

Bruno Fernandes Zenóbio de Lima

O percurso diacrônico das construções com o pronome se na Língua Portuguesa como um processo de gramaticalização.

Tese apresentada o Programa de Pós-Graduação em Estudos Lingüísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Doutor em Lingüística.

Área de concentração: Lingüística
Linha de Pesquisa: Estrutura Gramática da Linguagem
Orientador: Prof. Dr. Lorenzo Teixeira Vitral.

Belo Horizonte

Faculdade de Letras da UFMG

2006

Miragem

*Uma miragem
na linha eterna do tempo:
uma palavra, o sinal....
A ligação com o eterno
no desvelamento do efêmero.*

O Principiante

*Sua mão mal se movimenta,
custa a escorregar pela mesa,
caracol no jardim da ciência,
desenrolando letra a letra
a obscura linha de seu nome.*

*Ah, como é leve o átomo puro,
e ágil o equilíbrio do mundo,
e rápido, e célere, o curso
e o céu, do destino de tudo!*

*Mas na terra o pálido aluno
Devagar escreve o seu nome.*

Cecília Meireles

*Aos humanos, só é possível caminhar
numa selva de códigos íntimos e bizarros,
tão próprios quanto a sua essência,
tão destrutivos quanto o seu fim.*

*Miragem e fantasia!
Para a redenção do mundo em sua ilusão
para a invenção sempre sagrada,
vestimenta de símbolos de poeira...*

Bruno Zenóbio

Apreensão

*Eis a lição da vida: a apreensão.
Não unicamente do que é apreendido,
ou daquele que apreende,
mas de sua fusão
nem cega, nem objetiva,
nem absurda, nem redentora.*

*Fusão que é o imbricamento de contrários,
dynamitando ilusórios caminhos opostos.
A eternidade é a amplidão da finitude.
A língua é a economia do homem.
O saber para sempre!
Eis a única prova de superação da humanidade.*

Bruno Zenóbio

AGRADECIMENTOS

Gostaria de agradecer ao professor Lorenzo Vitral por ter acreditado neste trabalho, pela constante dedicação e fé na pesquisa científica.

Agradeço às Professoras Jânia Martins Ramos e Maria Antonieta A. M. Cohen pelas precisas observações no exame de qualificação.

Agradeço também aos Professores Odete Pereira da Silva Menon, Charlotte Marie Chamberland Galves, Fábio Bonfim Duarte e Maria do Carmo Viegas, pelas considerações finais.

Agradeço ao Professor Anelito pelas imprescindíveis orientações nos cálculos estatísticos, ao colega e amigo Eduardo Tadeu Amaral pela atenciosa revisão, e ao também colega e amigo Marcos Feitosa pela tradução do resumo.

Agradeço à Universidade Federal de Minas Gerais e em especial ao Programa de Pós-Graduação em Linguística – POSLIN, pela oportunidade de realização desse trabalho.

Agradeço ao apoio da família, em especial, meus pais Ana e Marco, e minhas tias Ângela e Angélica.

SIGLAS DOS TEXTOS UTILIZADOS

PERÍODO ARCAICO

CDJ - Cronica d'el Rei Dom Joham
TLL – Trecho do Livro de Linhagens
VER – Vereações
PED - Carta que o Jfante dom pedro emujou a el rey de Brujas.
CAR - A carta de Pero Vaz de Caminha

PERÍODO MODERNO

AVE - Novelistas e contistas portugueses dos séculos XVII e XVIII.
CJB - Cartas dos primeiros jesuítas no Brasil
TTB - Tratado da Terra do Brasil – Gandavo
COB - Cultura e opulência do Brasil na lavra do açúcar. Engenho real moente e corrente
SES – Carta de Sesmaria ao Coronel Mathias Barbosa da Silva

PERÍODO CONTEMPORÂNEO

CFP - Hoje em Dia - Crônicas Futebol e Política de Minas
BUL - Bulas de remédio
DGC - Desenvolvimento de Gráficos de Controle Aplicados
SAR – Sarapalha - Guimarães Rosa

ENTREVISTAS

CMP - Transcrições das gravações realizadas em Campanha, Minas Novas e Paracatu

ÍNDICE DAS TABELAS E GRÁFICOS

TABELAS:

- Tabela 1:** Descrição das porcentagens e frequências por tipo no *Trecho do Livro de Linhagens*, 56
- Tabela 2:** Descrição das porcentagens e frequências por tipo na *Cronica d'el Rei Dom Joham*, 57
- Tabela 3:** Descrição das porcentagens e frequências por tipo em *Vereações*, 58
- Tabela 4:** Descrição das porcentagens e frequências por tipo na *Carta que o Jfante dom pedro emujou a el rey de Brujas*, 59
- Tabela 5:** Descrição total das porcentagens e frequências por tipo no período arcaico, 60
- Tabela 6:** Descrição das porcentagens e frequências por tipo nas *Cartas dos primeiros jesuitas no Brasil*, 62
- Tabela 7:** Descrição das porcentagens e frequências por tipo em *Tratado da Terra do Brasil*, 63
- Tabela 8:** Descrição das porcentagens e frequências por tipo em *Novelistas e contistas portugueses dos séculos XVII e XVIII*, 64
- Tabela 9:** Descrição das porcentagens e frequências por tipo em *Cultura e opulência do Brasil na lavra do açúcar. Engenho real moente e corrente*, 65
- Tabela 10:** Descrição das porcentagens e frequências por tipo em *Carta de Sesmaria ao Coronel Mathias Barbosa da Silva*, 66
- Tabela 11:** Descrição das porcentagens e frequências por tipo no período moderno, 67
- Tabela 12:** Descrição das porcentagens e frequências por tipo em *Hoje em Dia - Crônicas Futebol e Política de Minas*, 69
- Tabela 13:** Descrição das porcentagens e frequências por tipo em *Hoje em Dia - Crônicas Futebol e Política de Minas*, 69
- Tabela 14:** Descrição das porcentagens e frequências por tipo em *Bulas de remédio*, 70
- Tabela 15:** Descrição das porcentagens e frequências por tipo em *Desenvolvimento de Gráficos de Controle Aplicados*, 71
- Tabela 16:** Descrição das porcentagens e frequências por tipo no período contemporâneo, 73
- Tabela 17:** Descrição das porcentagens do grupo reflexivo nos períodos analisados, 74
- Tabela 18:** Descrição das frequências relativas do grupo reflexivo nos períodos analisados, 75
- Tabela 19:** Descrição das porcentagens dos tipos do grupo não-reflexivo nos períodos analisados, 77
- Tabela 20:** Descrição das frequências relativas dos tipos do grupo não-reflexivo nos períodos analisados, 78
- Tabela 21:** Descrição das frequências relativas dos tipos do pronome *se* nos períodos analisados, 80
- Tabela 22:** Descrição das porcentagens e frequências relativas por texto no período contemporâneo, 87
- Tabela 23:** Descrição das porcentagens e frequências relativas por texto no período moderno, 91

Tabela 24: Descrição das porcentagens e frequências por tipo em *Projeto POBH*, 102

Tabela 25: Descrição das porcentagens e frequências por tipo em *Transcrições das gravações realizadas em Campanha, Minas Novas e Paracatu*, 103

Tabela 26: Descrição da frequência por verbo em relação à frequência total, 106

Tabela 27: Descrição da frequência por verbo em relação à frequência total de cada verbo, 107

GRÁFICOS

Gráfico 1: Linha das porcentagens totais do grupo reflexivo nos períodos analisados, 75

Gráfico 2: Descrição diacrônica das frequências relativas do grupo reflexivo nos períodos analisados, 76

Gráfico 3: Descrição diacrônica dos valores totais frequências relativas do grupo reflexivo nos períodos analisados, 76

Gráfico 4: Linha das porcentagens totais dos tipos o grupo não-reflexivo períodos analisados, 78

Gráfico 5: Descrição diacrônica das frequências relativas dos tipos do grupo não-reflexivo nos períodos analisados, 79

Gráfico 6: Descrição diacrônica dos valores totais das frequências relativas dos tipos do grupo não-reflexivo nos períodos analisados, 79

Gráfico 7: Linha de tendência e dos valores totais das frequências relativas dos tipos do pronome **se** nos períodos analisados, 81

Gráfico 8: Linha das distribuições de ocorrência esperadas, 82

Gráfico 9: Linha das distribuições de ocorrência reais, 83

SUMÁRIO

INTRODUÇÃO.....	VIII
-----------------	------

CAPÍTULO I: O PRONOME SE: CLASSIFICAÇÃO E CONSIDERAÇÕES DIACRÔNICAS

1.1 Apresentação	10
1.2 A classificação de Cunha (2001)	11
1.3 Luft (1987) e os pronominais	12
1.4 Mattoso e conceito de voz medial	13
1.5 A classificação de Kury	14
1.6 Aguiar e a evolução do pronome se	16
1.7 A análise diacrônica de Maurer Jr.	17
1.8 Said Ali e a negação da indeterminação	20
1.9 Naro e a reanálise do pronome se	21
1.10 Nunes e o percurso do pronome se	23
1.11 Considerações finais	25

CAPÍTULO II: SUPORTE TEÓRICO

2.1 Apresentação	28
2.2 Origens da teoria e definições de Gramaticalização	28
2.2.1 Continuidade e Gradualismo	36
2.2.2 Unidirecionalidade.....	37
2.3 Perspectiva formal dos processos de gramaticalização	38

CAPÍTULO III: METODOLOGIA

3.1 Apresentação	41
3.2 Problemas Teóricos	41
3.3 Critérios de Identificação dos Processos de Gramaticalização	42
3.4 Tratamento do Corpus	45
3.5 Classificação das ocorrências	47

CAPÍTULO IV: ANÁLISE QUANTITATIVA

4.1 Apresentação	55
4.2.1 Período Arcaico	56
4.2.1.1 Perfil por texto e tipo	56
4.2.2.1 Perfil geral	60
4.2.2 Período Moderno	61
4.2.2.1 Perfil por texto e tipo	63
4.2.2.2 Perfil geral	67
4.2.3 Período Contemporâneo	68
4.2.3.1 Perfil por texto e tipo	68
4.2.3.2 Perfil geral	72
4.3 Análise comparativa dos períodos	74
4.3.1 Grupo reflexivo	74
4.3.2 Grupo não-reflexivo	77
4.3.3 Análise global comparativa dos períodos e tipos.....	80
4.4 Cálculo da Probabilidade	81
4.4.1 Detalhamento do cálculo e Interpretação Lingüística	81
4.5 O fator gênero.....	84
4.5.1 Apresentação	84

4.5.2 Período contemporâneo	85
4.5.3 Período moderno	89
4.5.4 Conclusões	92

CAPITULO V: IMPLICAÇÕES TEÓRICAS E CONCLUSÕES

5.1 Apresentação	95
5.2 Síntese dos resultados e a gramaticalização de se	95
5.2.1 Análise das entrevistas	102
5.3 Graus de gramaticalidade do pronome se	107
5.4 O estatuto gramatical do grupo reflexivo.....	109
5.5 O estatuto gramatical do grupo não-reflexivo.....	115
5.6 Considerações finais.....	119

ABSTRACT

This paper consists basically of a diachronic analysis of the pronoun “se” in Portuguese, with the hypothesis that its trajectory may be dealt with as a grammaticalization process. Thus, this analysis aims to investigate how the pronoun “se” behaves throughout the history of the Portuguese language regarding its frequency, as well as its syntactical and semantic characteristics. For such, the theoretical framework adopted is the one proposed by Hopper & Traugott (1993) and Heine et al. on grammaticalization, as well as the methodology developed by Vianna (2000) and Vitral (2005). Based on the results of the analysis made, I intend to verify whether or not the historical course of the pronoun “se” could be dealt as a case of grammaticalization.

INTRODUÇÃO

Esse trabalho constitui-se, fundamentalmente, de uma análise diacrônica do pronome **se**, considerando, como hipótese, que sua trajetória seja tratada como um processo de gramaticalização. Para a realização de tal tarefa, parto do quadro teórico proposto por Hopper & Traugott (1993), e Heine *et alii*, sobre a gramaticalização, bem como a metodologia desenvolvida por Vianna (2000) e Vitral (2005).

Constitui-se, então, em objetivo central do trabalho: *investigar, dentro da perspectiva teórico-metodológica supracitada, como o pronome se se comporta ao longo da história da língua portuguesa, em termos de frequência e de suas características sintáticas e semânticas*. De acordo com os resultados aferidos por meio da nossa análise, pretendemos confirmar se o percurso histórico do pronome **se** poderia ser tratado como um caso de gramaticalização.

Além do objetivo central, outros mais específicos podem ser considerados também, dentre eles, alguns que não foram devidamente explicados pelos teóricos da literatura até o momento, como: a) explicar como ocorreu a reanálise *se-reflexivo* > *se-passivo* > *se-indeterminado*, ou seja, como o uso reflexivo do pronome **se** do português pôde dar origem a formas do grupo não-reflexivo; b) analisar como ocorreu, em termos de frequência, a expansão dos subtipos funcionais da forma **se** no sistema da língua portuguesa, comparando-os; c) considerar o surgimento de novas formas de sujeito indeterminado e como elas podem estar “ocupando” o lugar da forma **se**; d) aplicar, testar e corroborar a metodologia de Vitral (2005) que pode tornar possível a identificação objetiva dos processos de gramaticalização. Diante então dos objetivos acima referidos, está estruturado esse trabalho em cinco capítulos.

O capítulo I, intitulado: “O pronome **se**: classificação e considerações diacrônicas”, apresenta uma revisão da literatura sobre o assunto. Meu objetivo nesse

capítulo é apresentar os trabalhos mais relevantes para a problemática em questão, em suas diversas formas de contribuição.

O capítulo II, intitulado: “Suporte Teórico”, apresenta o quadro teórico da gramaticalização, considerando a origem do conceito, e abordando também seus princípios e mecanismos, bem como a apresentação de uma perspectiva formal de compreensão de tais fenômenos.

No capítulo III, que se intitula “Metodologia”, apresento as diretrizes metodológicas do trabalho. Dessa forma, abordo questões relativas aos *corpora*, que envolvem desde a justificativa para sua escolha, como a periodicidade, até, por exemplo, a descrição das estruturas e verbos mais recorrentes com o pronome *se* nos textos e períodos analisados.

Delineadas as diretrizes metodológicas, passo a apresentar no capítulo IV, intitulado “Análise Quantitativa”, a descrição dos resultados das análises quantitativas. Entretanto, inicio o capítulo com duas seções que explicam a natureza dos cálculos da probabilidade e da frequência, no intuito de esclarecer ao leitor, não familiarizado com tais cálculos, sua relevância para os estudos lingüísticos. Em seguida, apresento os resultados por tipos de *se* nos três períodos analisados.

Finalmente, no capítulo V, intitulado “Implicações teóricas e conclusões”, discuto a hipótese de gramaticalização do pronome *se* considerando seu percurso histórico, por meio de teorias lingüísticas e da análise crítica dos dados extraídos dos *corpora*.

CAPÍTULO I - O PRONOME SE: CLASSIFICAÇÃO E CONSIDERAÇÕES DIACRÔNICAS

1.1 Apresentação

5

Apresentados na introdução desse trabalho o meu objeto de análise, e os objetivos dessa investigação, faço então nesse capítulo uma breve revisão da literatura a respeito do pronome *se*. Essa revisão constitui-se, então, da apresentação das principais contribuições para a problemática diacrônica desse pronome.

10

Cabe, entretanto, ressaltar que as contribuições são de ordem diversa e fragmentada, pois se diferem tanto do interesse, quanto do quadro teórico dos autores em questão. Serão, assim apresentadas, recortando-se nelas o que há de relevante para análise, já que muitas são repetitivas, não trazendo nenhuma elucidação a mais para o problema.

15

Nesse conjunto de contribuições, há, por exemplo, os gramáticos que se preocupam mais em classificar os tipos do pronome *se*. Há, por outro lado, um grupo de lingüistas que se preocupam mais em compreender o uso do *se* questionando se ele tem tal estatuto pronominal, ou se é clítico, morfema, ou afixo, etc. Há, por último, outro grupo de lingüistas, os diacronistas, que se preocupam basicamente em compreender a origem e o desenvolvimento histórico do pronome *se*.

20

Será então seguindo essas orientações teóricas que exporei tais contribuições: primeiro os gramáticos, posteriormente, os sincronistas, e, finalmente, os diacronistas. Ressalte-se, entretanto, que às vezes essas perspectivas e interesses se fundem, como é o caso de Maurer Jr. (1951) que não só classifica e propõe um estatuto teórico para o pronome, como também explica seu desenvolvimento histórico a partir da voz medial.

25

Passo então às explicações.

1.2 A classificação de Cunha e Cintra (2001)

Início minha apresentação pelo trabalho de CUNHA e CINTRA (2001) intitulado
5 *Nova Gramática da Língua Portuguesa*. Sua escolha se deve ao fato de que, a meu ver, ele
concentra o maior número de usos do pronome **se**, sendo esses usos apresentados de forma
objetiva.

Isso, obviamente, não quer dizer que não sejam encontrados em outras gramáticas
ou compêndios os mesmos usos. Contudo, ressalto que, a meu ver, o trabalho supracitado
10 sintetiza o que foi encontrado em termos de uso do pronome **se** numa perspectiva mais
descritiva. Os tipos descritos pelos autores são apresentados abaixo¹:

(1)

a) Pronome em função de *Objeto direto*

Ele se feriu.

15 b) Pronome em função de *Objeto indireto*

Perguntava-se a si mesma sobre o ocorrido.

c) Pronome como *Sujeito de um infinitivo*

Paulo deixou-se cair ao chão.

d) *Pronome apassivador*

20 *Ouvm-se ainda os sons da madrugada.*

e) Pronome em função de *Símbolo de Indeterminação do Sujeito*

¹ (CUNHA E CINTRA, 2001, p.305-307)

Come-se bem naquela cidade.

f) *Palavra Estilística* (realce)

Pois ela se morreu fria.

g) Pronome como *Parte integrante de certos verbos*

5 *Ele arrependeu-se do que fez.*

Veremos ao longo das próximas exposições dos tipos do pronome **se**, que há uma grande variação na sua denominação. Por exemplo, para a função assinalada em (1g) utiliza-se as seguintes denominações: a) *pronominal*; b) *dinâmica*; c) *voz medial pronominal*. Há, por exemplo, tipos não assinalados por outros autores, como é o caso de *b*, já que se costuma
10 tratar todas as construções reflexivas com **se** como transitivas diretas.

1.3 Luft (1987) e os pronominais

Celso Pedro LUFT (1987) em seu *Dicionário Prático de Regência Verbal*
15 apresenta uma classificação dos tipos do pronome **se**, considerando especificamente seus usos pronominais. O propósito de Luft é propor uma classificação considerando todos os empregos do **se** (estilístico, pronominal, recíproco, etc.²) como sendo reflexivos. Essas formas reflexivas, segundo o autor, seriam diferentes expressões da forma medial reflexiva. Vejamos, então, abaixo a classificação do autor:

20 (2)

a) Reflexiva simples

Ele se viu.

b) Reflexiva recíproca

² O autor utiliza outra nomenclatura para as funções do pronome **se**, como se vê adiante.

Os amigos se cumprimentaram.

c) Reflexiva dinâmica (exprime dinamismo do sujeito em si mesmo)

Eu me levanto cedo

d) Reflexiva enfática (estilística ou metafórica)

5 *Ele se riu..*

e) Reflexiva essencial ou obrigatória (porque o verbo nunca se usa se o pronome)

Ninguém se queixou do que foi dito.

A meu ver, a contribuição de Luft para a literatura é agrupar os subtipos dos usos do pronome **se** com sentido medial reflexivo em um grande conjunto. Ou seja, de certa forma, os tipos apresentadas acima (2a-e), demonstram um certo grau de reflexividade, isto é, são todas originárias do antigo uso da voz medial reflexiva, que será mais bem analisada nas próximas seções. Desse modo, o autor as considera todas (2a-e) como pronominais, pois o sujeito é de alguma forma afetado pela sua própria ação, enquanto que os usos passivo e indeterminado do pronome, embora não considerados em tal trabalho e que apresentam outras conotações semânticas, se classificam em outros grupos.

