo seja o dobro das sentenças em que ele é anteposto. Isso também pode ser observado no gráfico 6 abaixo:



**Gráfico 6 – Concordância do argumento interno em relação à posição do argumento interno**

Os resultados da regressão logística, por outro lado, mostraram interação entre as variáveis ( $p.\text{valor} < 2.22e-16$ ). Uma hipótese a ser considerada para explicar essa interação é o fato de que a considerável preferência pela concordância entre verbo e argumento interno pode indicar que os falantes interpretam o argumento interno como sujeito.

É interessante observar, então, em quais bancos de dados essa discordância aparece com mais frequência, como revela a tabela 5:

**Tabela 5 – Concordância entre verbo e argumento interno por banco de dados**

Concordância		CEDAE	Redações Infantis	COMVEST	NURC	Folha de São Paulo	Total	%
[+Concordância]	N	0	14	156	525	255	950	98.54%
	%	0.00%	1.48%	16.46%	55.17%	26.90%		
[-Concordância]	N	1	2	4	7	0	14	1.46%
	%	7.14%	14.29%	28.57%	50.00%	0.00%		
Total	N	1	16	160	530	255	964	100.00%
	%	0.10%	1.66%	16.63%	55.09%	26.51%		

Apesar da quantidade de dados em que há falta de concordância ser muito pequena, o aumento da porcentagem de sentenças sem concordância nos dados do NURC e sua completa ausência nos dados da Folha de S. Paulo podem indicar que sentenças como (130), (131) e (132) tendem a ser menos aceitas pelos falantes. Se essas formas foram excluídas do material escrito que exhibe as formas de prestígio e, até mesmo na fala escolarizada aparece pontualmente, pode-se constatar que esse uso não tem boa avaliação pelos falantes. Além disso, os resultados da regressão logística indicaram interação entre as variáveis ( $p$ .valor = 0.019235).

(130) Passou-se uns minutos e foi quando eu falei pra minha prima freiar.  
(REDAÇÕES INFANTIS).

(131) E os aniversários de casamento se comemora, não? (NURC)

(132) só se faziam espetáculo dentro daquela linha. (NURC)

Melo (2012) também observou o comportamento dos argumentos internos em sentenças com *se* indeterminador e *se* apassivador como foi visto no capítulo 2. A autora encontrou em seus dados uma correlação entre a posição do argumento interno e sua animacidade. A fim de verificar se sua hipótese, de que a anteposição só é possível quando o argumento for [-animado], a tabela abaixo cruza essas variáveis:

**Tabela 6 – Posição do argumento interno por animacidade do argumento interno**

Animacidade		Anteposto ao verbo	Posposto ao verbo	Total	%
[+Animado]	N	3	13	16	1.66%
	%	18.75%	81.25%		
[-Animado]	N	316	632	948	98.34%
	%	33.40%	66.60%		
Total	N	319	645	964	100.00%
	%	33.16%	66.84%		

Apesar da pequena quantidade de dados com argumentos com traço [+animado], a tabela 5 indica que, assim como Melo (2012) observou em seus dados, há uma expressa preferência pela anteposição de argumentos [-animados] em detrimento dos argumentos [+animados]. De acordo com a autora:

Esse DP, em geral, não pode ter traço [+humano] visto que se isto ocorrer a interpretação da sentença é reflexiva, pois o DP não subiria para [SPEC-FP] em virtude de precisar primeiro preencher a grade temática de [SPEC-TP] para depois haver a subida de um nível e o DP atingir [SPEC-FP]. Em sentenças com DP [+inanimado], a necessidade de preencher [SPEC-TP] para depois alçar um DP para [SPEC-FP] faz com que no PB sentenças como as expostas em (40) sejam interpretadas como agramaticais. (MELO, 2012, p. 134)

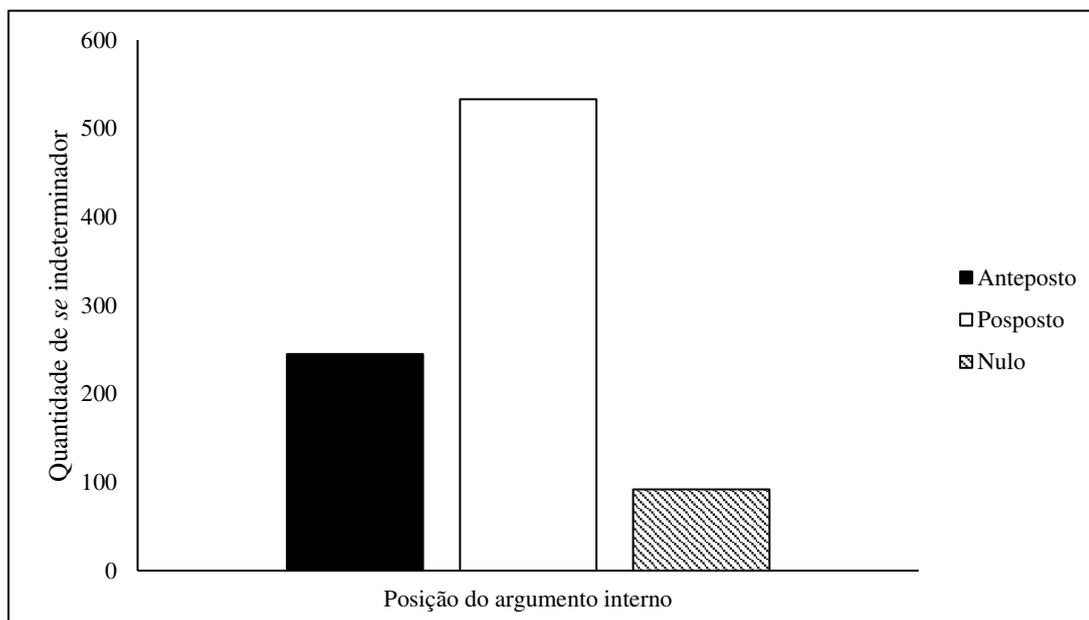
Outra hipótese da autora diz respeito à relação entre tempo verbo e anteposição do argumento interno. Sua hipótese afirma que, quando houver anteposição, a interpretação genérica depende de predicados genéricos, ou seja, de predicados construídos com verbos no presente ou no pretérito imperfeito. A tabela abaixo mostra a relação entre as duas variáveis:

**Tabela 7 – Posição do argumento interno por tempo verbal**

Posição do argumento		Pres.	Pret.Perf.	Pret. Imperf.	Fut. do Pret.	Infi.	Ger.	Pret. do Subj.	Pres. do Subj.	Fut. do Subj.	Total	%
Anteposto	N	201	25	29	4	38	4	3	15	0	319	33.16%
	%	63.01%	7.84%	9.09%	1.25%	11.91%	1.25%	0.94%	4.70%	0.00%		
Apassivador	N	408	43	100	9	48	20	6	10	1	645	66.84%
	%	63.14%	6.69%	15.55%	1.40%	7.47%	3.11%	0.93%	1.56%	0.16%		
Total	N	609	68	129	13	86	24	9	25	1	964	100.00%
	%	63.10%	7.07%	13.41%	1.35%	8.94%	2.49%	0.94%	2.60%	0.10%		

O que se observa é de fato uma preferência por ambos os tempos verbais. No entanto, há pouca diferença entre pretérito perfeito e pretérito imperfeito.

Por fim, a última consideração da autora aqui atestada é o aumento no uso de argumentos nulo em sentenças com *se* indeterminador. O que se observa é que ainda há mais argumentos internos pospostos que nulos no gráfico 7 abaixo:

**Gráfico 7 – Posição de argumento interno em sentenças com *se* indeterminador**

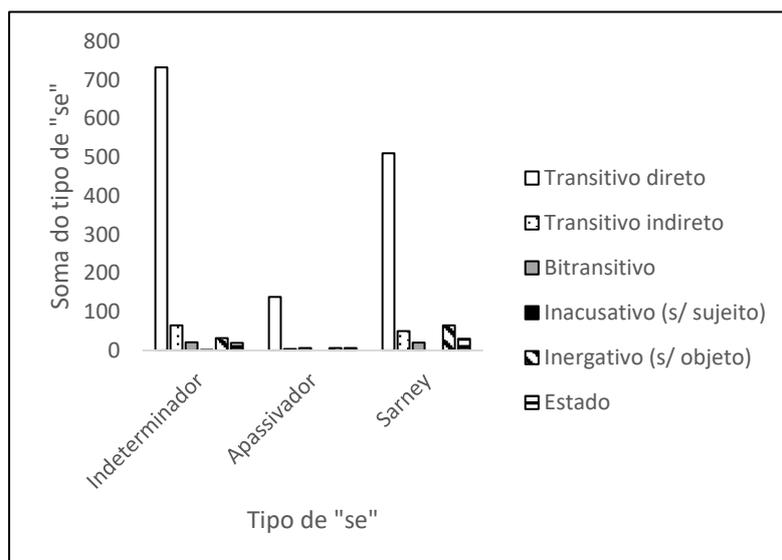
Findas as informações relativas aos argumentos internos das sentenças com *se* e suas características, serão vistas agora as particularidades dos verbos observadas nos dados. Ao examinar a tabela 8 abaixo, que descreve o comportamento dos verbos em relação à transitividade em cada para as sentenças com cada tipo de *se*, vê-se que as

categorias de verbos que permite o uso de *se* como recurso para a indeterminação do papel temático de sujeito, expandiram. A existência de sentenças com *se* indeterminador acompanhando verbos transitivos diretos, verbos transitivos indiretos, bitransitivos, inergativos e verbos de estado vai de encontro à prescrições da Gramática Tradicional. Como foi visto no segundo capítulo desta dissertação, o uso de *se* indeterminador deve ser restrito apenas a verbo intransitivos ou não acompanhados de objeto direto, enquanto o uso de *se* apassivador deve se limitar apenas aos transitivos diretos.

**Tabela 8 - Transitividade do verbo em relação ao tipo de *se***

Tipo de SE		Transitivo Direto	Transitivo Indireto	Bitransitivo	Inacusativo (sem sujeito)	Inergativo (s/ objeto)	Estado	Total	%
Indeterminador	N	732	65	21	0	32	19	869	51.06%
	%	84.23%	7.48%	2.42%	0.00%	3.68%	2.19%		
Apassivador	N	138	2	6	0	6	6	158	9.28%
	%	87.34%	1.27%	3.80%	0.00%	3.80%	3.80%		
Sarney	N	510	50	20	0	65	30	675	39.66%
	%	75.56%	7.41%	2.96%	0.00%	9.63%	4.44%		
Total	N	1380	117	47	0	103	55	1702	100.00%
	%	81.08%	6.87%	2.76%	0.00%	6.05%	3.23%		

A tabela 8, então, mostra que a maior porcentagem de ocorrências, para todos os tipos de *se*, aparece em sentenças com verbos transitivos diretos. Vê-se essa distribuição também no gráfico 8:



**Gráfico 8 – Transitividade do verbo por tipo de *se***

A interação entre as variáveis também foi atestada pelos resultados da regressão logística (p.valor = 0.001617). O uso de *se* apassivador, por outro lado, parece seguir as prescrições da norma uma vez que há apenas duas ocorrências de *se* apassivador com verbos transitivos indiretos:

(133) Então servem-se da água do rio. (NURC)

(134) Porém utilizam-se de humor. (NURC)

Logo, embora, as mudanças linguísticas envolvendo *se* indeterminador e *se* apassivador tenham ampliado seus contextos de usos, há ainda um padrão na escolha por verbos transitivos diretos. Nota-se também que não há nenhum verbo inacusativo entre os dados já que é necessário que o verbo tenha sujeito para que esse papel temático seja indeterminado.

Outro fator que apresenta regularidade é o tempo verbal dos verbos que acompanham *se*. Optou-se aqui por analisar essa variável, primeiramente, para determinar se o uso de *se* Sarney prevalece na forma gerundiva ou na forma infinitiva. Em seguida, estendeu-se a observação para todos os tipos de *se* a fim de verificar padrões de ocorrência, entretanto, os resultados da regressão logística não indicaram interação entre as variáveis (p.valor = 0.9997) . Para *se* indeterminador e *se* apassivador, a forma no presente é, de longe, a preferida. O *se* Sarney, por sua vez, aparece mais em sentenças com verbos no infinitivo. Isso pode ser observado na Tabela 9:

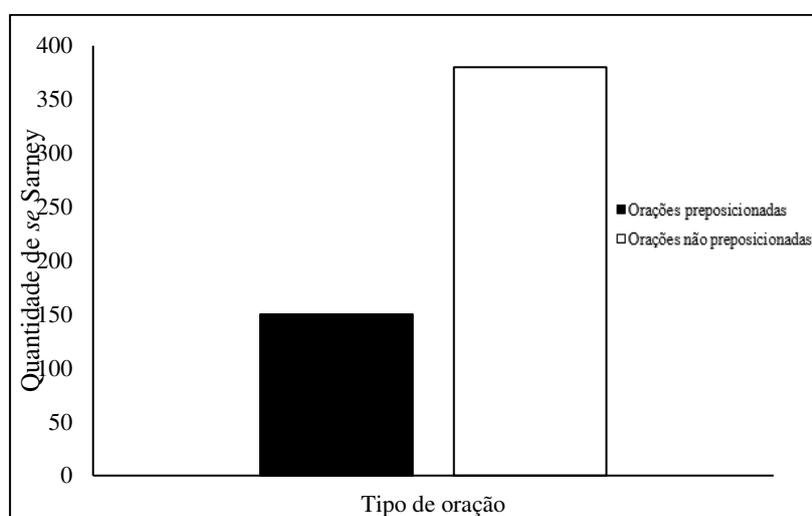
**Tabela 9 – Tempo verbal por tipo de *se***