1.4 Mattoso e o conceito de Voz Medial

Mattoso Câmara Jr. no seu *Dicionário de Lingüística e Gramática* faz uma explanação do conceito de voz medial. Conceito que, a meu ver, apresenta grande relevância para a compreensão diacrônica e do funcionamento do pronome **se**.

Para o autor, a voz medial corresponde: “em português a uma construção em que a forma do verbo na voz ativa se adjunge ao pronome átono, referente à pessoa do sujeito; *Eu me feri, Tu te feriste, Ele se feriu*, etc. É, pois, uma construção pronominal (Câmara

196:164)”.Ou seja, de certa forma, mesmo nas variações que a voz medial assume, sempre se caracteriza por um envolvimento ou uma integração do sujeito à sua própria ação, ou seja, *o sujeito é, de alguma forma, afetado pela sua própria ação.*

A voz medial apresenta as seguintes acepções: (1) voz medial reflexiva, (2) voz
5 medial dinâmica, (3) voz medial estilística e (4) voz medial passiva, assim respectivamente
exemplificadas: (1) *Eu me feri* (sendo pronome objeto de uma ação do verbo); (2) *Eu me
levantei* (sendo o pronome o centro de uma ação verbal transitiva, que parte dele mas não sai
do seu âmbito; (3) *Ele se riu* (sendo o pronome responsável pela intensificação da ação
realizada pelo sujeito; (4) *Dessa maçã se vende bem* (sendo o pronome o sujeito de uma ação
10 recebida de um agente exterior indeterminado). Essas construções são todas de interesse da
análise aqui proposta, já que o pronome *se* aparece empregado em todas elas.

Em relação às origens da voz medial, Mattoso afirma que a construção medial
surge nas antigas línguas indo-européias, especificamente no sânscrito e no grego, como um
artifício para indicar, por meio de flexões, a integração do sujeito nas ações que dele se
15 originam como, por exemplo, o gr. *Λυομαι* “eu me solto”.

1.5 A classificação de Kury

Outro relevante trabalho, em termos de classificação dos tipos do pronome *se*, é o
20 de Adriano da Gama Kury, intitulado *Lições de Análise Sintática*. O objetivo do autor é
apresentar uma consistente delimitação e classificação dos tipos do pronome *se*, dividindo-as
em dois grupos: a) os tipos do grupo reflexivo, e; b) o grupo não-reflexivo.

Em relação à voz medial ou reflexiva, como denomina o autor, pode ser definida
como: “quando a ação denotada por um verbo transitivo direto é simultaneamente exercida e
25 recebida pelo mesmo ser, diz-se que o verbo, então acompanhado de pronome, está na voz
medial ou reflexiva”(KURY,1951,p.32). Salienta, entretanto, o autor que a NGB

(Nomenclatura Gramatical Brasileira) também considera outros casos como também sendo relativos à voz medial. Vejamos a classificação dos tipos mediais (3a-e), de acordo com o gramático:

(3)

5 a) **Voz Reflexiva**: aparece com verbos transitivos diretos que possuem como objeto o pronome de qualquer pessoa gramatical: *Eu me lavo, Ele se lava*;

b) **Voz medial recíproca**: quando a ação de um verbo transitivo direto sempre no plural é exercida reciprocamente: *Eles se amavam, Eles se cumprimentaram*.

10

c) **Voz medial dinâmica**: Exprime uma ação que o sujeito executa identicamente em si e em outra pessoa, sem haver necessariamente uma idéia de reflexividade: *Ele arremessou-se sobre inimigo; Afastei-me do fogo*;

15

d) **Realce**: Exprime a vivacidade ou espontaneidade do movimento da ação executada pelo sujeito: *Ele ria-se à toa; Foi-se embora*.

20

e) **Voz medial pronominal**: na qual o pronome não tem nenhum valor expressivo, ou seja, constitui-se como um elemento fossilizado, e inseparável do verbo: *Ele se arrependeu do que disse; Ele se queixou do atendimento*.

25

Outro emprego do pronome **se** assinalado por Kury é na formação da voz passiva pronominal. O autor considera que quando numa oração na voz ativa com verbo transitivo direto o agente é indeterminado e o paciente é inanimado, e, por isso, incapaz de exercer a função de sujeito, essa construção pode admitir uma voz passiva, como em *Construíam-se muitos edifícios*. Caracteriza-se, então, a voz passiva por uma *perspectiva* diferente de foco da

oração, ou seja, o foco (perspectiva) que é estabelecido nas construções ativas no sujeito, passa a ser, no caso da construção passiva, no objeto.

Com a evolução histórica da língua portuguesa, as passivas pronominais estenderam seu emprego a todos os tipos de verbos, como os transitivos indiretos e intransitivos, como por exemplo em orações com sujeito indeterminado: *Obedeça-se às normas do jogo; Também se vive bem lá*. A esse tipo de construção, que se distingue das passivas pronominais pela ausência de sujeito, Kury denomina de passiva-indeterminada.

Entretanto, o autor, em momento algum, torna equivalentes as denominações de sujeito indeterminado e passiva-indeterminada, já que considera as passivas indeterminadas como orações desprovidas de sujeito, e as passivas pessoais como providas de sujeito, mesmo que esses estejam indefinidos. É proposta então a idéia de que o pronome *se* integra à forma impessoal dos verbos, ou seja, apresenta uma *função indeterminadora*.

A freqüência, por sua vez, dessa conjugação indeterminada com pronome *se* com verbos intransitivos e transitivos indiretos levou a uma extensão do emprego das passivas indeterminadas com verbos transitivos diretos.

1.6 Aguiar (1942) e a “evolução” do pronome *se*

Com o objetivo além de simplesmente classificar os usos do pronome *se*, destaco o trabalho de Martins de Aguiar, *Notas e Estudos de Português* (1942), como um dos pioneiros na tarefa de compreender o percurso diacrônico do pronome *se*, ou sua “evolução”, como se refere o autor, a esse percurso. Aguiar traça, então, cinco estágios que comporiam essa “evolução”.

O autor indica que o primeiro uso do pronome **se**, que surge já no latim, é o de pronome reflexivo (estágio 1), como em *O homem cortou-se*. Segundo ele, as construções reflexivas envolvem tanto uma leitura ativa, quanto uma passiva, e que essa última parece nos
5 “impressionar” mais. Dessa impressão, então, se passa ao estágio 2 da “evolução”, que é a passiva. Ou seja, como a forma passiva nos “impressiona” mais do que a ativa, ela prevalece a partir das construções reflexivas, como em *Fritam-se os ovos*, ou seja, a leitura seria *Os ovos são fritos* ou *Os ovos se fritam*, e não os *Os ovos se fritam a si mesmos*.

Devido à indefinição de agente em frases como: *Comem-se os bolos*, o pronome **se**
10 passa a ser interpretado na função de indeterminação do sujeito (agente). Esse seria então o Estágio 3 da evolução, e que consistiria na função de pronome indeterminador do agente.

Os estágios 4 e 5 são os que o pronome desempenha a função de indeterminador do agente em verbos intransitivos e transitivos. Correspondem esses estágios respectivamente à expansão do uso da indeterminação do agente com verbos intransitivos (4) *Trabalha-se muito aqui*, e transitivos (5) *Vende-se casas no condomínio*.
15

Desconsiderando as naturais restrições metodológicas do trabalho, já que não era intuito do autor a observância de outros aspectos da evolução do pronome, cabe aqui ressaltar, entretanto, que ele representa uma incipiente discussão a respeito do percurso diacrônico do pronome **se**, fato que, a meu ver, merece ser mencionado por sua relevância na literatura sobre
20 o assunto.

1.7 A análise diacrônica de Maurer Jr.

Uma das análises mais relevantes do percurso diacrônico do pronome **se** é o
25 trabalho de Maurer Jr, intitulado *A Propósito da Evolução Semântica do Pronome se em*

português. Nele o lingüista apresenta desde uma descrição histórica do pronome até uma discussão a respeito dos problemas teóricos envolvidos na sua classificação.

Segundo o autor, no período do latim vulgar o uso do pronome **se** era bastante restrito. Seria então somente a partir do período medieval que as inovações de uso do pronome surgiriam, no período do latim vulgar, derivadas da forma latina reflexiva **se**. O autor identifica três inovações do uso pronome **se** já nesse período, vejamos:

(4)

- a) **Estilístico**: Uso do pronome com verbos intransitivos para realçar a espontaneidade da ação, verificada em praticamente todas as línguas românicas. Vejam-se os exemplos dos verbos *ir-se* e *sair-se*, respectivamente: rum. *a se veni*, *a se sui*; it. *andarsi*, *fuggirsi*; fr. ant. *soi venir*, *soi fuir*; esp. *irse*, *venirse*.
- b) **Passivo**: Emprego do pronome **se** somente na terceira pessoa para a expressão da voz passiva, como na frase latina “*mela rotunda... todo anno servare se possunt*”. Essa forma passiva do pronome apresenta restrições nas línguas românicas, pois só se empregam preferencialmente quando o sujeito é inanimado, como: port. *Vendem-se as casas*, esp. *Se escriben las cartas*; it. *Dopo che l’adunanza si fu sciolta*. Entretanto, apenas o português e o espanhol apresentam a possibilidade de sujeito pessoal: port. *Louvam-se os grandes homens*.
- c) **Indeterminado**: Emprego do pronome **se** com verbos intransitivos para a expressão de um agente exterior indefinido. Esse tipo de construção indeterminado não é só comum ao português, mas também é recorrente no espanhol, italiano e romeno, como os exemplos a seguir: rum. *se sta bine aici*; it. *si sta bene qui*; esp. *se vive bien en esta ciudad*; port. *vive-se bem aqui*.

Mas qual seria a origem dessas construções? Para Maurer Jr., a resposta está no Indo-Europeu. Segundo o autor, como a *língua mãe* não tinha propriamente uma voz passiva, ou seja, só tinha voz ativa, ela desenvolveu por intermédio de um sistema de afixos a
5 expressão de noções passivas.

Os verbos então passaram a aceitar um complemento circunstancial que indicava a relação da ação expressa pelo sujeito com ele mesmo. Por exemplo, as formas do grego λουομαι “lavar-se”; διδασκομαι “faço que me ensinem”, exprimem, respectivamente, relações de reflexividade e passividade.

10 Assim, considerando os vários empregos da voz média, a “passividade” pode ser expressa por desinências verbais empregadas circunstancialmente, já que ela, a idéia de passividade, implica o fato de que o sujeito é afetado pela ação praticada por um agente exterior. Assim, enunciando o agente dessa ação, tem-se uma voz passiva perfeita, como na frase já citada: “*mela rotunda... todo anno servare se possunt*”.

15 Da passiva medial então se desenvolve “espontaneamente” o uso indeterminado do pronome *se*, que não deve ser confundido com o de sujeito. Para Maurer Jr. o pronome *se*, nas frases indeterminadas como *Vive-se bem aqui*, nada mais é que um morfema que *indetermina* a construção verbal. Segundo o lingüista, basta separar o pronome do verbo para se perceber que nada subsiste do seu uso indeterminado. O que acontece na evolução da pronome *se* é que
20 ele, junto ao verbo, desempenha uma função passiva, indeterminada, etc.

Segundo o autor, o desenvolvimento desse afixo indeterminado explica-se basicamente por um *esvaziamento* do sentido passivo, que gera as seguintes inovações: a) a construção passiva passa a admitir verbos intransitivos, como a forma *sai-se* no português; b)

mesmo na função *indeterminadora*, pode receber objeto direto elíptico, como em *Bebe-se muito aqui*.

1.8 Said Ali e a ausência da indeterminação

5 Said Ali em *Difficuldades da Língua Portuguesa* apresenta uma peculiar análise do pronome *se*. Essa peculiaridade se deve, basicamente, ao fato de que o autor nega a existência tanto do valor passivo do pronome, quanto do indeterminado. Veja-se o que é dito:

10 Que a princípio a forma reflexiva (ou média) coexistiu com a ativa, enquanto a passiva não se conhecia ainda. Ações praticadas por seres humanos não podiam ser enunciatas pela linguagem sem a indicação do agente. Quando, porém, o agente humano era desconhecido ou não convinha mencioná-lo, a linguagem servia-se deste expediente: personalizava o objeto se era um ente animado, e fingia-o praticar a ação sobre si mesmo. Certa mercadoria, por exemplo, devia ser vendida, ignorando-se o vendedor; dizia-se simplesmente: tal mercadoria vende a si própria. Pouco a pouco, porém, a mera forma reflexa em casos desse gênero começou a sugerir a idéia de um agente humano indeterminado.³

20 Considerando-se as palavras acima, conclui-se que o autor não está de acordo com a idéia da expansão dos usos gramaticais do *se*. Admitindo-se o fato de que em qualquer expressão de uma ação deveria haver também a especificação de um agente, então, quando tal ação tem um agente indefinido e exterior, o objeto é personalizado, dando a impressão de se praticar a ação sobre si mesmo, como em *Vende-se uma bola*, que poderia ser entendida também como “*A bola vende a si mesma*”, causando, segundo o autor, a impressão ilusória de passividade ou da presença de um agente exterior indefinido. Ou seja, na passagem acima, o autor ignora a interpretação indeterminada de orações como *Vende-se uma bola*, analisando-as como reflexivas.

30

³ *apud* (MAURER, 1951, p.50).

1.9 Naro e a reanálise do pronome *se*

Anthony J. Naro em: *The Genesis of the reflexive impersonal in portuguese: a study in syntactic change as a surface phenomenon*; e em sua tese doutoral intitulada *History of Portuguese and Impersonals* apresenta uma lúcida contribuição, dentro do quadro da perspectiva da gramática gerativa transformacional, a respeito da problemática diacrônica do pronome *se* na língua portuguesa.

Naro identifica três funções do pronome *se* na língua portuguesa. A primeira é caracterizada pelo seu emprego na terceira pessoa não-nominativa, sendo idêntica em referência a outro elemento da frase. Esse emprego é reconhecido pelo autor como reflexivo, como em: *O pai vestiu-se*.

Outro emprego do pronome *se* reconhecido por Naro é na formação da voz passiva, denominado de *se-passivo*. Consideram-se passivas as construções em que há concordância do verbo com o seu argumento interno. São casos típicos de *se-passivo* em construções da língua portuguesa: *Vendem-se estes móveis; Alugam-se casas*.

Segundo Naro (1976:779) outro tipo de emprego bastante comum na fala de indivíduos tanto no Brasil, quanto em Portugal, é o *se-impessoal* com verbos intransitivos e transitivos, embora as gramáticas normativas de ambos os países não considerem normalmente esse último tipo de construção correta. Vejamos os respectivos exemplos abaixo em (5a-b):

(5)

a) Vende-se estes carros.

b) Vive-se bem em Sabará.

Segundo Naro (1976), as construções com o *se-impessoal*, sem concordância, são, na língua portuguesa, historicamente posteriores às com *se-passivo*, e ganham aceitação mais geral no português a partir da primeira metade do século XV até se consolidarem em torno da metade do século XVI, como afirma (NARO, 1976, p.788): “There can be no doubt the
5 historical chronology: the *se-passive*, with agreement and agent phrase, precedes the *se-impersonal*, without agreement or agent phrase, by several centuries.”

Entretanto, para Naro (1968) o processo de surgimento do *se-impessoal* não é arbitrário, e não se constitui como uma mera derivação *errônea* ou subtipo do *se-passivo*, como afirmam alguns autores. Tal fato pode ser elucidado pela improdutividade das
10 construções com pronome *se* e a preposição *por*, como: * (a) *Escreve-se cartas pelas mãos*, * (b) *Observa-se mesas pelos olhos*. Ou seja, essas expressões deveriam ser produtivas para justificar a extensão da regra de uso do pronome *se* em frases que apresentam sujeito.

A improdutividade das construções de (a) e (b) indica também, segundo o autor, que a produtividade das construções com *se-impessoal* não se formou por um processo de
15 extensão analógica das estruturas *se-passivas*, já que as indeterminadas não admitem especificação de agente e as passivas sim.

Naro afirma então que a reanálise que gerou as construções com *se-impessoal* só foi possível por meio da ausência de sujeito em frases nas quais o verbo no plural não concorda com o argumento interno. O autor salientou, por fim, que o apagamento (*deletion*)
20 do sujeito só é possível em frases que esse é indefinido ⁴.

⁴ Notice also that this blocking means that only *se-passives with deleted deep subject are eligible for re-analysis as actives*. According to the universal deletion conditions, the deletion could occur only if the deep subject were an indefinite since there is no other NP in the sentence to which the deep subject is in general identical. (NARO, 1968, p.145)

1.10 Nunes e o percurso do pronome **se**

Jairo Nunes, em seu trabalho intitulado *O Famigerado SE: uma análise diacrônica das construções com SE apassivador e indeterminador*, apresenta, dentro do quadro gerativista, uma relevante contribuição, do ponto de vista teórico e metodológico, para a compreensão do percurso diacrônico do pronome **se**.

Considerando a problemática do desenvolvimento histórico do pronome **se**, e partindo da análise apresentada no trabalho de Naro (1976), Nunes (1990) propõe descrever o percurso histórico do pronome **se** na língua portuguesa, considerando dois aspectos fundamentais: a) a posição do argumento interno, e; b) a concordância do argumento interno com o verbo. Outro objetivo de Nunes é, a partir da descrição de Naro para a reanálise *se-passivo* > *se-indeterminado*, apontar a expansão do *se-indeterminado*, usualmente encontrado em construções com verbos intransitivos, para também com verbos transitivos.

Nunes afirma que a expansão *se-passivo* > *se-indeterminado* é caracterizada por uma reanálise da categoria vazia na posição de sujeito das passivas pronominais, de modo que em construções como *Aluga-se casas*, o pronome **se** passa a ser interpretado como agentivo, e não como uma partícula apassivadora, como no caso das sentenças passivas.

Essa reanálise forçou a interpretação do SN posposto como objeto direto, portador do caso acusativo e não mais como sujeito, caso nominativo e responsável pela concordância.

De acordo com as análises de Nunes (1990), e também o trabalho de Nascimento (1984)⁵, no caso do *se-apassivador*, o pronome **se** absorve o papel temático do argumento externo e o caso acusativo; já em relação ao *se-indeterminador*, o pronome tem somente a função de indeterminar a categoria vazia na posição sujeito, e é esta última que recebe o papel temático e o caso nominativo reservados ao sujeito.

⁵ O autor também estudou o pronome **se** dentro de uma perspectiva gerativa. Esse trabalho não será comentado aqui.

Desse modo, o processo de reanálise do *se-passivo* que inicialmente ocorreu com verbos intransitivos, expande-se, segundo Nunes (1990), passando a ocorrer também com verbos transitivos. Essa expansão gramatical, segundo o autor, seguiu a seguinte ordem:

5 (6)

a) verbos transitivos diretos usados intransitivamente:

Come-se muito no inverno.

b) verbos intransitivos:

Trabalha-se muito neste lugar.

10

c) verbos transitivos preposicionados:

Precisa-se de empregados.

d) verbos de ligação:

É-se feliz quando é jovem.

e) verbos ergativos:

15

Chegou-se tarde a reunião.

f) verbos em construções passivas perifrásticas:

Aqui se é visto por todos.

Outro fenômeno recente e de grande relevância na indeterminação do sujeito na língua portuguesa observado por Nunes (1990) é o apagamento do pronome *se* em
20 construções do tipo: *Nos nossos dias não usa mais saia*⁶, muito comum no português europeu com sentido determinado, se referindo a *ela* ou a *ele*, mas que no português brasileiro podem ser encontradas com sentido indeterminado.

25

⁶ Cf. (NUNES, 1990, p.99),

1.11 Considerações finais

Mesmo diante da contribuição dos estudos supracitados a respeito do percurso histórico – reanálise - do pronome *se* na língua portuguesa, ainda persistem questões que necessitam ser mais bem esclarecidas, e que, por sua vez, se constituem nos objetivos desse trabalho.