Tipo de SE		Pres.	Pret. Perf.	Pret. Imperf.	Fut.do Pret.	Inf.	Ger.	Pret. do Subj.	Pres. do Subj.	Fut. do Subj.	Total	%
Indeterminador	N	616	67	137	14			9	26	1	870	51.03%
	%	70.80%	7.70%	15.75%	1.61%	0.00%	0.00%	1.03%	2.99%	0.11%		
Apassivador	N	128	12	16					4		160	9.38%
	%	80.00%	7.50%	10.00%	0.00%	0.00%	0.00%	0.00%	2.50%	0.00%		
Sarney	N					530	145				675	39.59%
	%	0.00%	0.00%	0.00%	0.00%	78.52%	21.48%	0.00%	0.00%	0.00%		
Total	N	744	79	153	14	530	145	9	30	1	1705	100.00%
	%	43.64%	4.63%	8.97%	0.82%	31.09%	8.50%	0.53%	1.76%	0.06%		

Nunes (1990) já definia o *se* Sarney como aquele que acompanhava sentenças no infinitivo, sendo o “gerúndio” apenas uma extensão dos casos. De fato, foi o que se observou: apenas em 22% das sentenças com esse tipo de *se* o verbo estava no

gerúndio. Esse dado era esperado uma vez que a forma infinitiva apareceu primeiramente e é sobre a qual Nunes (1990) se debruça. A inserção de *se* com verbos no gerúndio seria, então, apenas um processo decorrente da inserção de *se* com verbos no infinitivo.

Ainda sobre o *se* Sarney, Cavalcante (2006) afirma que há três contextos que favorecem o aparecimento do *se* junto a verbos no infinitivo, como já citado no capítulo 2: orações preposicionadas, presença de agente da passiva e posição do argumento interno. Em relação ao primeiro contexto, o que se observou nos dados analisados confirmou o que afirmou a autora, como mostra o gráfico 9:



**Gráfico 9 – Quantidade de *se* Sarney em oração preposicionadas e não preposicionadas**

Nota-se que o número de ocorrências de *se* Sarney em orações preposicionadas, como (135) e (136) é maior que o dobro do número de sentenças como (137) e (138):

(135) A Escola pede para *se* tomar uma providência. (NURC)

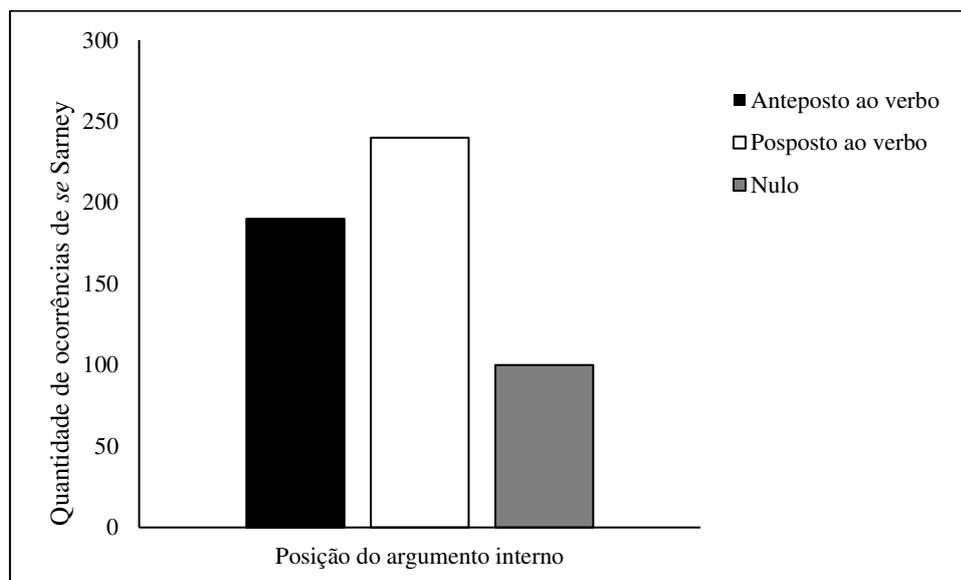
(136) Nunca houve filme obrigatório para *se* ver em cinema nacional.  
(NURC)

(137) Não poderia *se* cuidar de uma área desprezando a outra. (NURC)

(138) Tudo o que *se* fizer a fim de impedir a derrota passa a ser moralmente válido. (FOLHA DE S. PAULO)

Já em relação ao agente da passiva, não foi encontrada nenhuma sentença em que esse elemento estava expresso, resultado muito semelhante ao encontrado por Cavalcante (2006). A autora registrou apenas duas ocorrências e busca como justificativa a afirmação de Naro (1976) retomada por Nunes (1990): o agente da passiva aparece junto ao *se* em Português somente até o século XVI.

Quanto à posição do argumento, não se observou tanta diferença entre argumentos pospostos ao verbo e antepostos ao verbo, como mostra o gráfico 10 abaixo:



**Gráfico 10 – Quantidade de *se* Sarney por posição do argumento interno**

Cavalcante (2006) observou que sentenças com argumentos internos pospostos ao verbo ou nulos (não-expressos) se comportaram como contextos favorecedores do aparecimento de *se* junto de verbos infinitivos. Nos dados aqui analisados, entretanto, sentenças com argumento interno anteposto ao verbo somam quase o dobro da quantidade de sentenças com argumento interno nulo. Embora a posposição do argumento interno seja pareça ser uma preferência, a diferença numérica é pequena para que se faça uma constatação categórica a respeito desse contexto como um favorecedor.

Ainda retomando o trabalho de Cavalcante (2006), a autora apontou que *se* só acompanhava verbos infinitivos se esses fossem transitivos. Essa preferência também foi vista nos dados desta pesquisa, mas houve uma aparente expansão na categoria de verbos que permitem essa construção com *vê-se* no gráfico 11 abaixo:

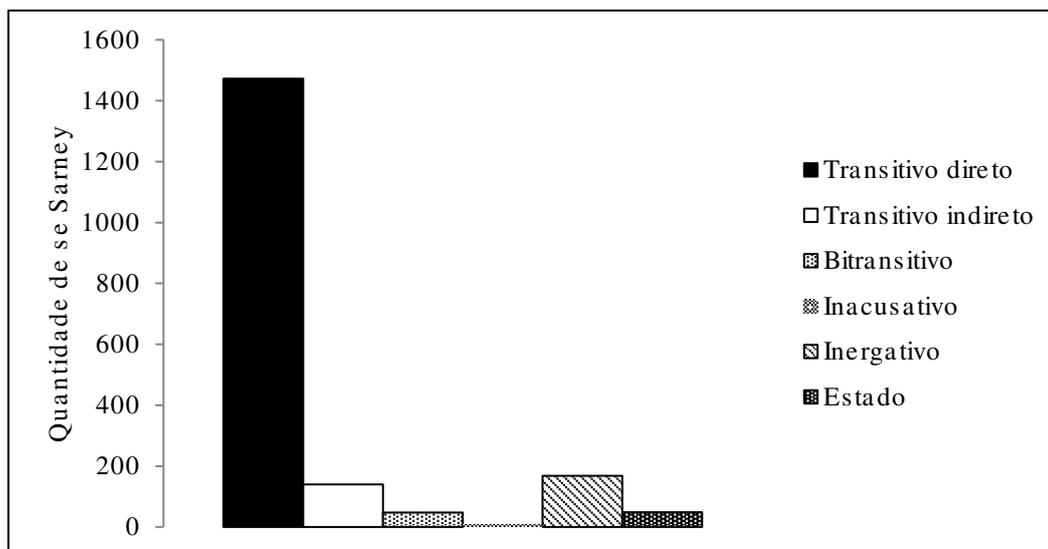


Gráfico 11 – Quantidade de *se Sarney* por transitividade verbal

Mais de 90% das sentenças registradas apresentavam verbos transitivos mas também forma registradas ocorrências com verbos inergativos, como (139), e com verbos de estado, como (140):

(139) Para se sair ou entrar em Belo Horizonte tem que pegar a rodovia.  
(NURC)

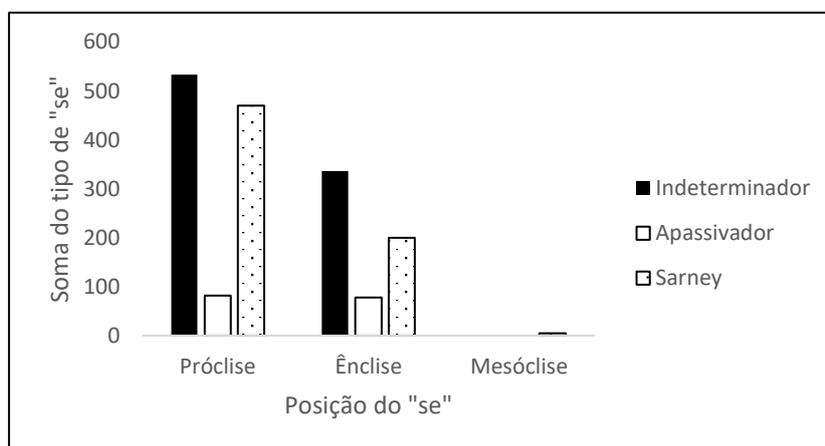
(140) Talvez pelo fato de se ser mulher. (NURC)

Uma vez que a posição dos clíticos também é um elemento que diverge entre o PB falado e o PB escrito, verificou-se a frequência de próclises, ênclises e mesóclises a fim de investigar se há relação entre os elementos preconizados pela normatividade escolar. A tabela 10 apresenta a distribuição da posição do clítico em relação ao verbo para cada tipo de *se*:

Tabela 10 - Posição de *se* por tipo de *se*

Tipo do SE		Pré-verbal	Pós-verbal	Mesóclise	Total	%
Indeterminador	N	533	336	1	870	51.03%
	%	61.26%	38.62%	0.11%		
Apassivador	N	82	78	0	160	9.38%
	%	51.25%	48.75%	0.00%		
Sarney	N	470	200	5	675	39.59%
	%	69.63%	29.63%	0.74%		
Total	N	1085	614	6	1705	100.00%
	%	63.64%	36.01%	0.35%		

Vê-se que a preferência de próclise é generalizada para todos os tipos de *se*. Logo, não há interação entre as variáveis como revela os resultados da regressão logística (p.valor = 0.9951). Essa preferência é esperada uma vez que o PB prefere a posição proclítica para clíticos (cf. Pagotto, 1993). Entretanto, o uso de ênclise e de mesóclise não tem relação com o *se* apassivador. Esses resultados também podem ser observados no gráfico 12 abaixo:



**Gráfico 12 - Posição de *se* por tipo de *se***

Ademais, o intrigante uso de mesóclise, estrutura rara que, mais que o *se* apassivador, se tornou uma exclusividade de materiais escritos instiga o olhar sobre a frequência de sentenças com *se* em cada uma das três posições por banco de dados. Segue abaixo a tabela 11, que representa essas ocorrências:

**Tabela 11 - Posição de *se* por banco de dados**

Tipo do SE		Pré-verbal	Pós-verbal	Mesóclise	Total	%
CEDAE	N	20	4	0	24	1.54%
	%	83.33%	16.67%	0.00%		
Redações Infantis	N	36	24	0	60	3.85%
	%	60.00%	40.00%	0.00%		
COMVEST	N	58	202	1	261	16.76%
	%	22.22%	77.39%	0.38%		
NURC	N	525	230	0	755	16.76%
	%	201.15%	88.12%	0.00%		
FOLHA DE SP	N	72	384	1	457	16.76%
	%	27.59%	147.13%	0.38%		
Total	N	711	844	2	1557	100.00%
	%	45.66%	54.21%	0.13%		

De fato, quanto às mesóclises, a hipótese acima se confirma: as duas ocorrências, abaixo transcritas, foram observadas em ambientes escritos:

(141) Então, no caso, far-se-ia uma farsa. (FOLHA DE S. PAULO)

(142) Poder-se-ia facilmente ficar impressionado. (COMVEST)

Os dados também indicam que a forma enclítica está ligada à escolarização. Nos dados de aquisição, ambos os usos de *se* após o verbo estão em sentenças cristalizadas ora por uma cantiga infantil, ora pela literatura infantil:

(143) Dona chica ca ca admirou se se. (CEDAE)

(144) Cabou-se a história. (CEDAE)

É somente nos primeiros anos escolares que esse tipo de estrutura surge, ainda sem que a criança tenha aprendido a segmentar as palavras de acordo com o que a norma da escrita prevê:

(145) Lavou as mãos e foi se embora. (REDAÇÕES INFANTIS)

(146) Pitacuça sentouse ao lado de um velho. (REDAÇÕES INFANTIS)

(147) ...da seguinte maneira: dobrase uma ponta assim. (REDAÇÕES INFANTIS)

Entretanto, ao fim da escolarização novamente, as porcentagens referentes ao uso de ênclise aumentam e, nos materiais escritos, sua ocorrência praticamente dobra em relação ao uso de próclise.

Para finalizar essa análise, convém olhar ainda para o registro fornecido pelo Microsoft Excel das formas que, em frequência, prevalecem sobre as demais. Por meio do cálculo da mediana dos códigos dos dados, foi possível observar um padrão nas sentenças em que cada tipo de *se* aparece. Curiosamente, esse padrão é bastante semelhante para todas as estruturas, salvo apenas as características que lhe são particulares, como o *se* Sarney que ocorre com verbos infinitivos e gerundivos. As estruturas com *se* indeterminador geralmente apresentam argumento interno [- animado]

posposto ao verbo e com concordância, verbos transitivos diretos no presente e *se* pré-verbal, tal qual a sentença (148):

(148) Não se comete exagero algum. (FOLHA DE SP)

Igualmente, o *se* passivador também aparece nos mesmos contextos: verbos transitivos diretos no presente, *se* pós-verbal, argumento interno [- animado] posposto ao verbo e com concordância, como vê-se em (149):

(149) Fecham-se as cortinas, acende as luzes. (NURC)

Esse último dado, além de ilustrar o contexto, exatamente idêntico ao anterior, em que *se* passivador aparece, revela ainda a competição de gramáticas entre a forma indeterminadora de *se* e seu apagamento. Logo após a realização de uma sentença com *se* passivador, o falante suprime o clítico e retira a concordância entre verbo e argumento interno.