Não está claro quando se tornou mais produtivo o processo de reanálise do pronome *se* na língua portuguesa, e como esse processo se desenvolveu. Ou seja, quando as construções com *se-indeterminado* passaram a ser consideradas produtivas? Qual a frequência de estruturas com *se-indeterminado* em períodos mais antigos dos analisados por Naro (1966)? Quão significativa deve ser a frequência delas para se atestar sua produtividade? Naro (1966) afirma não ter existido *se-indeterminado* no período medieval do português, fato que, a meu ver, merece ser mais bem investigado dentro de uma análise de frequência, e por meio de outros corpora.

Outro fator ainda inexplicado é como ocorreu a reanálise *se-reflexivo* > *se-passivo* > *se-indeterminado*, ou seja, como o pronome reflexivo *se* do português pôde dar origem a formas do grupo não-reflexivo, fato constatado por Rodrigues:

Elas [as construções com pronome *se*] não são, por assim dizer, um produto primário da língua, mas resultam de um processo mental um pouco longo. Foi preciso que o pronome reflexo *se* passasse a partícula apassivativa, e que depois a oração assim tornada passiva se considerasse transposta para a activa, ficando o *se* a servir de sujeito indeterminado. (*apud* NARO, 1968, p. 148)

Os estudos de Naro e Nunes não apontam uma solução para esses fatos, ou seja, como e quando as construções reflexivas com o pronome *se* foram reanalisadas em construções com *se-passivo* e *se-indeterminado*, e porque esse fenômeno foi possível

considerando aspectos sintáticos e semânticos das estruturas com *se-reflexivo*, *se-passivo* e *se-indeterminado*.

Outra questão relacionada à reanálise do pronome **se** é como ocorreu, em termos comparativos de frequência, a expansão de seus subtipos no sistema da língua portuguesa. Ou seja, quão significativo é o aumento da frequência dos subtipos do pronome em usos mais gramaticais, como, por exemplo, o *se-indeterminador*, nos vários períodos de nossa língua, o que será discutido no capítulo V.

Os trabalhos realizados sobre a reanálise do pronome **se** até o momento não descrevem, ou melhor, não objetivam analisar, em termos de comparação de frequência, sua expansão gramatical. Necessário se faz, a meu ver, para elucidar tais processos de reanálise, compreender, comparando mais precisamente em termos de aumento/diminuição de frequência, a expansão dessas diversas estruturas, ou subtipos, do pronome **se** ao longo da história da língua portuguesa.

Deve ser considerado também nesse processo de reanálise do pronome **se** o papel que diferentes gêneros textuais desempenham na frequência de uma determinada estrutura, fator esse sequer mencionado nas análises apresentadas. A meu ver, ambientes sintáticos e semânticos distintos podem estar associados à maior ou menor frequência de determinadas estruturas com o pronome **se**.

Um fato relevante ainda a ser destacado é que, para se tratar a reanálise do pronome **se**, considerando-se aspectos de sua frequência, faz-se necessário o desenvolvimento de uma metodologia apropriada, que não está ainda disponível aos pesquisadores. Ou seja, justifica-se assim não só a análise de frequência, que consistiria uma inovação para os processos de reanálise do pronome **se**, mas, também, o desenvolvimento de um instrumental teórico para se tratar tal fenômeno de modo apropriado, dentro dos princípios da gramaticalização.

Outro aspecto que merece também uma melhor elucidação é a coexistência de estruturas de uso mais antigas e mais recentes do pronome **se**, nesse caso, o *se-estilístico*, *se-reflexivo* e o *se-passivo* apresentam-se como formas mais antigas, e o *se-indeterminado*, com verbos transitivos, mais recente.

5 Ou seja, mesmo sendo verificada a reanálise do pronome **se**, as estruturas mais antigas, nesse caso, a Estilística, reflexiva e a passiva, são mantidas, convivendo “harmonicamente” com as ocorrências do *se-indeterminado* com verbos intransitivos e transitivos. Tal fato pode vir a ser um problema, principalmente, para o quadro variacionista, que prevê a queda de frequência das formas mais antigas em favor das inovadoras quando
10 elas têm o mesmo valor de verdade. Tais questões serão discutidas no capítulo X.

Desse modo, acredito que considerar os processos de reanálise inerentes ao percurso histórico do pronome **se** à luz dos princípios da gramaticalização poderá não só enriquecer a análise desse fenômeno, elucidando muitos dos problemas suscitados aqui pelo objeto de estudo, como também contribuir para o desenvolvimento de novos instrumentos
15 teóricos que poderão ser utilizados e estendidos a vários outros fenômenos que podem ser vistos como processos de gramaticalização.

CAPÍTULO II: SUPORTE TEÓRICO

2.1. Apresentação

5

Após ter apresentado no capítulo anterior dados a respeito da classificação, definição, e da trajetória diacrônica do pronome *se*, e ter aventado a hipótese de que essa trajetória seja tratada como um processo de gramaticalização, apresento neste capítulo o suporte teórico adequado para se tratar esses fenômenos.

10

Desse modo, primeiramente, como introdução ao suporte teórico, apresento um histórico da teoria da gramaticalização e seus pressupostos gerais, e, em seguida, apresento as noções de princípios, estágios e mecanismos dos processos de gramaticalização.

2.2 Origens da teoria e definições de Gramaticalização

15

Uma questão ainda muito controversa é a que diz respeito à explicação das causas e do sentido da mudança lingüística. Partimos então do início: *O que leva línguas a mudar? Quais fatores específicos estariam implicadas nesse processo?*

20

Basta um olhar desprezioso para a história de uma língua que poderá ser percebido o quão claro é o fato de que ela está em movimento, que suas unidades estão se transformando, porque o sistema de uma dada língua é, antes de tudo, um sistema que expressa todos os tipos de aspectos culturais de uma sociedade, aspectos que, por sua vez, estão em constante desenvolvimento. E numa sociedade que se movimenta, sua língua não deixa de refletir esse movimento também.

25

Do ponto de vista especificamente lingüístico, várias podem ser as explicações para a mudança. A mais comum tem um caráter evolucionista e não possui, a meu ver, qualquer fundamentação lingüística, se a tomarmos no sentido de que as línguas, assim como os homens, evoluem, mudam para melhor, tornando-se mais comunicativas, mais eficientes, mais reduzidas no que diz respeito aos seus componentes fonético, morfológico, e sintático, tornando-se, desse modo, mais simples, mais objetivas, ou mais claras.

Diante da questão da mudança lingüística, os lingüistas, no intuito de explicar a evolução das formas gramaticais, em geral, apontam três mecanismos fundamentais: a) analogia⁷; b) gramaticalização, e c) empréstimo. A analogia consiste na extensão de uma regra regular, como, por exemplo, ocorre com o empréstimo do inglês, o verbo *to delet*, que passa a ser em português **deletar*, e segue o paradigma flexional do presente do indicativo dos verbos da primeira conjugação, nesse caso, os verbos terminados em -ar, *eu deletei*, *ele deletou*, *nós deletamos*. Pelo fato de a primeira conjugação em português conter mais verbos, e seu paradigma ser mais regular, e freqüente, torna-se um modelo para a criação de novas formas gramaticais, assim como **scannear* < scan “copiar uma imagem”, **backupear* < backup “fazer uma cópia de segurança, um backup”, etc. Esse processo que cria novas formas gramaticais por meio de modelos regulares é conhecido pelos lingüistas como *analogia*. A investigação a respeito desses processos teve sua mais alta expressão nos trabalhos dos lingüistas neogramáticos, ao final do século XIX, a partir da proposição de leis fonéticas universais deduzidas pela comparações de línguas diferentes de uma mesma família.

Entretanto, o lingüista francês Antoine Meillet, procurando compreender os processos de mudança das línguas, e considerando fundamentalmente uma visão sociológica da linguagem, propõe o conceito de gramaticalização contrapondo-o ao da analogia para a explicação da criação das novas formas lingüísticas:

5

Tandis que l'analogie peut renouveler le détail des formes, mais laisse le plus souvent intact le plan d'ensemble du système existant, la 'grammaticalisation' de certains mots crée des formes neuves, introduit des catégories qui n'avaient pas d'expression linguistique, transforme l'ensemble du système.⁸

De uma forma muito diferente da analogia, os processos de gramaticalização também são responsáveis pela criação de novas formas gramaticais. Segundo Meillet, a gramaticalização pode ser definida como um processo pelo qual uma palavra autônoma passa a desempenhar uma função gramatical. Na frase do ing. “*I have a car*”, “Eu tenho um carro”, o verbo “have”, apresenta um sentido pleno, com sentido de “possuir”. Entretanto, na frase “*I have told you*”, “Eu disse a você”, o verbo “have” não apresenta o sentido lexical de “posse”, nesse caso, é um verbo auxiliar, que é responsável pela marcação do passado. Desse modo, como o verbo “have” em “*I have told you*” não apresenta sentido lexical, mas, sim, gramatical, em sua função de auxiliar, pode-se dizer que esse uso de “have” indica um processo de gramaticalização.

Ainda por oposição à analogia, os processos de gramaticalização podem alterar a percepção ou compreensão pelo falante da estruturação do sistema existente, ou seja, dos seus elementos e de suas respectivas categorias lexicais e funcionais que não são fixas no sistema lingüístico.

Ao processo que implica alteração da percepção da estrutura sintática da oração denomina-se *reanálise*. A título de exemplificação, consideremos o caso do período composto por duas orações [[devo] [viajar]], que é reanalisado na estrutura [[devo[viajar]], na qual *devo* passa a funcionar como auxiliar, diferentemente da primeira em que *devo* constitui-se a oração principal.

⁸ (MEILLET, 1948, p.8). “*Ao mesmo tempo que a analogia pode renovar o detalhe das formas, mas deixa por muitas vezes intacto o plano do conjunto do sistema existente, a 'gramaticalização' de certas palavras cria formas novas, introduz categorias que não tinham expressão lingüística, transforma o conjunto do sistema.*” (Tradução minha)

Meillet aponta dois aspectos fundamentais motivadores para a gramaticalização de um determinado elemento: a) diminuição do seu valor expressivo, e b) frequência. Segundo o lingüista francês, quanto mais uma determinada estrutura, ou palavra, é usada, mais ela perde em seu valor expressivo, tornando-se banal, ou um clichê. E segue: “Si donc un groupement de mots devient fréquent, s’il est souvent répété, il cesse d’être expressif, et il est reproduit de plus automaquement par les parlants (MEILLET, 1948, p.12).” Meillet cita o exemplo do verbo francês *laisser*, que devido à sua alta frequência com verbos no infinitivo como em fr. *je laisse venir*, perde o seu sentido próprio, e passa a desempenhar uma função auxiliar na frase.

A partir da contribuição do trabalho de Meillet, outros lingüistas como Vossler, Vendryes, Benveniste, Kurylowicz, entre outros, deram prosseguimento às investigações a respeito da gramaticalização, numa linha mais estruturalista.

Considerando as palavras de Kurylowics, admite-se que a gramaticalização apresenta uma gradação, ou seja, não necessariamente só pode se processar do léxico para a gramática, mas também de um nível menos gramatical para um mais gramatical. Lehmann (1982), por outro lado, apresenta alguns parâmetros que definem o que ele compreende como aspectos da gramaticalização fraca ou forte.

Joseph Vendryes (1968), em sua obra intitulada *Le langage*, embora não utilize o termo *gramaticalização*, faz uma menção explícita ao trabalho de Meillet sobre o assunto. O autor aborda questões semelhantes sob a denominação de *transformações morfológicas*. O lingüista considera que o sistema morfológico das línguas é instável, e enumera duas tendências gerais relativas a essas transformações. Vejamos quais são elas abaixo:

(7)

a) **analogia**: criação de uma forma por meio de um modelo que lhe é conhecido.

b) **transformação de palavras plenas em palavras vazias**: com o decorrer do tempo, as palavras plenas, com valor referencial, passam a desempenhar funções gramaticais, tornando-se vazias de significado.

5 Segundo Vendryes, muitos são os exemplos de esvaziamento de sentido, como por exemplo, no grego, as formas *μετα* (com), *μεσφα* (até) que provém de uma raiz que significa *metade*. Segundo o autor, o equivalente ao nome *homem* tornou-se também um instrumento gramatical no francês *on*, em *On dit*, (Diz-se) e no alemão *man*, em *Man sagt*, (Diz-se).

Karl Vossler (1968), em sua obra *Filosofia da Linguagem*, admite o conceito de gramaticalização introduzido por Meillet, entretanto, inserindo-o de um modo diferente no contexto dos processos gramaticais. Vossler acredita que a formação do sistema gramatical vai além dos processos de analogia e gramaticalização. Para o lingüista alemão, os mecanismos que interferem na constituição dos sistemas lingüísticos são seis, três uniformizadores e três diferenciadores. Os uniformizadores são: analogia, gramaticalização, e 10 a contaminação ou empréstimo; e os diferenciadores são: a mudança fonética, a mudança semântica e a diferenciação. Segundo Vossler, o fundamento desses seis processos gramaticais está na falta de atenção do falante ao usar a língua em diferentes contextos. 15

Analisando mais especificamente a gramaticalização, Vossler a considera como um processo de esvaziamento semântico, ou seja, palavras livres tornam-se, com seu uso, 20 morfemas presos a outras. Segundo o autor, quase todos os elementos que desempenham funções gramaticais provêm de palavras independentes, arcaizadas e fossilizadas pela excessiva freqüência do seu uso num sentido específico. O lingüista apresenta exemplos da língua alemã como *-heit*, *-tum*, *-bar*, etc. que atualmente são sufixos, como em *Freiheit*, (liberdade), mas que em períodos pretéritos dessa língua eram palavras independentes, ou 25 seja, com sentido pleno.

De acordo com Vossler, as palavras gramaticais são como vestígios, formas residuais de elementos lingüísticos lexicais desaparecidos em um sistema. Vossler, entretanto, tem uma visão bem mais ampla das causas desse processo de esvaziamento que a de Meillet, que vê na frequência o seu mais forte mecanismo. Para o lingüista germânico, esse esvaziamento de sentido está sujeito também a influências da lógica, estética, ética, bem como aspectos da geografia lingüística e da sociologia, já que o autor compreende a língua como uma atividade espiritual e uma expressão da individualidade humana.

O filósofo alemão Wilhelm Von Humboldt, num trabalho intitulado, *Über das Entstehen der grammatischen Formen und ihren Einfluss auf die Ideenentwicklung*, (Sobre a formação das formas gramaticais e sua influência no desenvolvimento das idéias) 1825⁹, defende a tese de que classes de palavras como preposições e conjunções têm sua origem em palavras reais, propondo quatro estágios de evolução que descrevem o processo de mudança dos itens lexicais até a formação de itens com funções gramaticais:

15

QUATRO ESTÁGIOS DA EVOLUÇÃO LINGÜÍSTICA HUMBOLDT (1825)

20

Estágio I: (o mais baixo estágio) idiomas, sentenças, orações.

Estágio II: ordem de palavras fixa oscilando entre assunto e forma do significado.

Estágio III: formas semelhantes, que são a expressão de relações.

Estágio IV: (o mais alto estágio) inflexão, palavras com valor especificamente gramatical

Este modelo, que foi conhecido como “teoria da aglutinação”, está estritamente relacionado com aquele desenvolvido por Schelegel, pois os estágios I e II correspondem à etapa lexical da língua, o estágio III é o de aglutinação, e o estágio IV, consiste no inflexional e gramatical. Essa talvez tenha sido a primeira noção a surgir no âmbito dos estudos lingüísticos do que se compreende hoje como gramaticalização.

⁹ *apud* (HEINE, HÜNNEMEYER E CLAUDI, 1991, p.6).

Jean Dubois (1973:85), no *Dictionnaire de Linguistique*, compreende a gramaticalização como a transformação de um morfema lexical em um gramatical ao curso de desenvolvimento de uma língua: “on parle de grammaticalisation, quand un morphème lexical, au cours de l’évolution d’une langue, ou dans la transformation d’une langue en autre, 5 est devenu un morphème grammatical”¹⁰.

J. Mattoso Câmara Jr. no *Dicionário de Filologia e Gramática*, define gramaticalização como a transformação de vocábulos lexicais ou palavras providas de semantema, em vocábulos de natureza gramatical. O autor afirma adiante que os vocábulos gramaticais, em sua totalidade, advêm do processo de gramaticalização, ou seja, formam-se a 10 partir de vocábulos lexicais.

Outra relevante contribuição para a definição do conceito de gramaticalização, e que apresenta especial relevância para a problemática do trabalho em questão, é a de Kurylowicz, que a define como: *a ampliação dos limites de um morfema, cujo estatuto gramatical avança do léxico para a gramática, ou de um nível menos gramatical para mais 15 gramatical.*¹¹

Considerando a gramaticalização como apenas um processo de mudança lingüística, faz-se necessário um modelo para subsidiar sua análise. Os autores da literatura, comumente, não a descrevem a partir de um modelo próprio, utilizando, de um modo geral, noções oriundas de quadros teóricos distintos, e buscando estabelecer princípios e 20 mecanismos envolvidos nos processos de gramaticalização.

Nessa discussão, entretanto, não há consenso, pois tais princípios e mecanismos são sempre propostos considerando-se os respectivos fundamentos das teorias dos quais são formulados. Desse modo, o intuito aqui neste presente trabalho é apresentar algumas

¹⁰ “Fala-se de gramaticalização quando um morfema lexical, ao longo de sua evolução em uma língua, ou na transformação de uma língua em outra, se torna um morfema gramatical” [tradução minha]

¹¹(CASTILHO, 1997, p. 29)

postulações gerais, que, a meu ver, são compartilhadas pela maioria dos teóricos da gramaticalização.

Lehmann (1982), buscando fundamentar uma teoria da gramaticalização, identificou cinco princípios fundamentais nos processos de gramaticalização, considerando,
5 mais especificamente, aspectos formais de análise:

(8)

a) **Paradigmatização**: de acordo com esse princípio, construções sintáticas unem-se como formas perifrásticas, conduzindo à formação de paradigmas progressivamente menores e homogêneos.

10

b) **Obrigatoriedade**: dentro de um determinado paradigma, a escolha entre seus membros está sujeita a regras gramaticais. O conjunto categorial torna-se gradativamente mais obrigatório nas sentenças da língua.

15

c) **Condensação**: quanto mais gramaticalizado um elemento, mais complexos são os constituintes com os quais ele pode combinar-se.

d) **Coalescência**: este princípio se manifesta por meio de processos de justaposição, por meio de cliticização, aglutinação e fusão.

20

e) **Fixação**: itens mais gramaticalizados tendem a ocupar uma posição fixa, inicialmente, na sintaxe, e, posteriormente, na morfologia.

Considerando o trabalho de Lehmann (1982), um item mais gramatical apresenta
25 paradigmas cada vez menores e homogêneos e está mais fixamente posicionado dentro da estrutura da oração, podendo combinar-se com um grande número de estruturas. Em termos

morfológicos, processos de cliticização, aglutinação, e fusão podem ser também considerados índices de gramaticalização.

A partir de Lehmann (1982), pode se afirmar que quanto mais gramatical é um elemento mais ele: a) perde traços semânticos; b) reduz o seu paradigma; c) é empregado
5 obrigatoriamente; d) se posiciona fixamente nos sintagmas ou palavras.

Outras contribuições significativas para a compreensão dos processos de gramaticalização podem ser enumeradas. Dentre elas, as fortemente influenciadas pela perspectiva funcionalista de Givón. Trabalhos como os de Paul Hopper, Heine, Traugott e outros, procuraram, dentro do quadro funcionalista, compreender mais precisamente os
10 princípios dos processos de gramaticalização, bem como propor estágios e mecanismos pelos quais a gramaticalização se desenvolveria. Princípios, estágios e mecanismos que serão mais bem descritos nas seções seguintes.

Considerados então explicitados os princípios que originam os processos de gramaticalização, passo a analisar *mais especificamente* as características dos processos de
15 gramaticalização. Destacaremos a continuidade, o gradualismo e a unidirecionalidade (irreversibilidade). Passemos a compreendê-los melhor nas seções seguintes.