Outro dado interessante que referente à elisão de *se* está transcrito abaixo:

(150) Não se pode admitir que seja por falta de instrumentos legais que não são apurados os atentados terroristas, nem se punem os responsáveis. (FOLHA DE SP)

Nota-se que, ao contrário do que Nunes (1990) verificou em seus dados diacrônicos, não há elisão de clítico em sentença coordenada com o uso da conjunção *nem*.

Retornando aos contextos sintáticos de aparecidos do *se*, novamente, *se* Sarney também ocorre nos mesmos contextos, como (151) e (152):

(151) Hoje se existem inseticidas, que não causam dano aos animais.  
(NURC)

(152) Está se usando esse critério. (NURC)

Por fim, apesar da regularidade de contextos, o *se* apareceu nas mais diversas estruturas. Na sessão abaixo, os dados vistos acima serão discutidos e suas implicações destacadas.

#### 4. Discussão

Diante do exposto, os dados encontrados parecem confirmar o que afirmou Nunes (1990):

O surgimento da passiva pronominal sem concordância, apontado por Naro<sup>20</sup>, desencadeou um processo de mudança linguística no português brasileiro em que a forma conservadora (com concordância) foi paulatinamente perdendo terreno para sua concorrente, culminando no momento sincrônico, em que deve sua sobrevivência na modalidade escrita à renitência da gramática normativa. (NUNES, 1990, p. 77)

Com efeito, viu-se que a forma indeterminadora de *se* tomou espaço na língua e parece ser preferida em relação à forma apassivadora independente do material em que a sentença é produzida, como mostra o gráfico 13:

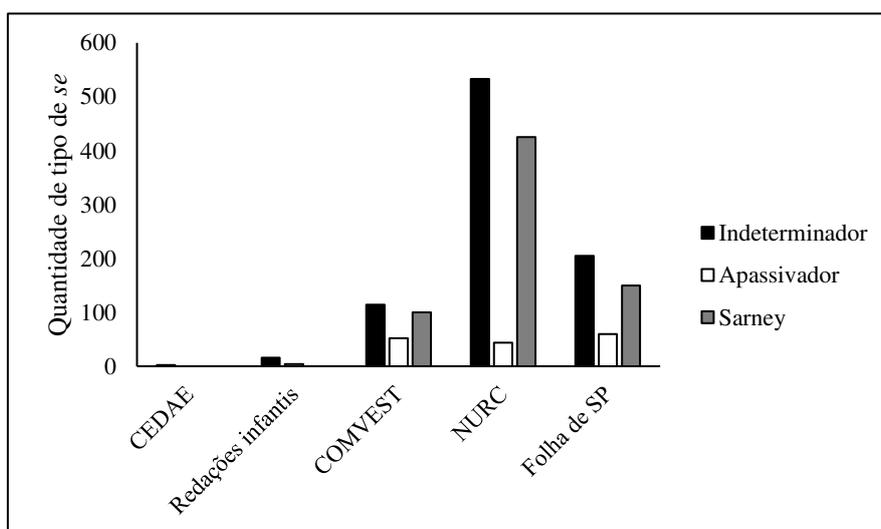


Gráfico 13 – Quantidade de ocorrência de cada tipo de *se* por Banco de Dados

Por outro lado, é possível fazer algumas ressalvas. Embora os testes de regressão logística não tenham indicado interação entre tipo de *se* e banco de dados (p.valor = 0.9977), *se* apassivador, apareceu em frequências muito semelhantes quando

<sup>20</sup> Nunes (1990) se refere ao trabalho de Naro (1976) em que o autor busca as origens de *se* indeterminador.

comparados dados de fala e de escrita. Isso parece indicar que seu uso ainda extrapola os limites da escrita. Também não era esperado que sua frequência fosse tão alta em redações escolares do século XXI. São 33% das sentenças com a forma apassivadora nessas redações contra 27% no NURC, como mostrou a tabela 1. Esse dado pode indicar que a variedade vernacular dos anos 1970 era mais próxima da norma que a atual, embora ainda não haja um estudo mais recente sobre o comportamento dessas sentenças para que se possa realizar um estudo comparativo que verifique se há, de fato, uma mudança em curso.

Ademais, verificou-se que uma das mudanças que Nunes (1990) apontou ocorreu massivamente: a inserção de *se* em sentenças com infinitivo e gerúndio ou o aumento do uso do *Se Sarney*, principalmente na fala. Essa forma também foi inserida em contextos escritos e aparece com uma frequência próxima da forma indeterminada do clítico tanto nos dados da COMVEST quanto nos dados da Folha de S. Paulo. Os resultados relativos ao *Se Sarney* também se comportaram de forma muito semelhante à observada por Cavalcante (2006). Dos três contextos favorecedores observados pela autora, dois foram confirmados: orações preposicionadas e ausência de agente da passiva. A posição do argumento interno não teve diferença significativa para ser considerada um contexto favorecedor tal qual propõe Cavalcante (2006). Tampouco observou-se uso categórico de verbos transitivos diretos. Assim como foi visto para os demais tipos de *se*, houve uma expansão das categorias que permitem o uso de *se* como ferramenta de indeterminação do papel temático de sujeito.

Quanto à supressão do clítico em sentenças finitas verificada por Nunes (1990), não foram coletados dados isoladamente, mas essa forma aparece competindo com a presença do clítico, como será descrito abaixo.

Por outro lado, apenas uma das afirmações de Melo (2012) foi verificada: a relação entre a posição do argumento interno e a animacidade. De fato, viu-se que em sentenças em que o argumento interno é anteposto, esses argumentos internos são [-animados] na maioria das sentenças.

O clítico *se*, se comportou então, de forma regular nas sentenças analisadas. As estruturas com *se* indeterminador e apassivador geralmente apresentaram os mesmos contextos: argumento interno [- animado] posposto ao verbo e com concordância, verbos transitivos diretos no presente e *se* pré-verbal.

Por fim, confirmou-se a hipótese principal do trabalho de que a aquisição de *se* indeterminador e de *se* apassivador dependem da escolarização. Os dados indicam que, ao contrário do *se* reflexivo, pronominal e recíproco, o *se* apassivador e o *se* indeterminador não são adquiridos naturalmente pelos falantes do PB. Seu aparecimento se dá primeiramente em material escrito e, somente próximo do fim da escolarização, seu aparecimento na fala é verificado. Inclusive, os testes estatísticos comprovaram interação entre o aparecimento do *se* e as variáveis *faixa etária* e *tipo de registro*. Em bancos de dados representativos da fala em período de aquisição e dos primeiros anos escolares, sua presença é inexpressiva, levando em torno de 12 anos de escolarização para que se faça significativa.

Também é possível afirmar que após a aquisição, esse conhecimento se assemelha a uma língua estrangeira, como propõe Kato, Correa & Cyrino (2009) e os falantes letrados alternam entre as formas de indeterminação adquiridas naturalmente, como o uso de expressões genéricas ou o apagamento de *se*, e o uso de *se* na forma de *code-switching*. Sentenças como (115) e (116) comprovam essa alternância de código. Enquanto em (115), o falante alterna entre o uso e a supressão de *se* indeterminador, o falante que enunciou (116) utiliza também a expressão genérica *a gente* para marcar a indeterminação do sujeito.

(115) Se cortava o bambu, tirava fatias de bambu. (NURC)

(116) Do jeito que se vive, a gente vive correndo. (NURC)

Atesta-se, portanto, que o conhecimento interno de um falante escolarizado do Português Brasileiro envolve duas gramáticas, tal qual Kato (2000) propôs: uma interna adquirida naturalmente, com inovações, e uma periférica adquirida na escola, conservadora e pautada nas normas da Gramática Tradicional. A primeira conteria outras formas de indeterminação e a segunda, o uso do *se* indeterminador e apassivador. Essa última gramática é acionada em situações de fala monitorada ou na escrita.

É possível concluir ainda com as considerações de Kato (1996) sobre a perda dos clíticos de terceira pessoa no PB:

O processo de letramento, seja através de dados positivos encontrados na imersão espontânea do aprendiz em textos escritos, seja através de dados negativos (intervenção corretiva da escola), apresenta, nos dados externos dos aprendizes (sua língua-E), uma evidência quantitativa do sucesso escolar

na recuperação dos fósseis, pelo menos na língua escrita. Contudo, a análise qualitativa revela indícios de que o processo difere muito da aquisição natural, pois esses dados apresentam inúmeros casos de hipercorreção e estratégias de esquiva, evidências de que há um comportamento consciente de monitoração do produto. Além disso, (...) há um conflito entre a Língua-I e o conhecimento meta-cognitivo do que é bem aceito socialmente (KATO, 1996, p. 25).

## CONSIDERAÇÕES FINAIS

A presente dissertação se propôs a estudar a aquisição e o comportamento do clítico *se* apassivador e indeterminador no PB. A principal hipótese levantada foi a de que a aquisição desse clítico não é natural para os falantes brasileiros por se tratar de um fósil linguístico recuperado pela escola, nos termos de Kato (1999, 2005a) e, dessa forma, esse processo de aquisição se dá tal qual uma L2. Para verificar essa hipótese reuniu-se um *corpora* composto por cinco bancos de dados, cada um deles representativo de um estágio da aquisição da linguagem e da escrita, como propôs (KATO, 1986), e as sentenças com *se* indeterminador e apassivador foram analisadas. O trabalho foi, então, dividido da seguinte maneira: o primeiro capítulo foi dedicado à revisão bibliográfica dos trabalhos teóricos referentes aos processos envolvidos na aquisição L2; o segundo capítulo, por sua vez, apresentou pesquisas sobre o estatuto teórico e sintático do clítico *se*; e o terceiro capítulo apontou e discutiu o resultado da coleta de dados realizada.

No primeiro capítulo, discutiu-se a problemática da aquisição de marcas estilísticas de séculos anteriores como L2 por já não fazerem parte do conhecimento interno do falante do PB. A partir do trabalho de Kato, Cyrino & Correa (2009), viu-se que os clíticos acusativos de terceira pessoa são adquiridos apenas via escolarização e, propõe-se que essa seja também a situação da aquisição do clítico *se* indeterminador e apassivador. Para discutir essa aquisição, foram apresentadas algumas das principais teorias acerca da aquisição de L2. Assumiu-se a hipótese do Acesso Parcial tal qual defendida por Kato (2005, 2003) para a aquisição da gramática escolar como L2, ou seja, o falante tem acesso parcial à GU durante a aquisição de L2 uma vez que ele produz sentenças que não estavam contidas em seu input. Também foram importantes para a investigação do fenômeno as contribuições teóricas de Roeper (1999) sobre o Bilinguismo Universal, de Hershensohn (2000) sobre o Construcionismo Linguístico, de Marcelino (2007) sobre o estágio final da aquisição de L2 e de Giusti (apud PIRES, 2017) sobre a aquisição do registro escrito em italiano. Em seguida, viu-se que, adquirida a gramática ensinada na escola, o falante escolarizado alterna entre essa e a gramática adquirida naturalmente pelos falantes do PB na forma de Code-switching, tal qual conceituado por Poplack (2001), Riehl (2005) e Toribio (2017).

No segundo capítulo, foram discutidas as principais teorias sobre o estatuto teórico do *se* apassivador e do *se* indeterminador a partir de três ópticas: da Gramática Tradicional, da Gramática Funcionalista e da Gramática Gerativa. Viu-se, por meio dos trabalhos de Rocha Lima (1994), Bechara (1992, 2009), Cipro Neto & Infante (2003), Cegalla (2007) e Cunha & Cintra (2001), que o *se* na Gramática Tradicional ainda é um dos principais recursos recomendados para realizar a indeterminação do sujeito. Sua denominação aparece como índice de indeterminação do sujeito, para *se* indeterminador, e partícula apassivadora, para *se* apassivador. Já os autores da Gramática Funcionalista – Castilho (2010), Bagno (2011), Neves (2000) e Azeredo (2008) – propõem que essa marca, embora ainda existente na língua, já não é a preferida pelos falantes uma vez que outras formas de indeterminação como o uso de expressões genéricas e verbos no infinitivo tem tomado lugar na língua. Em seguida, foram retomados alguns estudos gerativistas seminais sobre o *se* como recurso de indeterminação do sujeito em outras línguas, como o de Hyams (1986), Cinque (1988), Raposo & Uriagereka (1996) e Dobrovie-Sorin (1998). Com base nos trabalhos de Everett (1985) e de Galves (1986), Nunes (1990), principal autor da Gramática Gerativa que fundamenta esta pesquisa, por sua vez defende que, devido a um processo de reanálise sintática, o *se* apassivador passou a ser interpretado como *se* indeterminador. Essa e as outras afirmações do autor foram tomadas como alicerce deste trabalho. O autor também afirma, após analisar dados diacrônicos, que houve uma mudança linguística em relação ao *se* indeterminador em que ele passou a ser inserido em sentenças com verbos no infinitivo e no gerúndio. Para investigar melhor essa mudança, valeu-se aqui do estudo de Cavalcante (2006) sobre as construções com *se* e verbos infinitivo. A autora propõe alguns contextos favorecedores desse tipo de estrutura, os quais foram verificados nos dados coletados. Por último, considerou-se também o estudo de Melo (2012) sobre a ordem dos argumentos internos em sentenças com *se* indeterminador e apassivador.

No capítulo 3, os dados do *corpora* mostraram que o *se* indeterminador e o *se* apassivador, de fato, aparecem primeiramente em materiais escritos, durante o período da escolarização. Enquanto o *se* pronominal, o *se* recíproco e o *se* reflexivo surgem na fala de crianças em período de aquisição, os outros dependem exclusivamente da escolarização. Outro ponto interessante que parece confirmar a

hipótese de Nunes (1990) é a frequência de *se* indeterminador muito superior à de *se* apassivador. Contudo, a frequência de *se* apassivador ainda foi superior à esperada inicialmente. Notou-se também que o uso de *se* junto a verbos no infinitivo parece ter se consolidado tanto na fala quanto na escrita.

Diante desses resultados, é possível, então, responder às seguintes perguntas de pesquisa propostas no início deste trabalho:

- 1) Como se dá a aquisição do clítico *se* para falantes de PB?
- 2) Como se comporta o clítico *se* para as funções de apassivador e indeterminador antes, durante e depois da aprendizagem da escrita?
- 3) Qual a relação desses resultados com as ideias de Kato (1999a; 2005a) sobre a aquisição da gramática escolar?