2.2.1 Continuidade e Gradualismo

20 Considerando que as línguas tendem continuamente a se renovar, vários mecanismos de mudança, dentre eles, a gramaticalização, desempenham um papel fundamental nesse processo.

Nessa renovação, a variação constitui-se em um aspecto fundamental. Autores como Hopper indicam que *a variação é uma consequência natural e necessária do caráter*
25 *gradual e contínuo das mudanças lingüísticas*. Assim, na perspectiva da gramaticalização, a mudança lingüística não deve ser vista como abrupta, como propõe Lightfoot (1992).

Sendo então inextinguível do sistema das línguas a variação e a mudança, incluindo nesse caso mais especificamente os processos de gramaticalização, torna-se difícil identificar e compreender claramente, como afirmam Heine et alii (1991), suas fases ou
5 estágios.

A esse caráter constante da continuidade e gradualidade da mudança está associada, em grande parte, a criatividade dos falantes e ouvintes ao inferirem sobre o significado das estruturas lingüísticas, bem como as motivações sociais de toda sorte que podem influenciar as mudanças lingüísticas. Desse modo, enquanto houver mudança social,
10 ou seja, uma sociedade vivaz, essa vivacidade *social* se refletirá no sistema lingüístico.

Entretanto, o processo de contínua renovação das línguas apresenta dois aspectos contraditórios: a) por um lado, os falantes podem escolher entre diferentes formas lingüísticas, o que pode levar à extinção de umas e à consolidação de outras; b) por outro, as inovações podem trazer ao sistema um conjunto de novas restrições, tornando arcaicas e
15 incompreensíveis formas antes produtivas, e fazendo das inovações, elementos consolidados no sistema.

2.2.2 Unidirecionalidade

20 Segundo Hopper & Traugott (1993), os processos de gramaticalização compreendem quatro estágios, como indicado abaixo:

(9)

1 Item lexical > 2 item gramatical > 3 clítico > 4 afixo
--

Considerando então os *estágios* da gramaticalização, os autores afirmam que ela só se realiza da esquerda da direita (considerando o quadro acima). Ou seja, um determinado item não poderia se desgramaticalizar, mas só poderia movimentar-se de um estágio inferior para um estágio superior, e não o contrário.

5 Entretanto, embora sejam aventados alguns casos que poderiam ser considerados como exemplos de desgramaticalização, esses não estão ainda devidamente explicitados pelos teóricos. Alguns dos discutíveis casos de desgramaticalização são, por exemplo, a nominalização de alguns sufixos, como o suf. *ismo*, em *Os ismos da ciência*.

10 2.3 Perspectiva Formal de Análise dos Processos de Gramaticalização

Vitral (2005), preocupando-se em investigar, primeiramente, qual concepção de gramática é subjacente à perspectiva da gramaticalização, e, em segundo lugar, que noção de mudança interessa a essa perspectiva, propõe uma perspectiva mais formalista de se
15 compreender tais fenômenos.

Desse modo, para “responder” ao primeiro ponto acima, ou seja, qual concepção de gramática seria subjacente à perspectiva da gramaticalização, é elaborada uma proposta num quadro mais formalista, dentro das propostas da gramática gerativa¹². Nessa visão, os processos de gramaticalização envolvem a recategorização de itens de uma determinada
20 categoria lexical para uma categoria gramatical, ou funcional (nos termos gerativistas). De acordo com essa perspectiva, as categorias lexicais e gramaticais são assim caracterizadas:

(10)

Categorias Lexicais: Nome (N), Verbo (V), Preposição (P), Adjetivo (A) e Advérbio (ADV).

Categorias Gramaticais: Determinante (D), Negação (NEG), Flexão (F), Auxiliar (AUX) e
25 Complementizador (C).

¹² cf. VITRAL; L, RAMOS, J. (2002).

Nos termos dessa proposta, um item apresentará gramaticalização se for demonstrada a passagem dele de uma categoria lexical para uma gramatical. Um exemplo elucidativo desse fenômeno é o caso do verbo *ter* (V), que passou a funcionar como (AUX). Nesse caso, a categoria AUX, segundo Vitral (2004), “ganha” um novo item para sua
5 expressão.

Em relação ao segundo ponto, ou seja, qual concepção de mudança interessa à perspectiva da gramaticalização, o autor afirma que, pelo fato de haver co-ocorrência das categorizações lexicais e gramaticais sem o mesmo “valor” de verdade, tal fato não pode ser compreendido como o mesmo tipo de fenômeno de interesse da teoria da variação e da
10 mudança (variacionismo). Desse modo, o autor, observando as especificações que a perspectiva da gramaticalização implica, propõe referir-se às re-categorizações lexicais e gramaticais por *Inovações Lingüísticas*, e não mais mudanças.

Segundo o autor, nos fenômenos de gramaticalização, a forma fonte é preservada, mas isso só ocorre nas primeiras duas fases, considerando o ciclo de Heine e Traugott. A
15 partir do terceiro estágio, podem já ocorrer, entretanto, formas fonéticas diferentes com o mesmo valor de verdade, como é o caso do pronome “você”, e de seu clítico equivalente “cê” (cf. Vitral (2003)), quando usados como segunda pessoa do singular. O autor conclui que a concorrência de formas constitui-se de um subcaso ou estágio do processo de gramaticalização.

20 Vitral (2004) afirma também que, considerando os objetivos do seu trabalho, que é estabelecer linhas gerais de uma metodologia específica para os processos de gramaticalização, a análise da frequência do item exerce um papel fundamental nessa tarefa. O autor considera que só há ganhos no diálogo entre as perspectivas da Gramaticalização, do Modelo Gerativo e do Variacionismo. Diálogo que, num futuro próximo, pode vir a dar forma
25 a um programa específico de pesquisa distinto reunindo resultados desses três campos de pesquisa.

Considerados os princípios e mecanismos acima, passo ao próximo capítulo que trata mais especificamente da metodologia de identificação de processos de gramaticalização.

CAPITULO III: METODOLOGIA

3.1 Apresentação

5 Após ter apresentado no capítulo precedente um histórico e alguns conceitos centrais da teoria da gramaticalização, passo a expor mais especificamente a proposta metodológica que, como dito, está, a meu ver, de acordo com os fenômenos e problemas teóricos apontados nas seções precedentes.

10 Desse modo, tomo como *referencial metodológico* para meu trabalho o de Vianna (2000) e o de Vitral (2005), que têm como ponto fundamental a proposição de uma metodologia quantitativa capaz de identificar, de forma objetiva, os processos de gramaticalização.

15 O desenvolvimento da metodologia proposta por Vitral constitui-se de um desdobramento de outro trabalho também realizado com propósitos semelhantes, o de Vianna (2000), esse que analisou o percurso diacrônico dos verbos modais *poder, dever, e querer*. Para demonstrar sua metodologia, nesse trabalho Vitral analisa o percurso diacrônico da forma verbal *ter*.

20 Considerados então os objetivos estabelecidos pelo autor, são apresentadas algumas questões teóricas e critérios de análise que envolvem os processos de gramaticalização.

3.2 Problemas teóricos

25 Vitral (2005) apresenta três razões principais que dificultam a caracterização de processos de gramaticalização, e que surgem como motivações para proposição de uma metodologia mais consistente do ponto de vista teórico: a) nem todas as etapas do processo são visíveis pelos procedimentos atuais utilizados na identificação desses processos; b) nem

sempre se sabe quais aspectos sintáticos devem ser considerados na caracterização deles; c) um item pode ocorrer simultaneamente desempenhando uma função lexical e gramatical numa mesma sincronia.

Vitral apresenta também outras questões teóricas que apontam a necessidade do desenvolvimento de uma metodologia própria para a identificação dos processos de gramaticalização: “é possível elaborar uma metodologia, composta de recursos quantitativos, que nos permita a identificação e descrição dos processos de gramaticalização?” (Vitral 2004: 2).

Em suma, o interesse do autor é propor uma metodologia quantitativa capaz de “diagnosticar” processos de gramaticalização. Para tal, propõe-se um estudo da trajetória da forma *ter*, já reconhecida por alguns autores como Cohen (1988), Mattos e Silva (1989), e Ribeiro (1993) como um fenômeno de gramaticalização e, desse modo, capaz de “testar” a metodologia proposta.

Após justificar-se quanto ao desenvolvimento de sua metodologia, Vitral apresenta os critérios que, a seu ver, podem identificar o grau de gramaticalização de determinados elementos e estruturas. Passemos, então, à seção seguinte.

3.3 Critérios de Identificação dos processos Gramaticalização

A gramaticalização, como dito acima, só pode ser atestada, se, após serem definidos determinados critérios, um dado item encontrar-se mais ou menos gramaticalizado num determinado período. Entretanto, para se verificar os processos de gramaticalização é preciso, primeiramente, recorrer-se a uma teoria de base que dê subsídios para a caracterização do processo.

Em seu trabalho, Vitral (2004) apresenta os critérios propostos para a identificação dos processos de gramaticalização. Esses critérios podem ser subdivididos em dois grandes grupos, a saber:

(11)

- 5 1º. Grupo: critérios sintáticos, critérios semânticos e critérios morfofonéticos;
 2º. Grupo: critérios de frequência.

Em relação aos critérios sintáticos, pertencentes ao primeiro grupo, o autor propõe que, para definir se um item se comporta como um de uma categoria lexical ou de uma
10 categoria gramatical, deve-se considerar, primeiramente, uma teoria sintática que ofereça subsídios de análise. Nesse caso o autor, então, baseia-se no trabalho de Chomsky (1995), como dito, dentro do quadro gerativo. No caso do verbo *ter*, o simples fato de ele poder ser seguido de um SV com participio “Eu tenho estado”, seria suficiente para caracterizá-lo nessa ocorrência como um indicador de gramaticalização e lexical com um SN.

15 No que diz respeito à morfofonética, compreende-se que as formas mais gramaticalizadas apresentam redução fônica, critério que é compartilhado pelos teóricos da literatura como sendo um indicador de processo de gramaticalização. Assim como na sintaxe, a aplicação desse critério pressupõe também uma teoria de base para sua sustentação.

20 No que diz respeito à semântica, compreende-se que as formas mais gramaticalizadas apresentam ampliação do número de sentidos de um determinado item e perda de complexidade semântica. Outro aspecto a ser ressaltado é que o item gramatical passa a apresentar um significado mais abstrato, em oposição ao lexical, mais concreto.

O segundo grupo de critérios para a caracterização dos processos de gramaticalização é o da frequência, como foi descrito acima¹³. Segundo o autor, a motivação para a proposição desses instrumentos de quantificação, ou seja, da apuração da frequência, é porque os então disponíveis ainda não são capazes de tornar objetiva a identificação dos processos de gramaticalização.

De um modo geral, os cálculos da frequência funcionam como um mecanismo capaz de comparar o comportamento de determinados elementos em funções gramaticais e lexicais nos períodos delimitados para a análise de acordo com os critérios determinados. No caso do trabalho de Vitral, como dissemos, foi procedida a análise de frequência do verbo *ter* na função lexical e gramatical, nos três períodos analisados. Caso a frequência da forma *ter* em função gramatical seja maior do que aquela em sentido lexical ao longo dos períodos analisados, poderá ser aventada a hipótese da existência de um processo de gramaticalização no percurso histórico dessa forma. Vejamos abaixo mais detalhes técnicos de procedimentos e etapas de análise propostos pelo autor:

- 15 1. análise da frequência geral do item I – somando-se, portanto, sua ocorrência nas funções gramatical e lexical em relação ao número total de palavras de cada texto que compõe o corpus de cada período, comparando posteriormente os valores encontrados nos períodos considerados.
- 20 2. análise da frequência do item I em função gramatical em relação ao total de ocorrências do item I - isto é, somados os números de ocorrência nos sentidos gramatical e, comparando posteriormente os valores encontrados nos períodos considerados.
- 25 3. análise da frequência do item I em função lexical em relação ao total de ocorrências do item I - isto é, somados os números de ocorrência em função gramatical e, comparando posteriormente os valores encontrados nos períodos considerados.
- 30 4. análise da produtividade do item I em função gramatical, isto é, verificação da frequência de I em função gramatical em relação à frequência de um item gramatical que, sabidamente desempenha a mesma função gramatical.

¹³ Como vimos, vários autores, dentre eles Antoine Meillet, já apontaram a estreita relação entre a frequência e os processos de gramaticalização, de modo que, quanto maior a frequência de uma determinada estrutura mais gramaticalizada ela é.

Vitral propõe, de acordo com seus procedimentos descritos acima, analisar a frequência geral de um item, e, posteriormente, analisá-la comparativamente em suas funções lexicais e gramaticais. Um dos indicadores da gramaticalização seria se se encontrasse, comparados os períodos, maior frequência da forma numa função gramatical do que numa
5 função lexical. Se essa frequência aumentasse gradativamente considerando os períodos analisados, a hipótese de gramaticalização seria, a modo de ver do autor, confirmada. Entretanto, deve-se ressaltar que a análise da frequência está diretamente ligada aos critérios estabelecidos por um modelo teórico que aponte o que são, e o que não são do ponto de vista teórico, as funções lexical e gramatical.

10

3.4 Tratamento do corpus

Vitral (2004) apresenta três diretrizes gerais para a composição do *corpus*, que são
15 elaboradas de acordo com as especificidades dos fenômenos os quais o autor propõe-se a analisar. São elas: a) diversidade de gênero, justificada pela necessidade de investigação de um item em ambientes semânticos diferentes; b) os textos que compõem o *corpus* devem ser aproximadamente do mesmo tamanho, ou seja, com um mesmo número de palavras, diretriz justificada pela necessidade de comparação de itens em termos de sua frequência¹⁴; c)
20 distância entre os textos, que possibilita comparar estágios ou períodos de mudanças mais consolidadas no sistema. O autor aponta ainda necessidade de se fazer conhecida a influência do gênero textual, investigada também pelo trabalho de Vianna.

Desse modo, foi recolhido um *corpus* que abrangesse a língua portuguesa em recortes representativos para o trabalho. Dividiu-se então a história da língua portuguesa em

¹⁴ Como veremos adiante, por meio do cálculo da frequência relativa é possível utilizar textos de tamanhos distintos.

três períodos: arcaico, moderno e contemporâneo. O período arcaico corresponde aos séculos XIII ao XV, o moderno, XVI ao XIX, e XX e XXI e contemporâneo.

Considerando então as diretrizes acima, apresento a seguir um quadro com a descrição (classificação) dos gêneros e tipos textuais utilizados e em seguida três quadros com

5 a descrição dos *corpora* neste trabalho. Vejamos:

Período Arcaico

Código	Corpus	Descrição	Gênero	Datação	Nº Palavras
CDJ	Crônica de D.João	Trecho da Crônica do Rei D. João	Crônica	1437/1450	9.746
TLL	Livro de Linhagens	Trecho do livro de Linhagens	Narrativa	Meados século XIV.	6.907
VER	Vereações do Funchal	Atas/Câmara de Vereadores Funchal - Portugal	Notarial	1485 e 1486	7.901
PED	Conselhos de Duarte	Cartas de Dom Pedro e do Conde Arraiolos	Carta	1426 a 1434	9.438

10

Período Moderno

Código	Corpus	Descrição	Gênero	Datação	Nº Palavras
CJB	Cartas dos primeiros jesuítas do Brasil	Cartas informativas	Carta	1550	9.738
TTB	Tratado da terra do Brasil Gândavo	Texto de caráter informativo	Texto Informativo	1550	9603
AVE	Aves Ilustradas (Novelistas e contistas portuguesas)	Texto moral e fabulas para religiosos nos mosteiros	Fábula	1738	10.925
COB	Cultura e Opulência do Brasil	Tratado sobre como conduzir um engenho de cana	Tratado	1711	10.383
SES	Carta de Sesmaria ao Coronel Matias Barbosa	Documentos cartoriais de Barra Longa - MG	Notarial	1736-1786	6.942

Período Contemporâneo

Código	Corpus	Descrição	Gênero	Datação	Nº Palavras
CFP	Jornal Hoje em Dia	Matérias e crônicas sobre futebol e política.	Crônica	2000	9.999
BUL	Bulas de remédio	Bulas de remédios Tylenol, Gardenal	Bula	1990 a 1994	6.630
DGC	Modelo Funcional da Regressão	Texto acadêmico	Acadêmico	1993 a 1995	26.539

5

Entrevistas

Código	Corpus	Descrição	Gênero	Datação	Nº Palavras
POB	Transcrições de Belo Horizonte	Projeto BH	Entrevista	2002	4.988
CMP	Transcrições do Sul de Minas	Entrevistas para a pesquisa de Amaral (2003)	Entrevista	2002	55.098

10 Como se vê, para o período contemporâneo incluímos textos de fala, na forma de entrevistas, com o intuito, como se verá adiante,, de analisar a frequência, e mais especificamente, o apagamento de **se** nesta modalidade.

3.5 Classificação das ocorrências

15

Nesta seção inicio a apresentação dos dados do *corpus* por meio da classificação e da caracterização dos tipos de **se** encontrados nas ocorrências. Desse modo, descrevo, primeiramente, exemplos dos tipos do pronome encontrados, e, posteriormente, faço uma breve exposição dos seus respectivos aspectos sintáticos e semânticos.

20

Os tipos de **se** foram subdivididos em dois grupos: a) o grupo reflexivo, composto pelas ocorrências dos tipos reflexivo, estilístico, e pronominal; b) o grupo não-reflexivo,

composto pelas ocorrências dos tipos passivo, indeterminado, e ambíguo. Essa classificação levou em consideração o conceitos de voz medial reflexiva, que compreenderia o primeiro grupo, e a voz medial passiva, que compreenderia o segundo¹⁵. Vejamos abaixo o quadro com as respectivas siglas:

5 (12)

Descrição dos tipos e siglas do pronome se

Grupo Reflexivo	Sigla	Grupo Não-reflexivo	Sigla
Reflexivo	Ref	Passivo	Pas
Estilístico	Est	Ambíguo	Amb
Pronominal	Pro	Indeterminado	Ind

10

Desse modo, em relação ao tipo reflexivo, foram encontrados: a) reflexivo simples; b) reflexivo recíproco, no qual a reflexividade se expressa por uma ação recíproca realizada pelos dois agentes; e c) reflexivo duplicado, no qual o pronome é duplicado com o intuito de reforçar a reflexividade. Vejamos os exemplos:

15 (13)

a) Reflexivo

e o time SE recuperou e a situação é boa. (CFP)

20

b) Reflexivo Recíproco

A troca disto SE vendião huns aos outros, (TTB)

c) Reflexivo Duplicado

o negro SE sentiu-se marginalizado... (CMP)

25

Quando à estrutura temática, os verbos em estrutura reflexiva atribuem dois papéis temáticos simultaneamente, isto é, [agente] e [paciente], que estão coindexados por meio do pronome **se**. Notou-se também a partir das ocorrências do *corpus* que o agente é sempre interpretado como [+animado]. Quanto à estrutura temática, o verbo é, portanto, sempre [+bi-argumental].

30

¹⁵ Para mais detalhes a respeito desses conceitos conferir o capítulo I.

Nas orações em que se encontram o se-reflexivo, o pronome **se** apresenta a possibilidade de sua substituição pelas formas *a si mesmo*, *a si mesma*, etc, e recebe caso acusativo, o que não acontece com os tipos estilístico e pronominal, como no exemplo abaixo:

5 (14)

a) o time SE recuperou... (CFP)

Que pode ser interpretado como:

10

b) o time SE recuperou [*a si mesmo*]

Tal critério foi então o utilizado para a determinação das ocorrências do corpus.