A aquisição do clítico *se* indeterminador e apassivador no PB só foi verificada após 12 anos de escolarização. Antes disso, sua presença é inexpressiva. Por não se tratar de uma aquisição natural, ela acontece tal qual uma L2 e esse conhecimento passa a compor uma gramática periférica, tal qual propõe Kato (1999a; 2005a), que é acionada em situações de fala monitorada ou de escrita formal. Nos primeiros anos da escolarização, seu aparecimento se restringe a atividades de cópia ou a sentenças isoladas produzidas por crianças filhas de pais muito escolarizados que tiveram em seu input esse tipo de estrutura. No vernáculo, o clítico aparece em alternância com outras formas de indeterminação, na forma de *code-switching*. Na escrita mais escolarizada, representada por textos do jornal Folha de São Paulo, seu uso é o que mais se aproxima da norma. Contudo, apesar da Gramática Tradicional restringir os usos do *se*, os contextos em que a partícula apareceu foram muito diversificados, o que revela que não há apenas uma distância entre o que se fala e o que se escreve no PB, mas também entre a norma e a escrita.

Diante disso, como afirmou Galves (1986), o *se* ainda fará correr muita tinta. Seria interessante, então, para complementar esta pesquisa, coletar dados mais recentes tanto de fala quanto de escrita a fim de verificar se se trata de uma mudança em curso no PB. A área também carece de estudos que comparem a aquisição do clítico *se* à de outras formas de indeterminação para traçar com mais precisão o comportamento da competição entre essas formas. Além disso, a investigação sobre quais elementos compõem a gramática nuclear e a gramática periférica do falante do PB tem se tornado

um campo de pesquisa muito fértil bem como as formas em que o fenômeno de *code-switching* se dá em situações diferentes já que cada dia mais a língua aprendida na escola e a língua adquirida na escola estão se distanciando.

## REFERÊNCIAS BIBLIOGRÁFICAS

ALEXIADOU, A; DORON, E. **The syntactic construction of two non-active Voices: Passive and Middle.** Journal of Linguistics 48, 2012. p 1-34.

AZEREDO, J. C. **Gramatica Houaiss da Lingua Portuguesa.** São Paulo: Publifolha. 2008.

BAGNO, M. **Gramática pedagógica do português brasileiro.** São Paulo: Parábola, 2011.

BECHARA, E. **Moderna Gramática Portuguesa.** 36 ed.. São Paulo: Companhia Editora Nacional, 1997.

BECHARA, E. **Moderna Gramática Portuguesa.** 37<sup>a</sup>. Ed. Atualizada pelo Novo Acordo Ortográfico. Rio de Janeiro: Lucerna: 2009.

CASTILHO, A. T. **Nova Gramática do Português Brasileiro.** São Paulo: Editora Contexto. 2010.

CAVALCANTE, S. R. O. **O uso do se com infinitivo na história do português: do português Clássico ao Português Europeu e Brasileiro Modernos.** Tese de Doutorado. Campinas. Unicamp. 2006.

CEGALLA, D. P. **Novíssima gramática da língua portuguesa.** 46<sup>a</sup> ed. São Paulo: Companhia Editora Nacional, 2007.

CHOMSKY, N. **Lectures on Government and Binding.** Dordrecht: Foris. 1981.

CHOMSKY, N. **Knowledge of Language: It's nature, origin and use.** Praeger, New York. 1986.

CORRÊA, V. R. **Oração relativa: O que se Fala e o que se Aprende no Português do Brasil.** Tese de Doutorado, Campinas, Unicamp. 1989.

CYRINO, S. M. L. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: ROBERTS, I.; KATO, M. A. (orgs.) **Português Brasileiro: uma viagem diacrônica**, Campinas: Editora da UNICAMP. 1993 pps

CINQUE, G. **On si constructions and the theory of arb.** Linguistic Inquiry 19, 1988, p. 521-582.

CIPRO NETO, P. & INFANTE, U. **Gramática da Língua Portuguesa.** São Paulo: Editora Scipione. 2003

CUNHA, C. & LINDLEY, C. **Nova Gramática do Português Contemporâneo.** Rio de Janeiro: Nova Fronteira, 3a ed..2001.

DOBROVIE-SORIN, C. **Impersonal se constructions in Romance and the assivization of unergatives.** Linguistic Inquiry 29, 1998. p. 399-437.

DRACHMAN, G. Some properties of clitics (with special reference to Modern Greek), in: ALEXIADOU, A. & HALL, A. T. (eds.) **Studies on Universal Grammar and Typological Variation.** Amsterdam: John Benjamins, 1997. p. 219–248.

DUARTE, M. E.L.; KATO, M. A.; BARBOSA, P. Sujeitos indeterminados em PE e PB. Anais do II Congresso Internacional da ABRALIN, vol. I, p. 405-409, 2001.

GALVES, C. **Ensaio sobre as gramáticas do português.** Campinas, SP: Editora da Unicamp. 2001.

HYAMS, N. The Interpretation and Acquisition of Italian Impersonal si. In: BORER, H. **Syntax and Semantics 19: The Syntax of Pronominal Clitics**, ed. Hagit Borer. Orlando: Academic Press. 1986. p. 219 - 240.

KATO, M.A. **No Mundo da Escrita**. São Paulo: Editora Ática. 1986.

KATO, M.A. **VARIAÇÃO SINTÁTICA E ESTILO**. Cadernos de Estudos Lingüísticos (UNICAMP), CAMPINAS, v. 22, 1992, p. 127-137.

KATO, M. A. **Metalinguistic reflections on Code-Switching**. In: BARBARA, L; SCOTT, M (orgs). *Reflections on Language Learning*. Clevedon: Multilingual Matters. 1994. p. 122-137.

KATO, M. A. **Português Brasileiro falado: aquisição em contexto de mudança lingüística**. In: DUARTE, M. E.; LEIRIA, I. (eds). *Actas do Congresso Internacional sobre o Português*. Vol II, 1996, pp.211-237.

KATO, M. A. Aquisição e aprendizagem da língua materna: de um saber inconsciente para um saber metalingüístico. In: MORAES, J; GRIMM-CABRAL, L. (orgs) **Investigações a linguagem: ensaios em homenagem a Leonor Scliar-Cabral**. Florianópolis: Editora Mulher. 201-225. 1999a.

KATO, M.A. **Variação e mudança no Português Brasileiro: problemas de aquisição**. A Cor das Letras (UEFS), Feira de Santana, n.3, 1999b, p. 159-178,

KATO, M. A. **Questões atuais da aquisição de L1 na perspectiva de Princípios e Parâmetros**. Cadernos de Estudos Lingüísticos, v. 16. 1999c. p. 11-16.

KATO, M. A. The permeability of grammars in bilingual language acquisition. In: **Bilingualism: Language and Cognition**, 4(01), 2001. 31-32.

KATO, M. A. A gramática do letrado: questões para a teoria gramatical. In: MARQUES, M. A.; TEIXEIRA, J.; LEMOS, A. S. (Orgs.). **Ciências da linguagem: trinta anos de investigação e ensino**. Braga: Universidade do Minho. 2005a. p. 131 – 145.

KATO, M.A. **A contribuição chomskiana para a compreensão da aprendizagem de L2.** *Trabalhos em Lingüística Aplicada, Campinas*, v. 44, n.2, 2005b, p. 185-199.

KATO, M. A. Acquisition in the context of language change: the case of the null subject in Brazilian Portuguese. In: RINKE, E.; KUPISCH, T. (eds.): **The Development of Grammar: Language Acquisition and Diachronic Change** (in Honour of Jürgen M. Meisel) Hamburg Series on Multilingualism (HSM 11). Amsterdam/ Philadelphia: John Benjamins. 2011. p. 309-330.

KATO, M. A. A contribuição social do modelo formal chomskiano. In: VIAN JR, O.; CATABIANO, C. (Org.). **Linguagem e suas múltiplas faces.** 1ª ed. São Paulo: Mercado de Letras, 2013, p. 13-26.

KATO, M A. **A variação no domínio dos clíticos no português brasileiro.** *Lingüística, Montevideo* , v. 33, n. 1, p. 135-145, 2017 .

KATO, M. A.; TARALLO F. Anything YOU can do in Brazilian Portuguese. In: JAEGLI, O.; SILVA-CORVALAN, C. (orgs) **Studies in Romance Linguistics.** Foris. 1986. p. 343-358.

KATO, M. A.; TARALLO F. **Harmonia Transistêmica:** Variação intra e interlingüística. *Preedição*, 5. Campinas. 1989.

KATO, M. A.; CYRINO, S. M. L.; CORRÊA, V.R. Brazilian Portuguese and the recovery of lost clitics through schooling. In: PIRES, A.; ROTHMAN, J. (eds.) **Minimalist Inquiries into Child and Adult Language Acquisition:** Case Studies across Portuguese. Berlin/New York: Mouton De Gruyter. 2009. p. 245-272.

KENEDY, E. **Curso básico de linguística gerativa.** São Paulo: Contexto, 2013.

KROCH, A. **Mudança sintática**. Traduzido por Silvia Cavalcante. 2003. Edição online. Disponível em: <<http://www.ling.upenn.edu/kroch>>. Acesso em 20-06-2015.

LUFT, C. P. **Gramática Resumida**. Porto Alegre: Editora Globo. 1976.

MAGALHÃES, T. M. V. **A aquisição de pronomes sujeitos no PB e no PE**. Letras de Hoje. Porto Alegre, v. 42, n. 1, p. 97-112, 2007. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/viewFile/673/488>>. Acesso em: 25/06/2015.

MAGALHÃES, T. M. V. **Aprendendo o Sujeito Nulo na Escola**. UNICAMP: Dissertação de Mestrado. 2000.

MANZINI, M. R. On Italian 'si'. In: BORER, H. **Syntax and Semantics 19: The Syntax of Pronominal Clitics**, ed. Hagit Borer. Orlando: Academic Press. 1986. p. 241 - 262.

MANZINI, M. R. & SAVOIA, L. Clitics: Cooccurrence and mutual exclusion patterns. In: RIZZI, L. (ed.). **The structure of CP and IP**. New York: Oxford University Press. 2004. p. 211-250.

MELO, E. S. **Construções com se: evidências da emergência de gramática do Português Brasileiro**. Dissertação de Mestrado. Rio de Janeiro. UFRJ. 2012.

NEGRÃO, E. V. ; VIOTTI, E. Estratégias de impessoalização no português brasileiro. In: José Luiz Fiorin; Margarida Petter. (Org.). **África no Brasil: a formação da língua portuguesa**. São Paulo: Contexto, 2008, p. 179-203.

NEGRÃO, E. V. ; VIOTTI, E. **Elementos para a investigação da semântica do clítico se no português brasileiro**. Cadernos de Estudos Linguísticos, Campinas, v. 57, n. 1. 2015.

NEVES, M. H. M. **Gramática de usos do português**. São Paulo : Ed. UNESP. 2000

NEVES, M. H. M. **Guia de uso do português: Confrontando regras e usos.** So Paulo : Ed. UNESP. 2003.

NUNES, J. M. **O Famigerado se: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador.** Dissertação de Mestrado. Campinas. Unicamp. 1990.

NUNES, J. Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em português brasileiro. In: ROBERTS, I.; KATO, M.(org.). **Português brasileiro: uma viagem diacrônica.** Campinas: UNICAMP, 1996.pps

O ESTADO DE S. PAULO. **Manual de redação e estilo.** São Paulo: O Estado de S. Paulo, 2010.

PAGOTTO, E. Clíticos, mudança e seleção natural. In: ROBERTS, I. & kato, M. A. **Português brasileiro: uma viagem diacrônica,** Campinas, Editora da Unicamp: 185-206. 1993.

PIRES, L. C. **Aquisição da escrita-padrão do português brasileiro em esfera escolar: um estudo transversal.** Tese de Doutorado. Brasília. UnB. 2015.

POPLACK, S. Code-switching (linguistic). In: **International encyclopedia of the social and behavioral sciences.** Edição on-line. 2001. Disponível em: <<http://aix1.uottawa.ca/~sociolx/CS.pdf> > Acesso em 24-06-2015.

RAPOSO, E. URIAGEREKA, J. Indefinite SE. In: DIKKEN, Marcel Den (ed.) **Natural Language and Linguistic Theory.** New York: Springer, 1996. p. 749–810.

ROCHA LIMA, C. H. **Gramática normativa da língua portuguesa.** 32. ed. Rio de Janeiro: José Olympio, 1994.

MARCELINO, M. **O Parâmetro de Composição e a Aquisição/Aprendizagem de L2**. Tese de Doutorado. Campinas. Unicamp. 2007.

RIEHL, C. M. Code-switching in Bilinguals: Impacts of Mental Processes and Language Awareness. In: **ISB4**: Proceedings of the 4th International Symposium on Bilingualism. Edição on-line. 2005. Disponível em: <<http://www.lingref.com/isb/4/151ISB4.PDF>> Acesso em 24-06-2015.

TORIBIO, J. A. **Structural approaches to code-switching**. In: Selected papers from the 45th Linguistic Symposium on Romance Languages (LSRL), Campinas, Brazil. 2017

VERGNAUD, J. R. Letter to Noam Chomsky and Howard Lasnik (1976). In: ROBERT, F; LASNIK, H. (eds.) **Syntax**: Critical concepts in linguistics. Vol. 5. 2006. 21-34. London: Routledge.

WEHRLI, E. On Some Properties of French Clitic se. In: BORER, H. **Syntax and Semantics 19**: The Syntax of Pronominal Clitics, ed. Hagit Borer. Orlando: Academic Press. 1986. p. 263 – 284

WHITE, L. **Second Language Acquisition and Universal Grammar**. Cambridge: Cambridge University Press. 2003.

XAVIER, G. R.. **Português Brasileiro como segunda Língua**: um Estudo sobre o Sujeito Nulo. Tese de Doutorado. Campinas. Unicamp. 2006.