Há ainda, como dissemos, construções em que aparece o *se estilístico*. Nesses casos, tem-se um verbo intransitivo: *Riu-SE a ninfa e disse* (AVE)

15

Nas orações em que se encontra o *se-estilístico*, o pronome **se** exprime a vivacidade ou espontaneidade do movimento da ação executada pelo sujeito, mas não recebe nenhum caso, diferentemente do reflexivo que recebe acusativo, como dito anteriormente. Vejamos o exemplo:

20

(15)

a) e com isto SE tornou para a sua pousada.(AVE)

b) e com isto SE tornou ~~a si mesmo~~ para a sua pousada

25

Assim, não é possível em (15b) a interpretação reflexiva, já que o verbo *tornar* não é um verbo transitivo direto, portanto, não exigindo complemento:

Foi percebida uma preponderância dos verbos de movimento em estruturas com *se-estilístico* encontrados nos *corpora*, embora isso não se estenda a todos os casos. Os verbos de movimento encontrados são: *meter (entrar), andar, sair, vir, avançar, passar, volver, tornar, começar de, ir (hir), partir, descer e continuar*.

5 A seguir, temos um exemplo do *se* nos chamados verbos pronominais em que, como se sabe, o verbo exige um complemento preposicionado¹⁶. Trata-se então do tipo pronominal:

(16)

10 *sem que o demónio SE esquecesse de o molestar com terríveis visões (AVE)*

Dentre os aspectos caracterizadores do *se-pronominal*, quatro são fundamentais: a) o pronome *se* nos verbos pronominais é parte integrante do radical; b) tal pronome não tem função sintática, ou seja, é uma forma fossilizada nos radicais dos verbos; c) o complemento é sempre preposicionado; d) nas estruturas pronominais o *se* não recebe caso. Ou seja, 15 diferentemente dos demais tipos de *se* aqui analisados que possuem função sintática e significação gramatical, o *se-pronominal*, pelo menos no português atual, não dispõe de qualquer traço sintático ou semântico.

Outro aspecto caracterizador do *se-pronominal* é que ele parece estar restrito a 20 uma classe geral de verbos, os denominados *verbos psicológicos*. Tais verbos psicológicos são assim chamados devido ao fato de eles denotarem um estado emocional e terem, obrigatoriamente, um sujeito experienciador. Segundo Cançado (1996), tais verbos podem ser subdivididos em subclasses de acordo com suas características semânticas e sintáticas. Vale ressaltar que as formas com *se-pronominal* parecem ser compatíveis com todas as classes 25 descritas pela autora, o que faz que o rótulo generalista aqui utilizado *verbos psicológicos* abarque todas as ocorrências encontradas no nosso corpus de análise. Vejamos um exemplo:

¹⁶ O *se* estilístico pode aparecer também com verbos no infinitivo, quer flexionado quer não flexionado: *com dificuldade para SE andarem por serem despinhados e (SES); El rei dom Afonso de Portugal, jazendo pera SE sair deste mundo, (TLL)*

(17)

a) *o gerente SE preocuparia com a utilização das mil horas a (CFP)*

5 Impossível de ser interpretado como:

b) Função Reflexiva: *o gerente SE preocuparia (~~a si mesmo~~) com a utilização*

Desse modo, quanto aos seus traços semânticos, o *se-pronominal* pode ser
10 caracterizado como: em relação ao aspecto sujeito, ele apresenta sempre os traços
[+animado] e [+experienciador], é sempre um ser animado que vivencia um estado
emocional; quanto ao verbo, é sempre [+bi-argumental], ou seja, há a exigência de dois papéis
temáticos, o de [+experienciador] e o de [+experienciado].

Passemos agora o grupo não-reflexivo. No caso do *se-passivo*, é preciso distinguir
15 dois subtipos: f) a passiva sintética: no caso em que a forma verbal apresenta concordância
com argumento interno; e a g) médio-passiva. Vejamos abaixo os respectivos exemplos¹⁷:

(18)

a) Passiva Sintética

20 *obtêm-SE limites de controle mais apertados, (DGC)*

b) Médio-passiva

Estas bananas crião-SE em cachos, (TTB)

25

Dentre as características fundamentais do tipo *se-passivo*, podem ser citadas: a) a
possibilidade de paráfrase utilizando-se o verbo *ser*; b) a concordância entre o termo
interpretado como objeto e o verbo; c) o verbo é sempre [+transitivo]; d) há presença de um

¹⁷ O *se* estilístico pode aparecer também com verbos no infinitivo, quer flexionado quer não flexionado: *com dificuldade para SE andarem por serem despinhados e (SES); El rei dom Afonso de Portugal, jazendo pera SE sair deste mundo, (TLL)*

agente, mas esse é indeterminado; e) o pronome **se** recebe caso acusativo. Vejamos um exemplo do corpus:

(19)

a) que não **SE esperam** efeitos indesejados sobre o lactente.(BUL)

5

Impossível de ser lida como:

b) Função Reflexiva: que não SE esperam *a si mesmos* efeitos indesejados.

10

Mas podendo ser parafraseado pelo verbo ser em:

c) Forma Passiva: que não *são esperados* efeitos indesejados sobre o lactente.

Em relação ao *se-indeterminador*, distinguimos as construções seguintes: (a) com
15 verbo intransitivo na terceira pessoa do singular; (b) verbo transitivo direto seguido de oração
subordinada; (c) com verbo transitivo direto na terceira pessoa do singular e objeto realizado
foneticamente¹⁸(trata-se, como se sabe, das ocorrências ainda condenadas pelo
normativismo); (d) com a forma infinitiva do verbo *ter* sem possibilidade de leitura passiva;
(e) com locução verbal contendo o verbo *ser ou estar* + participio; (f) com formas gerundivas:
20 de verbos transitivos indiretos ou funcionando como auxiliar; (g) com verbos transitivos
indiretos + preposição DE. Vejamos os exemplos abaixo:

(20)

(a) com verbo intransitivo na terceira pessoa do singular:

25

Nesta capitania se vivia seguramente nos peccados (CJB)

(b) com verbo transitivo direto seguido de oração subordinada:

¹⁸ A interpretação de sujeito indeterminado aparece também com verbo modal formando uma locução: *pode-SE cometer duas espécies de erros*, (DGC)

Diz-se que os processos podem ser alterados (DGC)

(c) com verbo transitivo direto e objeto realizado foneticamente:

Geralmente não SE conhece m e s (DGC)

5

(d) com a forma infinitiva do verbo *ter* sem possibilidade de leitura passiva:

para SE ter uma decisão marcada pela certeza real, incontestes. (DGC)

(e) com locução verbal contendo o verbo *ser ou estar* + participípio:

10

digamos , dado outra, digamos U. E outras vezes SE está interessado (DGC)

(f) Com formas gerundivas: de verbos transitivos indiretos ou funcionando como auxiliar:

15

Partindo-SE da idéia estatística de que todo teste de hipótese pode ser (DGC)

(g) Com verbos transitivos indiretos + preposição DE

20

trata-SE da mais charmosa ... (CFP)

Já o *se-indeterminador* apresenta as seguintes características sintáticas e semânticas: a) a impossibilidade de paráfrase utilizando-se o verbo *ser*; b) a discordância entre o termo interpretado como objeto e o verbo; c) não há presença de agente específico, pois a ação verbal apresenta um aspecto genérico; e) o pronome *se* não recebe caso. Vejamos alguns exemplos do *corpus*:

25

(21)

a. tem-SE os seguintes estimadores (DGC)

30

b. necessita-SE somente do conhecimento da probabilidade (DGC)

c. Sabe-SE que o principal propósito de um diagrama de dispersão (DGC)

Como se observa nas descrições acima, a distinção entre as orações do grupo não-reflexivo não se restringe unicamente ao aspecto presença/ausência de concordância sujeito e verbo como pensam e afirmam alguns teóricos da literatura sobre o assunto.

35

Com o intuito de tornar ainda mais clara a distinção *se-passivo* e *se-indeterminador*, vamos considerar alguns aspectos da análise de Raposo e Uriagereka (1999).

Segundo esses autores, a interpretação das orações passivas são: a) compatíveis com referência específica de tempo; b) compatíveis com a existência de um sujeito singular que satisfaça sua descrição; c) incompatíveis com referência genérica; d) incompatíveis com contextos que dispensam especificidade de referência temporal.

5 Por outro lado, de acordo com Raposo e Uriagereka, as orações com *se-indeterminador* são: a) incompatíveis com referência específica de tempo; b) compatíveis com referência genérica; c) compatíveis com contextos que dispensam especificidade de referência temporal; d) incompatíveis com a existência de um sujeito singular que satisfaça sua descrição¹⁹.

10 Algumas construções, no entanto, permitem-nos propor mais de uma análise. Consideraremos que nesses casos temos estruturas ambíguas, como no seguinte exemplo em que, o verbo estando na terceira pessoa do singular e o objeto estando também no singular, não há como aferir se ocorre concordância entre o objeto e o verbo, o que caracterizaria uma passiva sintética, ou se temos o *se* indeterminador do sujeito:

15 (22)
(a) construções ambíguas: *se-passiva* ou *se-indeterminador* do sujeito
donde SE colnsegue mais proveito são assusces, (TTB)

20 É de se supor, porém, que o que chamamos de ambigüidade resulta apenas da nossa dificuldade de discernir as duas análises: ao empregá-las, o usuário da língua seguramente realizou, internamente, uma análise para orações como as que estão acima, atribuindo-lhes uma das estruturas pertinentes.

¹⁹ Veremos mais detalhes sobre os aspectos semânticos do *se* no capítulo cinco.

CAPÍTULO IV – ANÁLISE QUANTITATIVA

4.1 Apresentação

5

Após ter apresentado no capítulo precedente as diretrizes metodológicas do trabalho em questão, passo neste capítulo à exposição dos resultados da quantificação dos *corpora*.

10

Neste trabalho, a quantificação das ocorrências dos *corpora* divide-se em duas partes, que correspondem, por sua vez, aos diferentes objetivos e naturezas dos cálculos estatísticos utilizados: a) primeiro, apresento os resultados das porcentagens, frequências absolutas e relativas (por mil) considerando as ocorrências dos *corpora* nos três períodos analisados, por texto, e o total do período; b) posteriormente, apresento os resultados do teste qui-quadrado, que calcula a probabilidade de a alteração da frequência dos tipos do pronome *se* ser relevante, ou seja, demonstra, em termos estatísticos, se a variação da frequência está relacionada *diretamente à interferência* dos períodos analisados.

15

Com o intuito de esclarecimento dos conceitos estabelecidos, apresento abaixo as definições de *frequência absoluta* e *relativa* utilizadas neste trabalho, bem como suas fórmulas. Vejamos as definições abaixo:

20

- *Frequência absoluta*: constitui-se da divisão do número de ocorrências por função pelo total de ocorrências por período.

25

- *Frequência Relativa*: constitui-se da divisão do número de ocorrências por função pelo total de palavras do texto multiplicado. O valor retornado é a frequência da função no texto ou período analisado a cada mil palavras.

Os conceitos acima podem ser sintetizados pelas fórmulas abaixo:

(23)

$$\text{Frequência Absoluta} = \frac{F^X}{\Sigma F}$$

5

$$\text{Frequência Relativa} = \frac{F^X}{\Sigma (\text{Palavras})} * 1000$$

10 Passo então à apresentação dos resultados da análise quantitativa dos textos dos períodos medieval, moderno e contemporâneo.

4.2.1 Período Arcaico

15

4.2.1.1 Perfil por texto e tipo

Início a apresentação dos resultados da análise quantitativa com o texto mais
 20 antigo do *corpus*, que se intitula *Trecho do Livro de Linhagens* (TLL). O texto em questão contabiliza um total de 69 ocorrências com o pronome **se**, sendo que só foram registradas três ocorrências de **Ind**, quatro **Pro** e cinco **Amb**. Vejamos a tabela abaixo com os dados do texto:

25 **Tabela 1**
Descrição das porcentagens e frequências por tipo no Trecho do Livro de Linhagens

TLL	Freq Ab.	Porc. %	Freq.%
Ref	30	43,5	4,34
Est	17	24,6	2,46
Pro	4	5,8	0,58
Pas	10	14,5	1,45
Amb	5	7,2	0,72
Ind	3	4,3	0,43
Total	69	100,0	9,99

30

35

Percebe-se pela tabela 1 que, em termos percentuais, o tipo **Ref** é o que mais ocorre no texto analisado, correspondendo a 43% do total das ocorrências, seguido do tipo **Est**, com 24,6%. Somado, o grupo reflexivo (**Ref**, **Est** e **Pro**) corresponde a 73,9% do total das ocorrências, o que constitui um dado muito significativo.

5 Por outro lado, em contraposição às altas taxas de ocorrência do grupo reflexivo, observa-se uma baixa ocorrência do grupo não-reflexivo, **Pas**, 14,5%, **Amb**, 7,2%, e **Ind**, 4,3%, que correspondem, somados aos demais tipos não-reflexivos, a 26,1%.

Em relação à frequência do pronome **se** no texto, em todos os tipos analisados, ela é de 9,99%. O grupo reflexivo tem frequência 7,38%, enquanto o grupo não-reflexivo,
10 2,61%. Passemos à análise descritiva do próximo texto do período arcaico.

O próximo texto a ser analisado se intitula *Cronica d'el Rei Dom Joham* (CDJ) que contabiliza um total de 90 ocorrências com o pronome **se**, sendo que foram registradas menos de dez ocorrências dos tipos **Pas** e **Ind**, respectivamente, sete e cinco. Vejamos a tabela abaixo com os dados do texto:

15

Tabela 2
Descrição das porcentagens e frequências por tipo na *Cronica d'el Rei Dom Joham*

20

CDJ	Freq Ab.	Porc. %	Freq.%
Ref	22	24,4	2,26
Est	23	25,6	2,36
Pro	18	20,0	1,85
Pas	7	7,8	0,72
Amb	15	16,7	1,54
Ind	5	5,6	0,51
Total	90	100,0	9,2

25

Observa-se pela tabela 2 que, em termos percentuais, o tipo **Est** é o que mais ocorre no texto analisado, correspondendo a 25,6% do total das ocorrências, seguido do tipo
30 **Ref**, com 24,4%, e de **Pro** com 20%. Somado, o grupo reflexivo (**Ref**, **Est** e **Pro**) corresponde

a 70% do total das ocorrências, o que constitui, assim como no texto anterior, um dado muito significativo para nossa análise lingüística.

Por outro lado, em contraposição às altas taxas de ocorrência do grupo reflexivo, observa-se uma baixa ocorrência do grupo não-reflexivo em relação ao total, **Pas** 7,8%, e **Ind** 5,6%, o que corresponde, somados aos demais tipos não-reflexivos, a 30,1% do total das ocorrências.

Em relação à frequência do pronome **se** no texto, em todos os tipos analisados, ela é de 9,2 a cada mil palavras (‰). O grupo reflexivo tem frequência 6,47‰, enquanto que o grupo não-reflexivo somado chega somente a 3,52‰. Passemos à análise do próximo texto do período arcaico.

O texto seguinte a ser analisado se intitula *Vereações* (VER), de 1486, é um dos menos antigos do período arcaico. O texto contabiliza um total de 52 ocorrências com o pronome **se**, sendo que não foi registrada nenhuma ocorrência aos tipos **Est** e **Pro**. Vejamos a tabela abaixo com os dados do texto:

Tabela 3
Descrição das porcentagens e frequências por tipo em *Vereações*

VER	Freq Ab.	Porc. %	Freq.‰
Ref	15	28,8	1,9
Est	0	0	0
Pro	0	0	0
Pas	12	23,1	1,5
Amb	18	34,6	2,3
Ind	7	13,5	0,9
Total	52	100	6,6

Observa-se pela tabela 3 que, em termos percentuais, o tipo **Ref** é que mais ocorre no texto analisado, correspondendo a 28,8% do total das ocorrências, enquanto que, por outro lado, os demais tipos do grupo reflexivo não ocorrem.

Considerando somente as ocorrências de **Ref** 28,8%, o restante se constitui dos tipos do grupo não-reflexivo, ou seja, somados os tipos do grupo reflexivo apresentam um valor menor, em termos de ocorrência, que as demais funções.

Em contraposição às baixas taxas de ocorrência do grupo reflexivo, percebe-se uma alta ocorrência de **Pas** 23,1%, **Amb** 34,6%, e **Ind** 13,4%. Os tipos do grupo não-reflexivo correspondem somados a 71,1% do total das ocorrências.

Em relação à frequência do pronome **se** no texto, em todos os tipos analisados, ela é de 6,58%. Nos tipos do reflexivo, só **Ref**, tem frequência 1,9%, enquanto que os restantes, o grupo não-reflexivo (grupo não-reflexivo), somados chegam a 4,68%. Passemos à análise descritiva do próximo texto do período arcaico.

O último texto a ser analisado do período em questão se intitula *Carta que o Jfante dom pedro emujou a el rey de Brujas* (PED). O texto contabiliza um total de 103 ocorrências com o pronome **se**. Vejamos a tabela 4 abaixo com os dados do texto:

Tabela 4
Descrição das porcentagens e frequências por tipo na *Carta que o Jfante dom pedro emujou a el rey de Brujas*

PED	Freq Ab.	Porc. %	Freq.%
Ref	21	20,4	2,23
Est	4	3,9	0,42
Pro	9	8,7	0,95
Pas	21	20,4	2,23
Amb	41	39,8	4,34
Ind	7	6,8	0,74
Total	103	100	10,91

Observa-se pela tabela 4 que, em termos percentuais, o tipo **Ref**, entre os tipos do grupo reflexivo, é o que mais ocorre em PED, correspondendo a 20,4% do total das ocorrências, enquanto que, por outro lado, os demais tipos do grupo reflexivo ocorrem em menor frequência, **Pro** 8,7% e **Est** 3,9%. Somados, os tipos do grupo reflexivo correspondem

a 33% do total, apresentando um valor menor, em termos de ocorrência, que os demais tipos do grupo não-reflexivo.

Note-se que, em contraposição às baixas taxas de ocorrência dos tipos do grupo reflexivo, percebe-se uma alta ocorrência dos tipos **Amb** 39,8% e **Pas** 19,4%. Os tipos do grupo não-reflexivo correspondem, somados, a 67% do total das ocorrências.

Em relação à frequência do pronome **se** no texto, em todos os tipos analisados, ela é de 10,91%. Os tipos do grupo reflexivo somados têm frequência de 3,6%, enquanto que os tipos restantes, o grupo não-reflexivo, somados chegam a 7,31%. Passemos à análise geral do período.

4.2.2.1 Perfil geral do período arcaico

Nesta seção é apresentada a quantificação de todo o período arcaico, considerando os números totais das ocorrências dos quatro textos analisados. Os textos, em seu conjunto, contabilizam um total de 314 ocorrências distribuídas em todos os tipos analisados. Vejamos a tabela abaixo com os dados do texto:

Tabela 5
Descrição total das porcentagens e frequências por tipo no período arcaico.

Total	Freq Ab.	Porc. %	Freq.%
Ref	88	28,0	2,64
Est	44	14,0	1,32
Pro	31	9,9	0,93
Pas	50	15,9	1,50
Amb	79	25,2	2,37
Ind	22	7,0	0,66
Total	314	100	9,40

Considerando os dados da tabela 5 acima, pode-se concluir que, em termos percentuais, o tipo reflexivo, entre todos os tipos, é o que mais ocorre, correspondendo

sozinho a 28% do total das ocorrências, enquanto que os demais tipos do grupo reflexivo ocorrem em menor frequência, **Est** 14% e **Pro** 9,9%. Somados, os tipos do grupo reflexivo correspondem a 51,9% do total.

Em contraposição às taxas significativas de ocorrência do grupo reflexivo, sejam
5 ressaltados, entre as ocorrências do grupo não-reflexivo, os valores de **Amb** 25,2%,
Pas 15,9%, que se destacam. Somados, os tipos do grupo não-reflexivo correspondem a
41,1% do total das ocorrências dos textos analisados.

A frequência global do pronome **se** no período arcaico é 9,4%. O tipo com maior
índice de frequência no período é **Ref** 2,64%, seguido de **Amb** 2,37%. Os tipos do grupo
10 reflexivo somados têm frequência total de 4,89% no período arcaico, enquanto que os tipos
restantes, o grupo não-reflexivo, somados chegam a 4,51%. Passemos à análise descritiva dos
resultados do período moderno.

4.2.2 Período Moderno

15

4.2.2.1 Perfil por texto e tipo

20

Início a apresentação dos resultados da análise quantitativa do período moderno
com o texto que se intitula *Cartas dos primeiros jesuítas no Brasil* (CJB). O texto em questão
contabiliza um total de 123 ocorrências com o pronome **se**, sendo que foram registradas
somente três ocorrências dos tipos **Est** e **Ind**. Vejamos a tabela abaixo com os dados do texto:

25

Tabela 6
Descrição das porcentagens e freqüências por tipo nas
Cartas dos primeiros jesuítas no Brasil

5

10

15

CJB	Freq Ab.	Porc. %	Freq.‰
Ref	37	30	3,80
Est	3	2	0,31
Pro	17	14	1,75
Pas	22	18	2,26
Amb	41	33	4,21
Ind	3	2	0,31
Total	123	100	12,63

20

Observa-se pela tabela 6 que, em termos percentuais, o tipo **Ref**, entre os tipos do grupo reflexivo, é o que mais ocorre em CJD, correspondendo a 30% do total das ocorrências. Por outro lado, dentre os demais tipos do grupo reflexivo que ocorrem em menor freqüência, só um apresenta valores significativos, que é **Pro** 14%. Somados, os tipos do grupo reflexivo correspondem a 46% do total, apresentando um valor um pouco inferior, em termos de ocorrência, ao dos demais tipos do o grupo não-reflexivo.

25

Considere-se que, em contraposição às altas taxas de ocorrência dos tipos **Ref** e **Pro**, os tipos do grupo não-reflexivo também apresentam valores significativos no período. Deve-se ressaltar, entre as ocorrências do grupo não-reflexivo, os valores de **Amb**, 33% e **Pas** 18%. Somados, os tipos do grupo não-reflexivo correspondem a 53% do total das ocorrências do texto analisado.

30

Em relação à freqüência do pronome **se** no texto, em todos os tipos analisados, ela é de 12,63%. Os tipos do grupo reflexivo somados têm freqüência de 5,86%, enquanto que os restantes, o grupo não-reflexivo, somados chegam a 6,77%.

35

O texto seguinte a ser apresentado se intitula *Tratado da Terra do Brasil*, de Pero Magalhães Gandavo (TTB). O texto em questão contabiliza um total de 205 ocorrências com o pronome **se**. Vejamos a tabela 7 abaixo com os dados do texto:

Tabela 7
Descrição das porcentagens e frequências por tipo em
Tratado da Terra do Brasil

5

10

TTB	Freq Ab.	Porc. %	Freq.‰
Ref	70	34	7,29
Est	10	5	1,04
Pro	12	6	1,25
Pas	47	23	4,89
Amb	53	26	5,52
Ind	13	6	1,35
Total	205	100	21,35

15

Note-se pela tabela 7 que, em termos percentuais, o tipo **Ref**, entre os tipos do grupo reflexivo, é a que mais ocorre em TTB, correspondendo a 34% do total das ocorrências. Por outro lado, os demais tipos do grupo reflexivo ocorrem em menor frequência, não apresentando valores significativos, inferiores a 10% como **Pro** 6%, o segundo mais frequente entre os tipos do grupo reflexivo. Somados, os tipos do grupo reflexivo correspondem a 45% do total, apresentando, como o texto anterior, valor um pouco inferior, em termos de ocorrência, ao dos demais tipos do grupo não-reflexivo.

Em contraposição à alta taxa de ocorrência do tipo **Ref**, os tipos do grupo não-reflexivo também apresentam valores significativos no período. Ressaltem-se, entre as ocorrências do grupo não-reflexivo, os valores de **Pas** 21% e **Amb** 21,5%. Somados, os tipos do grupo não-reflexivo correspondem a 55,1% do total das ocorrências do texto analisado.

Em relação à frequência do pronome **se** no texto, em todos os tipos analisados, ela é de 21,35‰. O tipo com maior índice de frequência é **Ref** 6,66‰, seguida de **Amb** 4,58‰. Os tipos do grupo reflexivo somados têm frequência de 9,57‰, enquanto que os restantes, o grupo não-reflexivo, somados chegam a 11,78‰. Passemos ao terceiro texto a ser analisado.

O texto seguinte se intitula *Novelistas e contistas portugueses dos séculos XVII e XVIII* (AVE). O texto em questão contabiliza um total de 137 ocorrências com o pronome **se**.

Vejamos a tabela 8 abaixo:

Tabela 8
Descrição das porcentagens e frequências por tipo em
Novelistas e contistas portugueses dos séculos XVII e XVIII

AVE	Freq Ab.	Porc. %	Freq.‰
Ref	21	15	1,92
Est	9	7	0,82
Pro	49	36	4,49
Pas	13	9	1,19
Amb	30	22	2,75
Ind	15	11	1,37
Total	137	100	12,54

5

10

15

A tabela 8 indica que, em termos percentuais, o tipo **Pro**, entre os tipos do grupo reflexivo, é a que mais ocorre em AVE, correspondendo a 36% do total das ocorrências. Por outro lado, dentre os demais tipos do grupo reflexivo que ocorrem em menor frequência, somente **Ref** com 15% apresenta valores significativos. Somados, os tipos do grupo reflexivo correspondem a 58% do total, apresentando valor superior, em termos de ocorrência, ao dos demais tipos do grupo não-reflexivo.

20

Em contraposição à alta taxa de ocorrência o tipo **Pro**, ressaltam-se, entre as ocorrências do grupo não-reflexivo, os valores de **Amb** 22% e **Pas** e **Ind** com 11% cada. Somados, os tipos do grupo não-reflexivo correspondem a 42% do total das ocorrências do texto analisado.

25

A frequência do pronome **se** no texto, em todos os tipos analisados, é de 12,54‰.

O tipo com maior índice de frequência é **Pro** 4,49‰, seguido de **Amb** 2,75‰.

30

Os tipos do grupo reflexivo somados têm frequência de 7,23%, enquanto que os do grupo não-reflexivo, somados chegam a 5,31%. Passemos ao terceiro texto, o penúltimo a ser analisado.

Trata-se de *Cultura e opulência do Brasil na lavra do açúcar: Engenho real moente e corrente* (COB) e contabiliza um total de 253 ocorrências com o pronome **se**. Vejamos a tabela 9 abaixo com os dados do texto:

Tabela 9
Descrição das porcentagens e frequências por tipo em
Cultura e opulência do Brasil na lavra do açúcar. Engenho real moente e corrente

AVE	Freq Ab.	Porc. %	Freq.%
Ref	44	17	4,24
Est	4	2	0,39
Pro	41	16	3,95
Pas	38	15	3,66
Amb	115	45	11,08
Ind	11	4	1,06
Total	253	100	24,37

Observando a tabela 9 é percebido que, em termos percentuais, os tipos **Pro** e **Ref** predominam entre os tipos do grupo reflexivo, correspondendo, respectivamente, a 16% e 17%, do total das ocorrências. Os tipos do grupo reflexivo somados correspondem a 35% do total, apresentando valor muito inferior, em termos de ocorrência, ao dos demais tipos do grupo não-reflexivo.

Em contraposição às baixas taxas de ocorrência do grupo reflexivo, considerem-se, entre as ocorrências do grupo não-reflexivo, os valores de **Amb** 45% e **Pas** 15% que se destacam na quantificação. Somados, os tipos do grupo não-reflexivo correspondem a 64% do total das ocorrências do texto analisado.

Considerando a frequência global do pronome **se** no texto, em todos os tipos analisados, ela é de 24,37%. O tipo com maior índice de frequência é **Amb** 11,08%, seguido de **Ref** 4,24%. Os tipos do grupo reflexivo somados têm frequência total de 8,58%, enquanto

que os restantes, o grupo não-reflexivo, somados chegam a 15,79%. Passemos ao próximo texto do período moderno a ser analisado.

O último texto a ser apresentado do período moderno se intitula *Carta de Sesmaria ao Coronel Mathias Barbosa da Silva* (SES). O texto em questão contabiliza um total de 98
5 ocorrências com o pronome **se**. Vejamos a tabela 10 abaixo com os dados do texto:

Tabela 10
Descrição das porcentagens e frequências por tipo em
Carta de Sesmaria ao Coronel Mathias Barbosa da Silva

SES	Freq Ab.	Porc. %	Freq.‰
Ref	3	3	0,43
Est	2	2	0,29
Pro	7	7	1,01
Pas	31	32	4,47
Amb	42	43	6,05
Ind	13	13	1,87
Total	98	100	14,12

De acordo com a tabela 10 é percebido que, em termos percentuais, os tipos do grupo reflexivo não predominam no texto SES, correspondendo somados a somente 12% do total, e apresentando, desse modo, um valor muito inferior, em termos de ocorrência, ao dos demais tipos do grupo não-reflexivo somados.

Em contraposição às baixíssimas taxas de ocorrência do grupo reflexivo, sejam ressaltados, entre as ocorrências do grupo não-reflexivo, os valores de **Amb** 43%, e **Pas** 32%, que se destacam. Somados, os tipos do grupo não-reflexivo, correspondem a 88% do total das ocorrências do texto analisado, ou seja, quase à sua totalidade.

A frequência global do pronome **se** no texto, em todos os tipos analisados, é de 14,37%. O tipo com maior índice de frequência é **Amb** 6,05%, seguido de **Pas** 4,47%. Os tipos do grupo reflexivo somados têm frequência total de 1,73%, enquanto que os do grupo não-reflexivo, somados chegam a 12,39%. Passemos ao perfil geral do período moderno.

5

4.2.2.2 Perfil geral do período moderno

Nessa seção é apresentada a quantificação de todo o período moderno, considerando os números totais das ocorrências dos cinco textos analisados. Os textos, em seu conjunto, contabilizam um total de 816 ocorrências distribuídas em todos os tipos analisados. Vejamos a tabela abaixo com os dados do texto:

15

Tabela 11
Descrição das porcentagens e frequências por tipo no período moderno

Total	Freq Ab.	Porc. %	Freq.%
Ref	175	21	3,68
Est	28	3	0,59
Pro	126	15	2,65
Pas	151	19	3,17
Amb	281	34	5,90
Ind	55	7	1,16
Total	816	100	17,15

Considerando os dados da tabela 11 acima, pode-se concluir que, em termos percentuais, o tipo **Ref**, entre todos os tipos do grupo reflexivo, é o que mais ocorre, correspondendo sozinho a 21% do total das ocorrências, seguida de **Pro** 15%, enquanto que o tipo **Est** ocorre em menor frequência, isto é, somente 3%. Somados os tipos do grupo reflexivo correspondem a 39% do total.

Em contraposição às taxas significativas de ocorrência do grupo reflexivo, ressaltam-se, entre as ocorrências do grupo não-reflexivo, os valores de **Amb** 34%, o tipo

25

mais recorrente nos *corpora* do período, seguido de **Pas** 19% que se destacam. Somados, os tipos do grupo não-reflexivo correspondem a 61% do total das ocorrências do texto analisado.

A frequência global do pronome **se** no período arcaico é 17,15%. O tipo com maior índice de frequência no período é **Amb** 5,90%, seguido de **Ref** 3,68%. Os tipos do grupo reflexivo somados têm frequência total de 6,96% no período moderno, enquanto que os restantes, o grupo não-reflexivo, somados chegam a 10,19%. Passemos à análise descritiva dos resultados do período contemporâneo.

10

4.2.3 Período Contemporâneo

15

4.2.3.1 Perfil por texto e tipo

20

Início a análise do período contemporâneo por um conjunto de crônicas intitulado *Hoje em Dia - Crônicas Futebol e Política de Minas* (CFP). As crônicas selecionadas contabilizam um total de 58 ocorrências com o pronome **se**, sendo que não foram registradas ocorrências do tipo **Est**. Vejamos a tabela 12 abaixo com os dados do texto:

25

Tabela 12
Descrição das porcentagens e frequências por tipo em
Hoje em Dia - Crônicas Futebol e Política de Minas

30

35

Total	Freq Ab.	Porc. %	Freq.%
Ref	19	33	1,90
Est	0	0	0,00
Pro	20	34	2,00
Pas	3	5	0,30
Amb	11	19	1,10
Ind	5	9	0,50
Total	58	100,0	5,80

De acordo com a tabela 12 é percebido que, em termos percentuais, os tipos do grupo reflexivo predominam no texto CFP, **Ref** 33% e **Pro** 34%, correspondendo somados a 67% do total. Note-se que os tipos do grupo reflexivo apresentam, desse modo, um valor bem superior, em termos de ocorrência, ao dos demais tipos do grupo não-reflexivo somados.

5 Em contraposição às altas taxas de ocorrência do grupo reflexivo, sejam ressaltados, entre as ocorrências do grupo não-reflexivo, os valores significativos de **Amb** 19%, e **Ind** 9%, que se destacam. Somados, os tipos do grupo não-reflexivo correspondem a 33% do total das ocorrências do texto analisado.

10 A frequência global do pronome **se** no texto, em todos os tipos analisados, é de 5,8%. O tipo com maior índice de frequência é **Pro** 2,0%, seguido de **Ref** 1,9%. Os tipos do grupo reflexivo somados têm frequência total de 3,9%, enquanto que os restantes, o grupo não-reflexivo, somados chegam a 1,9%. Passemos ao próximo texto do período contemporâneo.

15 Prossigo à apresentação dos resultados do período contemporâneo analisando o texto Sarapalha, de Guimarães Rosa. O texto contabiliza um total de 49 ocorrências com o pronome **se**, sendo que não foi registrada nenhuma ocorrência de **Pas** e somente uma **Ind**. Vejamos a tabela 13 abaixo com os dados do texto:

20

Tabela 13
Descrição das porcentagens e frequências
por tipo em Sarapalha

25

POB	Freq Ab.	Porc. %	Freq.%
Ref	12	25	0,24
Est	6	12	0,12
Pro	28	57	0,57
Pas	0	0	0
Amb	2	4	0,04
Ind	1	2	0,02
Total	49	100,0	0,99

30

De acordo com a tabela 13 observa-se que, em termos percentuais, os tipos do grupo reflexivo predominam em SAR. Os valores são: **Pro** 57%, **Ref** 24% e **Est** 12% correspondendo somados a 94% do total. Note-se que, em relação aos demais textos do período, os tipos do grupo reflexivo apresentam um valor bem superior, em termos de
5 ocorrência, ao dos demais tipos do grupo não-reflexivo somados. Isso se deve claramente, a meu ver, ao gênero em questão, como veremos adiante.

Entre as ocorrências do grupo não-reflexivo, há dois valores poucos significativos, que são **Amb** 4,1% e **Ind** 2%, e **Pas** que não ocorre.

A frequência global do pronome **se** no texto, em todos os tipos analisados, é de
10 0,99%. O tipo com maior índice de frequência é **Pro** 0,57%, seguida de **Ref** 0,24%.

Os tipos do grupo reflexivo somados têm frequência total de 0,93%, e os do grupo não-reflexivo chegam a 0,6%. Passemos ao próximo texto do período contemporâneo.

O próximo texto do período contemporâneo a ser analisado constitui-se de um conjunto de bulas de remédio intitulado *Bulas de remédio* (BUL). As bulas selecionadas
15 contabilizam somente 11 ocorrências com o pronome **se**, sendo que só foram registradas ocorrências dos tipos **Ref**, **Pro**, **Pas** e **Amb**. Vejamos a tabela 14 abaixo com os dados do texto:

20

Tabela 14
Descrição das porcentagens e frequências por tipo em
Bulas de remédio

BUL	Freq Ab.	Porc. %	Freq.%
Ref	4	37	0,61
Est	0	0,0	0,00
Pro	4	36	0,61
Pas	2	18	0,30
Amb	1	9	0,15
Ind	0	0	0
Total	11	100,0	1,67

25

A tabela 14 indica que, em termos percentuais, os tipos do grupo reflexivo predominam em BUL, **Ref** 36,4% e **Pro** 36,4%, correspondendo somados a 72,8% do total. Note-se que, como nos dois textos anteriores, os tipos do grupo reflexivo apresentam um valor bem superior, em termos de ocorrência, ao dos demais tipos do grupo não-reflexivo somados.

Entre as ocorrências do grupo não-reflexivo, há somente dois valores, que são **Pas** 18,2 %, e **Amb** 9,1% sendo que os outros tipos não ocorrem. Somados, os tipos do grupo não-reflexivo correspondem a 37,2% do total das ocorrências do texto analisado.

O pronome **se** no texto tem frequência global, em todos os tipos analisados, de 1,67‰. Os tipos com maior índice de frequência são **Ref** e **Pro** com 0,61‰ cada. Os tipos do grupo reflexivo somados têm frequência total de 1,22‰, e os do grupo não-reflexivo chegam a 0,45‰, que é o valor de **Ind**. Passemos ao último texto do período contemporâneo.

O último texto do período contemporâneo a ser analisado intitula-se *Desenvolvimento de Gráficos de Controle Aplicados* (DGC). Ele contabiliza 310 ocorrências com o pronome **se**, sendo que só não foram registradas ocorrências do tipo **Est** e somente duas **Ref**. Vejamos a tabela 15 abaixo com os dados do texto:

Tabela 15
Descrição das porcentagens e frequências por tipo em
Desenvolvimento de Gráficos de Controle Aplicados

DGC	Freq Ab.	Porc. %	Freq.‰
Ref	2	0,6	0,08
Est	0	0,0	0,00
Pro	13	4,2	0,49
Pas	27	8,7	1,02
Amb	133	42,9	5,01
Ind	135	43,5	5,09
Total	310	100,0	11,68

A tabela 15 indica que, em termos percentuais, os tipos do grupo reflexivo em DGC apresentam ocorrência muito baixa, **Pro** 4,2% e **Ref** 0,6%, sendo que os demais tipos do grupo reflexivo não ocorrem. Eles correspondem somados a 4,8% do total. Observe-se que, de forma bem distinta dos textos anteriores, os tipos do grupo reflexivo apresentam também um valor muito superior, em termos de ocorrência, ao dos tipos do grupo não-reflexivo somados.

Entretanto, as ocorrências do grupo não-reflexivo apresentam valores muito mais significativos, como **Ind** 43,5 %, e **Amb** 42,9% sendo que os outros tipos que ocorrem também apresentam valores significativos. Os tipos do grupo não-reflexivo somados correspondem a 95,2% do total das ocorrências do texto analisado, ou seja, quase a sua totalidade.

A frequência global do pronome **se** no texto, em todos os tipos analisados, é 11,68%. Os tipos com maior índice de frequência são **Ind** 5,09‰ e **Amb** 5,01‰. Os tipos do grupo reflexivo somados têm frequência total de 0,57‰ e os do grupo não-reflexivo chegam a 11,11‰. Passemos ao perfil geral do período contemporâneo.

Adiante, no capítulo cinco, apresentamos a quantificação dos dados do *corpus* de entrevistas.

4.2.2.2 Perfil geral do Período Contemporâneo

É apresentada nessa seção a quantificação de todo o período contemporâneo, considerando os números totais das ocorrências dos cinco textos analisados. Os textos, em seu conjunto, contabilizam um total de 428 ocorrências, sendo que foram encontradas somente seis ocorrências do tipo **Est**. Vejamos a tabela abaixo com os dados dos textos:

5

Tabela 16
Descrição das porcentagens e frequências por tipo no período contemporâneo

Total	Freq Ab.	Porc. %	Freq.%
Ref	37	9	0,36
Est	6	1	0,06
Pro	65	15	0,63
Pas	32	8	0,31
Amb	147	34	1,42
Ind	141	33	1,37
Total	428	100	4,15

10

Considerando os dados da tabela 16, observa-se que, em termos percentuais, os tipos do grupo reflexivo no período contemporâneo têm baixa representatividade, **Pro** 15% e **Ref** 9% são mais significativos. Eles correspondem somados a 25% do total. Observe-se que, no geral, apesar de o grupo reflexivo apresentar valores muito superiores em alguns textos específicos, em termos gerais de ocorrência, eles não superam os tipos do grupo não-reflexivo somados no período contemporâneo.

15

Entretanto, os valores gerais no período contemporâneo das ocorrências do grupo não-reflexivo apresentam índices muito mais significativos, como **Ind** 33%, e **Amb** 34%, sendo que **Pas** ocorre também em valores significativos, 8%. Os tipos do grupo não-reflexivo somados correspondem a 75% do total das ocorrências dos textos analisados, ou seja, quase a sua totalidade.

20

A frequência global do pronome **se** no período contemporâneo, em todos os tipos analisados, é 4,15%. Dentre os tipos com maior índice de frequência, **Ind** 1,37% e **Amb** 1,42% se destacam. Os tipos do grupo reflexivo somados têm frequência total de 1,05%, e os do grupo não-reflexivo chegam a 3,1%. Passemos à comparativa geral dos períodos analisados, considerando o tipo.

25

4.3 Análise comparativa dos períodos

4.3.1 O Grupo Reflexivo

5

Nessa seção são apresentados os valores relativos a todos os períodos analisados, considerando, entretanto, o critério do tipo. Optou-se, desse modo, por descrever separadamente os números do grupo reflexivo e do grupo não-reflexivo, para posterior comparação.

10

Na tabela 17 abaixo apresento as porcentagens por tipo em cada um dos períodos analisados, vejamos:

Tabela 17
Descrição das porcentagens do grupo reflexivo nos períodos analisados

15

%	Arcaico	Moderno	Contemporâneo
Ref	28	22	11
Est	14	3	1
Pro	10	15	12
Total	52	40	24

Considerando os dados da tabela 17, pode-se concluir, em termos percentuais, que:

20

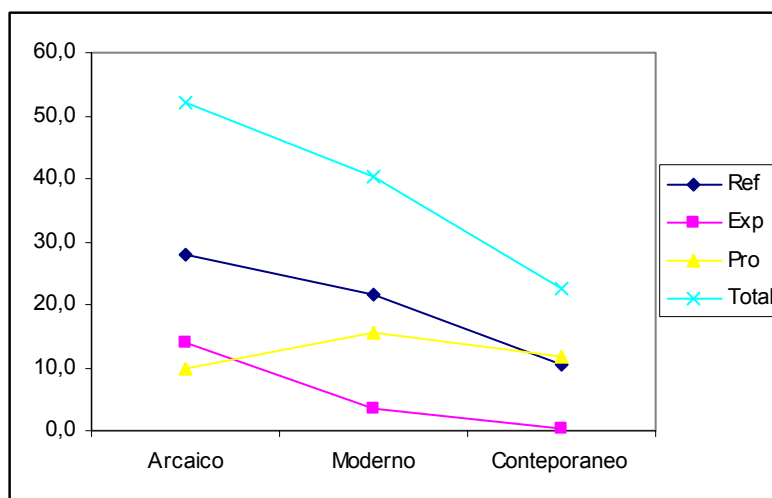
- a) os tipos **Est** e **Ref** apresentam queda gradativa de ocorrência ao longo dos períodos; b) os valores de ocorrência de **Pro** aumentam do arcaico para o moderno, mas diminuem em relação ao contemporâneo; c) o tipo **Est** é o que apresenta maior queda, de 14% para 1,4%; d) os valores totais apresentam queda gradativa, que pode ser apresentada pelo gráfico abaixo:

25

30

35

Gráfico 1
Linhas das porcentagens totais do grupo reflexivo nos períodos analisados



5

Passemos agora à apresentação da tabela 18, que contém os dados relacionados à frequência relativa do grupo reflexivo nos três períodos analisados:

10

Tabela 18
Descrição das frequências relativas do grupo reflexivo nos períodos analisados

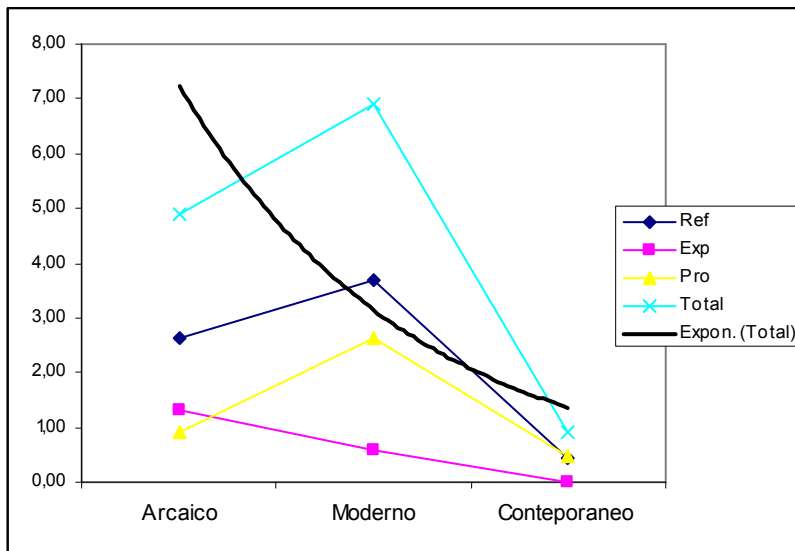
F. Rel (%)	Arcaico	Moderno	Contemporâneo
Ref	2,64	3,68	0,43
Est	1,32	0,59	0,06
Pro	0,93	2,65	0,48
Total	4,88	6,91	0,97

15

Observando os valores da tabela 18, pode-se concluir, em termos de frequência relativa, que: a) o tipo **Est** 1,32% no arcaico e 0,06% no contemporâneo apresenta queda gradativa de frequência ao longo dos períodos; b) a frequência de **Pro** aumenta três vezes do arcaico para o moderno, mas diminui cinco vezes do moderno para o contemporâneo; c) o tipo **Ref** aumenta 32% do arcaico para o moderno, mas cai em torno de nove vezes do moderno para o contemporâneo; d) os valores totais das frequências apresentam uma forma de parábola, sendo que a tendência geral é de queda, o que pode ser verificado pelos gráficos 2 e 3 abaixo:

5

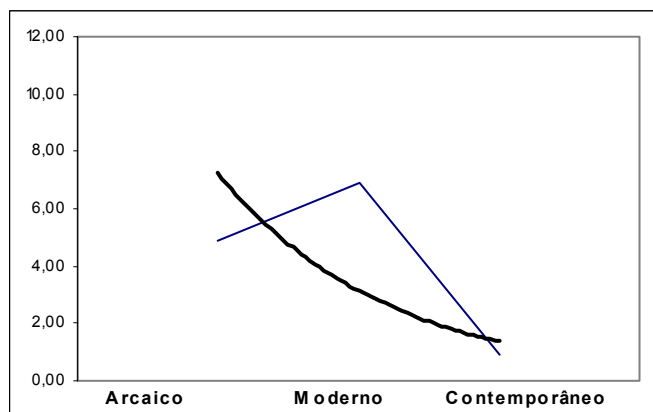
Gráfico 2
Descrição diacrônica das freqüências relativas do grupo reflexivo nos períodos analisados



10

Gráfico 3
Descrição diacrônica dos valores totais das freqüências relativas do grupo reflexivo nos períodos analisados

15



Passemos agora à próxima seção que apresenta os dados relativos às tipos o grupo

20 não-reflexivo nos períodos analisados.

4.3.2 Grupo não-reflexivo

5 Nesta seção apresento os valores das porcentagens e da frequência relativa nos três períodos analisados nos tipos do grupo não-reflexivo. Na tabela 19 abaixo apresento as porcentagens por tipo em cada um dos períodos analisados, vejamos:

10

Tabela 19
Descrição das porcentagens dos tipos do grupo não-reflexivo nos períodos analisados

15

%	Arcaico	Moderno	Contemporâneo
Pas	16	19	7
Amb	25	34	35
Ind	7	7	35
Total	48	60	77

20

Observando-se os dados da tabela 19, pode-se concluir, em termos percentuais, que: a) os tipos **Amb** e **Ind** apresentam aumento de ocorrência ao longo dos períodos, comparando-se os valores do arcaico e do contemporâneo, mesmo que **Ind** mantenha valores estáveis nos períodos arcaico e moderno; b) os valores de ocorrência de **Pas** aumentam do arcaico para o moderno, mas diminuem em relação ao contemporâneo, apresentando uma distribuição em forma de parábola; c) os valores totais das porcentagens apresentam um aumento gradativo, que pode ser observado pelo gráfico 4 abaixo:

25

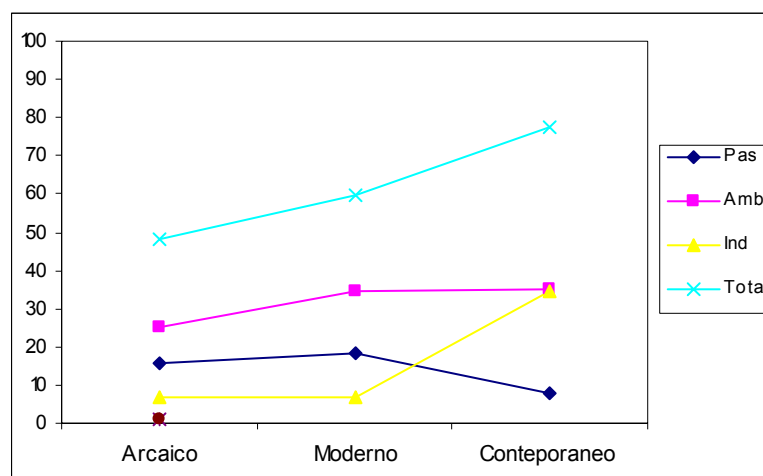
30

35

40

Gráfico 4
Linha das porcentagens totais do
grupo não-reflexivo períodos analisados

5



10 Passemos agora à apresentação da tabela 20, que contém os dados relativos à
 frequência relativa dos tipos do grupo não-reflexivo nos três períodos analisados:

Tabela 20
Descrição das frequências relativas dos tipos
o grupo não-reflexivo nos períodos analisados

15

F Rel (%)	Arcaico	Moderno	Contemporâneo
Pas	1,50	3,17	0,31
Amb	2,37	5,90	1,43
Ind	0,66	1,16	1,40
Total	4,52	10,23	3,15

20 Considerando-se os valores da tabela 20, pode-se concluir, em termos de
 frequência relativa, que: a) o tipo **Ind** apresenta aumento gradativo de frequência ao longo dos
 períodos; b) a frequência de **Amb** aumenta do arcaico para o moderno, mas diminui do
 moderno para o contemporâneo, o que indica uma distribuição de frequência em forma de
 parábola; d) considerados, os valores totais das frequências também apresentam uma parábola
 25 como indicado pelos gráficos 5 e 6 abaixo:

Gráfico 5
Descrição diacrônica das freqüências relativas
do grupo não-reflexivo nos períodos analisados

5

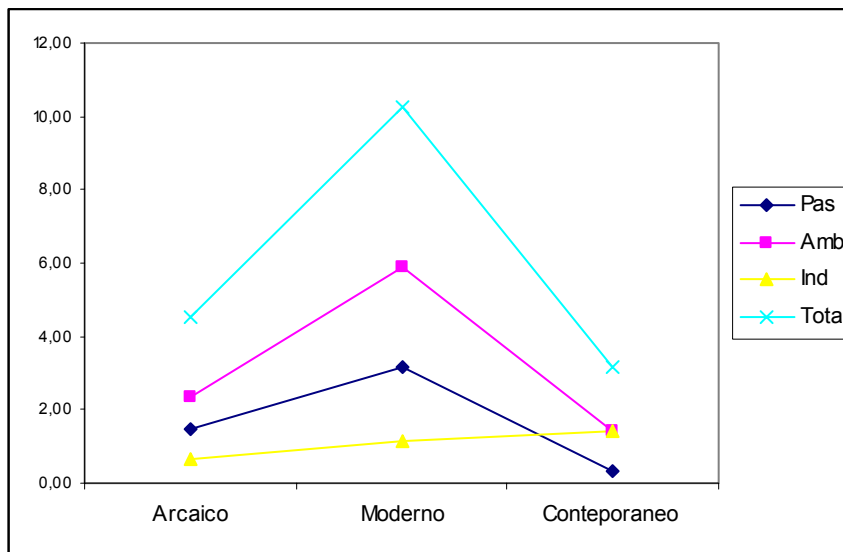
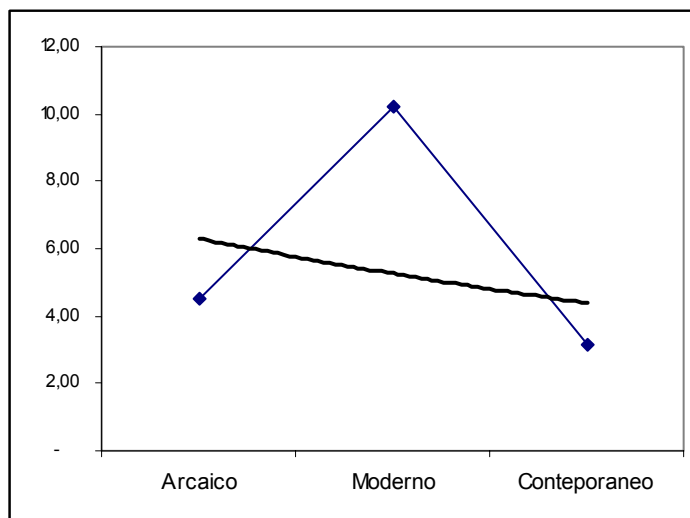


Gráfico 6
Descrição diacrônica dos valores totais das freqüências relativas
do grupo não-reflexivo nos períodos analisados

10



Passemos agora à apresentação da análise global comparativa que contém os dados

15 relativos à freqüência relativa do grupo reflexivo e o grupo não-reflexivo nos três períodos analisados.

20

4.3.3 Análise global comparativa da freqüência dos períodos e tipos

Nesta seção apresento os valores das freqüências do grupo reflexivo e o grupo não-reflexivo, e o valor geral da freqüência do pronome **se** nos períodos analisados. Veja-se a tabela 21 abaixo:

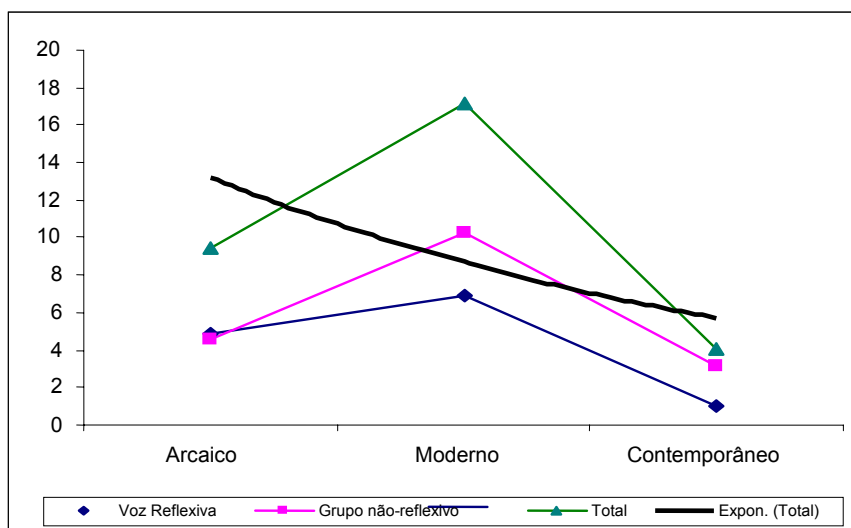
Tabela 21
Descrição das freqüências relativas
dos tipos do pronome se nos períodos analisados

Freq. Rel. (‰)	Arcaico	Moderno	Contemporâneo
Grupo Reflexivo	4,88	6,91	1,05
Grupo Não-reflexivo	4,52	10,23	3,10
Total	9,40	17,15	4,15

Observando a tabela 21 acima, é percebido que a freqüência do grupo reflexivo aumenta gradativamente do arcaico para o moderno e diminui de forma relevante no contemporâneo, o mesmo movimento que acontece com os tipos do grupo não-reflexivo. Ou seja, as freqüências desses tipos apresentam igualmente uma forma de parábola.

Com o intuito de verificar estatisticamente o movimento da freqüência total de **se** nos períodos analisados, apresento abaixo a *linha de tendência* da freqüência do pronome **se**, considerando todos os tipos:

Gráfico 7
Linha de tendência e dos valores totais das freqüências relativas dos tipos do pronome se nos períodos analisados



5

Note-se no gráfico acima que a linha ao centro indica um movimento linear, do arcaico para o contemporâneo, de tendência de queda da freqüência, considerando os valores do pronome se somados em todos seus tipos, sendo que todas as demais linhas apresentam uma linha geral de tendência de queda.

4.4 Cálculo da Probabilidade

15

4.4.1 Detalhamento do cálculo e Interpretação Lingüística

Antes de apresentar as interpretações lingüísticas do teste.qui-quadrado, detalho, a seguir, os objetivos do teste e como ele foi realizado.

20

O objetivo fundamental do teste.qui-quadrado é verificar se existe uma relação de dependência entre duas variáveis. No caso deste trabalho, o objetivo mais específico foi verificar se existia uma relação de dependência entre os períodos analisados e a freqüência. Ou seja, se o fato de ser um tempo x^j influenciaria uma freqüência y^j .

25

O teste qui-quadrado constitui-se de uma comparação da frequência real com a frequência esperada, essa última que seria uma projeção linear da primeira por meio de uma regra de três simples. Abaixo apresento os conceitos utilizados pelo programa Excel:

5

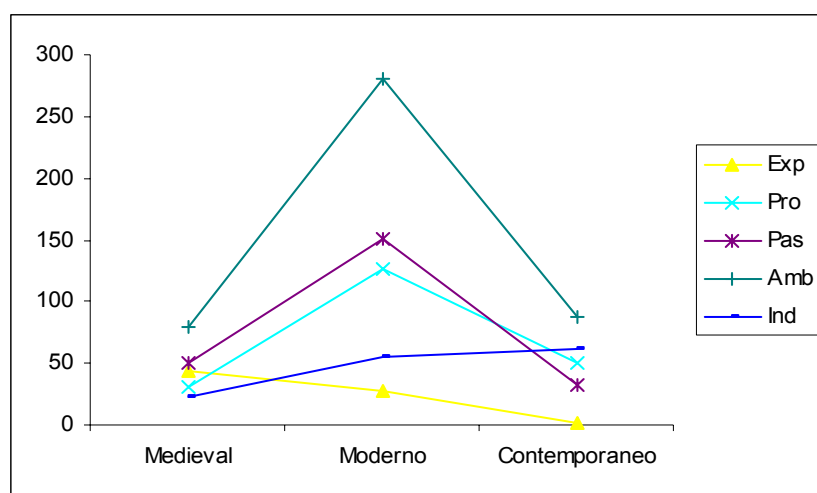
- *Intervalo real*: é o intervalo de dados que contém observações para serem comparadas com os valores esperados.
- *Intervalo esperado*: é o intervalo de dados que contém a razão entre o produto dos totais de linhas e dos totais de colunas e o total geral.

10

Comparando-se então o *intervalo real* com o *intervalo esperado*, pôde-se saber se a relação entre eles é relevante do ponto de vista estatístico. Vejamos, abaixo, os gráficos com os intervalos reais e os intervalos esperados:

15

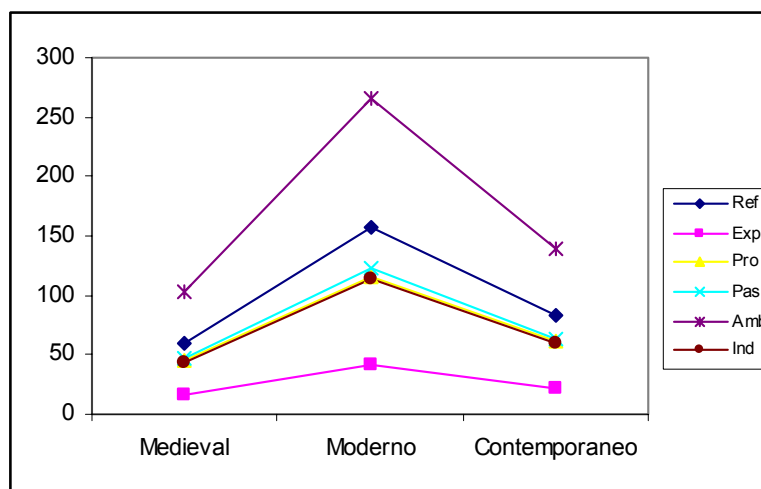
Gráfico 8
Linha das distribuições de ocorrência reais



20

5

Gráfico 9
Linha das distribuições de ocorrência esperadas



10 Era esperado pela progressão das médias que a frequência real se distribuisse de acordo com o gráfico 8, em concordância com as médias das frequências reais. Entretanto, ao se verificar o gráfico 9, é percebido que os vários tipos têm comportamentos distintos nos períodos analisados. Isso indica, *claramente*, que o *fator período* está diretamente ligado à *distribuição da frequência* dos tipos do pronome **se**. Veja-se, por exemplo, que o tipo Est apresenta uma queda gradual ao longo dos períodos, enquanto o tipo **Ind** aumenta sua frequência.

15 Por fim, após comparar, estaticamente, os valores esperados com os valores reais, o teste qui-quadrado gera um valor que indica a probabilidade de erro de não haver uma *correspondência* entre o *fator período* e a *frequência*. O valor gerado pelo teste é quase nulo, e corresponde a:

(24)

7,35083 (E-57)
E-57: vinte e seis zeros à esquerda.

25

O valor acima indica a probabilidade de haver ou não dependência entre os fatores analisados, sendo que, quanto mais próximo a zero, maior a relação de determinação entre eles. Desse modo, com um grau altíssimo, prova-se, por meio do teste que: a) a variação da
5 frequência é estaticamente relevante; b) a existência de um grau de dependência dos fatores período e frequência; c) a probabilidade altíssima de certeza das duas afirmações acima.

4.5 O fator gênero

10 4.5.1 Apresentação

Como indicado na metodologia, diferentes ambientes sintáticos e semânticos
15 poderiam, de certa forma, condicionar a ocorrência de determinados tipos do pronome *se*. Desse modo, meu objetivo nesta seção é apresentar os dados das frequências relativas considerando algumas observações sobre o gênero do texto e apontar possíveis condicionamentos desse fator na distribuição dos valores das frequências.

Entretanto, cabe ressaltar que não é possível uma comparação diacrônica entre os
20 textos, considerando o gênero deles, devido ao seu condicionamento histórico muito intenso, como, por exemplo, as cartas informativas sobre o Brasil quinhentista, que são elementos representativos de um dos gêneros mais utilizados no período moderno, enquanto que bulas de remédio são mais encontradas no período contemporâneo. Ou seja, se os gêneros apresentam um alto grau de motivação histórica isso impossibilita uma homogeneidade
25 tipológica dos textos do ponto de vista diacrônico, e uma conseguinte comparação entre eles.

Optou-se, assim, por se trabalhar somente com os períodos moderno e contemporâneo, pelo fato de eles apresentarem características mais claras e objetivas quanto aos gêneros considerados, já que no período medieval os gêneros ainda apresentavam muito hibridismo (misturavam traços de tipos narrativos, descritivos e argumentativos, o que

impossibilitaria uma classificação mais objetiva) e não estavam, de certa forma, bem constituídos do ponto de vista lingüístico, o que tornaria então a análise desse período imprecisa e sem objetividade.

Outra questão a ser destacada é que a análise apresentada é sincrônica. Serão
5 apresentados os valores das frequências por texto em cada período isoladamente, ou seja, não haverá comparação entre os períodos já que eles não apresentam uma perfeita simetria em relação aos gêneros, devido à citada motivação histórica deles.

O objetivo, então, de se analisar separadamente cada período é verificar se os
valores das porcentagens e frequência apresentam uma relevante variação de texto para texto,
10 mesmo nos próprios períodos analisados. Se isso ocorrer, a que tipo de motivação se deve essa variação? Acredito que o estilo individual, a variação dialetal, a idade, etc, poderiam influenciar nessa variação. Mas, a meu ver, o fator mais influente seria o gênero, que, por sua característica de restringir e, ao mesmo tempo, selecionar certos aspectos sintáticos e semânticos de diversos tipos de textos, proporcionaria um ambiente lingüístico específico
15 mais propício ao aparecimento de certas formas e estruturas específicas.

Desse modo, para explicitar melhor essa influência, apresento, na seção seguinte e numa perspectiva sincrônica, os dados das porcentagens e frequência relativas aos textos por período, iniciando pelo período contemporâneo.

20 4.5.2 Período contemporâneo

Meu objetivo nesta seção é confirmar, por meio dos dados do período contemporâneo, a hipótese da estreita relação entre a frequência de determinados tipos do pronome **se** e os gêneros dos textos analisados nos seus diversos ambientes sintáticos.

25 Para realizar tal tarefa, apresento, primeiramente, as características gerais dos gêneros encontrados no período contemporâneo de acordo com o texto. Meu intuito é delimitar as seleções lingüísticas que cada gênero apresenta e verificar, posteriormente, se

elas estão relacionadas a tipos específicos tipos do pronome **se**, considerando os valores da frequência e das porcentagens.

Desse modo, inicio minha caracterização pelo conjunto de crônicas retiradas do Jornal Hoje em Dia, intitulado *Crônicas Futebol e Política de Minas* (CFP). Trata-se, a meu
5 ver, do gênero crônica, que apresenta as seguintes características: a) relata de forma referencial fatos colhidos do cotidiano; b) enfatiza a verdade; c) usa a terceira pessoa; d) linguagem de acordo com variedade mais formal da língua; e) limitação do código, isto é, recursos estilísticos e vocabulário mais reduzido, o que aumenta a comunicabilidade.

O corpus intitulado *Bulas de remédio* (BUL) não foi caracterizado num gênero
10 específico, pois apresenta elementos próprios da descrição, ao mesmo tempo que apresenta também alguns aspectos do texto argumentativo, como a impessoalidade. Desse modo, por tratar-se de um gênero híbrido, optei por não considerá-lo nessa análise.

Analiso também o texto de Guimarães Rosa intitulado *Sarapalha*, que foi caracterizado como sendo do gênero literário, apresentando as seguintes características: a)
15 relata de forma artística e pessoal fatos do universo humano; c) objetiva a experimentação do código; d) apresenta elementos da narrativa: fatos, personagens, tempo e lugar; e) apresenta um único núcleo narrativo; f) apresenta aspectos das variedades coloquial da língua.

Por fim, a tese doutoral na área de engenharia de produção intitulada *Desenvolvimento de Gráficos de Controle Aplicados*, cuja autoria é de Silvia dos Santos de
20 Almeida, foi caracterizada como gênero científico e apresenta as seguintes características: a) defende-se um ponto de vista; b) o ponto de vista é fundamentado em argumentos pretensamente lógicos; c) estruturação básica: introdução (idéia principal), desenvolvimento (argumentos), conclusão (confirmação da idéia principal); d) linguagem de acordo com variedade padrão; e) o autor costumeiramente coloca-se de modo impessoal ou em primeira
25 pessoa; f) presença de palavras ou expressões impessoais que introduzem comentários (é provável que, trata-se de, não se pode observar que, convém salientar que, etc).

Por fim, os *corpora* de entrevistas intitulados *Projeto POBH (POB)* e *Transcrições das gravações realizadas em Campanha, Minas Novas e Paracatu (CMP)* foram caracterizados como gênero entrevista, que apresenta as seguintes características: a) pressupõe no mínimo dois interlocutores; b) apresenta estilo aspecto informal da língua; c) os interlocutores costumam ser preferencialmente protagonistas ou observadores no diálogo; d) os interlocutores colocam-se de modo pessoal, ou seja, primeira pessoa; e) costuma apresentar elementos da narrativa: fatos, personagens, tempo e lugar, entretanto, na perspectiva de um observador factual ou seja, de eventos não fictícios; f) o narrador costuma ser um experienciador ou observador de acontecimentos e fenômenos.

Consideradas então as características dos gêneros, apresento abaixo os dados das porcentagens e das frequências absolutas por texto no período contemporâneo:

Tabela 22
Descrição das porcentagens e frequências relativas
por texto no período contemporâneo

Tipo	Grupo Reflexivo		Grupo Não-reflexivo	
	%	‰	%	‰
CFP	67,24	3,90	32,76	1,90
SAR	93,87	0,93	6,13	0,06
POB	66,67	1,20	33,33	0,60
BUL	72,73	1,21	27,27	0,45
CMP	84,38	0,49	15,63	0,09
DGC	4,84	0,57	95,16	11,12

Por meio dos valores acima apresentados, note-se que o texto DGC, caracterizado pelo gênero argumentativo escrito, é o que apresenta o maior índice percentual de ocorrência dos tipos do grupo não-reflexivo, 95,12 %, enquanto que entre os tipos do grupo reflexivo correspondem a somente 4,84 % do total no mesmo texto. Os valores de frequência relativa dos tipos nesse texto são também muito relevantes: 11,12 para o grupo não-reflexivo e somente 0,57 para o grupo reflexivo.

Explica-se essa alta ocorrência do grupo não-reflexivo em DGC, a meu ver, pelo ambiente sintático e semântico típico aos textos argumentativos escritos como, por exemplo, a opção por formas mais impessoais e a utilização da forma padrão, que propicia o emprego de formas passivas.

5 De forma oposta aos valores acima referentes ao texto DGC, os valores de CMP, caracterizado pelo gênero entrevista, apresenta um alto índice de ocorrência do grupo reflexivo, 84,38%, enquanto que o grupo não-reflexivo corresponde a somente 15,63 % do total. Muito relevantes também são os valores das frequências relativas do grupo reflexivo e o grupo não-reflexivo, que são muito baixos: 0,49 para entre os tipos do grupo reflexivo e 0,09
10 para o grupo não-reflexivo.

Explica-se em CMP, como em DGC, a alta ocorrência de determinados tipos pelo ambiente sintático e semântico próprio aos textos, nesse caso, às entrevistas. Nelas, se opta por formas mais pessoais, e pela utilização da forma coloquial da língua. A baixa frequência relativa do pronome indica o movimento geral de queda de sua ocorrência em todos tipos.

15 De forma semelhante aos valores de CMP, cujo gênero é entrevista, o texto CFP, caracterizado como crônica, apresenta também um alto índice de ocorrência do grupo reflexivo, 67,24 %, enquanto que o grupo não-reflexivo corresponde a somente 32,76 % do total.

As frequências relativas do grupo reflexivo e o grupo não-reflexivo apresentam
20 valores altos relevantes: 3,6 para entre os tipos do grupo reflexivo(valor mais alto do período para esse tipo) e 1,9 para o grupo não-reflexivo.

Em CFP a alta ocorrência do grupo reflexivo está também, a meu ver, associada ao ambiente sintático e semântico dos textos de caráter mais narrativo. Nas crônicas, opta-se por formas mais pessoais, verbos de movimento, bem como a utilização da forma coloquial da
25 língua.

A alta frequência relativa do pronome no tipo reflexivo é motivada, a meu ver, por todos os aspectos do gênero, e mais ainda especificamente pelo tipo de verbo encontrado nesse tipo de gênero: o de movimento.

O texto SAR é o que apresenta, dentre os demais analisados do período contemporâneo, as maiores taxas de ocorrência do grupo reflexivo. Nele, o grupo reflexivo corresponde a 93,87% do total das ocorrências, enquanto que o restante, 6,13% corresponde à frequência do grupo não-reflexivo. Entretanto, apesar da preponderância do grupo reflexivo em SAR, ele não apresenta índice de frequência relativa tão significativo, 0,93%.

A partir dos dados acima, observa-se a relevante influência exercida em SAR de aspectos arcaicos da língua, que tendem, por sua vez, à preponderância do grupo reflexivo. Nota-se que o grupo reflexivo apresenta os tipos de ocorrência mais arcaicos, como Ref, Est e Pro.

4.5.3 Período moderno

15

Como realizado na seção anterior, apresento nesta a caracterização dos gêneros dos textos do período moderno, bem como os valores da frequência, no intuito de verificar a relação de motivação entre esses valores e o ambiente sintático e semântico específico dos gêneros.

20

Início a caracterização considerando os três primeiros textos do período, apresentando características gerais dos gêneros narrativos, o que não implica, entretanto, na postulação da existência de modos e gêneros “puros”, ou seja, conceber que num dado texto ocorra só um ou outro modo, num único gênero. Desse modo, percebi que, de modo geral, nos textos há a predominância de um modo e um gênero, o que não exclui que nesse mesmo texto

25

ocorram passagens descritivas e dissertativas.

O primeiro texto, intitulado *Novelistas e contistas portugueses dos séculos XVII e XVIII* (AVE), apresenta características eminentemente narrativas, específicas dos gêneros conto e novela. Já os dois seguintes *Cartas dos primeiros jesuítas no Brasil* (CJB) e *Tratado da Terra do Brasil* (TTB), mesclam características da narrativa, com alguns aspectos de

5 gêneros mais descritivos. De um modo geral, esses três textos apresentam as seguintes características: a) linguagem mais informal; b) sucessão temporal; c) verbos de movimento, percepção e caracterização de fenômenos e paisagens por meio de substantivos e adjetivos; c) pessoalidade no modo de expressão verbal; d) buscam relatar e documentar aspectos da realidade à qual se direcionam.

10 Outro texto analisado no período foi o intitulado *Cultura e opulência do Brasil na lavra do açúcar. Engenho real moente e corrente*, (COB), que foi caracterizado como gênero argumentativo escrito, e apresenta as seguintes características: a) defende-se um ponto de vista; b) o ponto de vista é fundamentado em argumentos; c) linguagem de acordo com variedade padrão; d) o autor costumeiramente coloca-se de modo impessoal ou em primeira

15 pessoa; e) presença de palavras ou expressões impessoais que introduzem comentários (*é provável que, trata-se de, não se pode observar que, convém salientar que*, etc).

O texto restante a ser caracterizado no período intitula-se *Carta de Sesmaria ao Coronel Mathias Barbosa da Silva* (SES), e é identificado como o gênero notarial, apresentando as seguintes características: a) forma padrão da linguagem; b) forma impessoal

20 ou primeira pessoa; c) conteúdo argumentativo. Considerados os textos, passemos à tabela abaixo com os valores das porcentagens e freqüências relativas do período moderno:

25

30

Tabela 23
Descrição das porcentagens e frequências relativas
por texto no período moderno

5	Tipo	Grupo Reflexivo		Grupo Não-Reflexivo	
	Texto	%	‰	%	‰
	CJB	46,34	5,85	53,66	6,78
	TTB	44,88	9,58	55,12	11,77
	AVE	57,66	7,23	42,34	5,31
	COB	35,18	8,57	64,82	15,80
	SES	12,24	1,73	87,76	12,39

10

Observando os valores acima apresentados, note-se que os textos SES e COB, caracterizados por gêneros que apresentam padrão mais formal e traços mais argumentativos e descritivos, têm os maiores índices percentuais e de frequência dos tipos do grupo não-reflexivo, 87,12 % (SES) e 64,82% (COB), enquanto que entre os tipos do grupo reflexivo, por outro lado, correspondem a somente 12,24 % em SES e 35,18 em COB.

15

Destacam-se também a relevância em SES e COB dos valores de frequência relativa dos tipos: 12,39 e 15,80 para o grupo não-reflexivo respectivamente, e a baixíssima frequência dentre os tipos do grupo reflexivo em SES 1,73.

20

As novelas e contos intitulados *Novelistas e contistas portugueses dos séculos XVII e XVIII* (AVE) apresentam a predominância do grupo reflexivo, 57,66 %, entretanto, em termos de frequência relativa, não há uma considerável diferença entre as médias, 7,23 e o grupo não-reflexivo, 5,31.

25

Os dois seguintes e últimos textos do período a ser analisados *Cartas dos primeiros jesuítas no Brasil* (CJB) e *Tratado da Terra do Brasil* (TTB), que mesclam características da narrativa e da descrição, apresentam ligeira preponderância dos tipos do grupo não-reflexivo, 53,66 % (CJB) e 55,12% (TTB). Destacam-se também os valores das frequências relativas 6,78 (CJB) e, principalmente, 11,77% (TTB), esse último próximo do valor de SES. Passemos à análise conclusiva do fator gênero.

4.5.4 Conclusões

5 Considerando as análises das duas seções anteriores, pode-se afirmar claramente a motivação que os gêneros dos textos possui na construção de um ambiente sintático e semântico que favorece a presença ou ausência de determinadas estruturas linguísticas.

Após então observar os valores de ocorrências por texto nos períodos analisados, pode-se perceber a associação de determinados traços mais específicos dos textos, como, por exemplo, o modo mais narrativo, que cria um ambiente favorável ao aparecimento de construções do tipo reflexivas, ou o modo mais argumentativo, que favorece a presença de construções impessoais, e por sua vez, daquelas com o pronome *se*, como *Observa-se que*, *Diz-se que*, etc. Apresento abaixo de modo mais sistemático minhas conclusões gerais:

15 1. *Quanto à linguagem formal ou informal*

1.1 Gêneros mais formais, que utilizam a norma-padrão, como as teses, dissertações, documentos notariais, apresentam predominância de ocorrência dos tipos do grupo não-reflexivo.

20

1.2 Gêneros mais informais, como as crônicas, novelas, contos, etc, apresentam predominância de ocorrência do grupo reflexivo.

2. *Quanto ao modo: narrativo, descritivo ou dissertativo*

25

2.1 Gêneros com predominância dos elementos da dissertação e da descrição, como ausência de tempo, argumentação, exposição e definição de idéias e

conceitos, etc criam um ambiente sintático e semântico propício às ocorrências dos tipos do grupo não-reflexivo.

5 2.2 Gêneros com predominância dos elementos da narrativa, como sucesso temporal, personagens, ações etc. criam um ambiente sintático e semântico propício à ocorrência do grupo reflexivo.

3. *Quanto às características semânticas*

10 3.1 Gêneros que apresentam conteúdos mais concretos, que narram ou descrevem fatos, personagens, acontecimentos, constituem um ambiente motivador das construções do tipo reflexivas.

15 3.2 Gêneros que apresentam conteúdos mais abstratos, que apresentam ou discutem conceitos, normas, análises, etc. constituem um ambiente motivador das construções o grupo não-reflexivo.

20 O que se conclui então é que há motivação da presença ou ausência de determinadas estruturas de acordo com: a) o grau de formalidade de um texto b) o modo e objetivo para o qual o texto é escrito; c) as características semânticas das palavras e construções do texto.

25 Outra conclusão relevante e de grande amplitude é que essa constatável motivação pode influenciar, de modo bem claro, a direção da análise lingüística. Se um lingüista, ao estabelecer o seu *corpus* de análise, não considerar o gênero como fator relevante na variação dos diversos ambientes sintáticos e semânticos presentes nos textos, poderá comprometer

seriamente sua análise. Imaginemos, por exemplo, se o *corpus* de análise desse trabalho fosse composto somente por textos com traços dissertativos, com certeza, seriam aferidas outras conclusões, muito mais distantes das realidades lingüística, do que uma análise mais global, que incluísse em seu escopo essa mesma variação.

CAPÍTULO V – ANÁLISE DO PROCESSO DE GRAMATICALIZAÇÃO DE SE

5 5.1 Apresentação

Neste quinto e último capítulo apresento a análise dos resultados deste trabalho, ou seja, os argumentos que comprovam a tese de tratamento do percurso diacrônico da forma pronominal *se* como um caso de gramaticalização.

10 Para tal, comento, em primeiro lugar, uma síntese da análise quantitativa dando a devida relevância aos aspectos relativos aos processos de gramaticalização. Em seguida, apresento uma caracterização sintática do grupo reflexivo e não-reflexivo, com o intuito de sustentar, dentro do quadro gerativo, qual grupo seria mais ou menos gramatical.

15 Como veremos adiante, a análise quantitativa em conjunto com a teórica apontam, claramente, que o percurso histórico de *se* na língua portuguesa pode ser caracterizado por um caso de gramaticalização.

5.2 Síntese dos Resultados e a Gramaticalização de *se*

20 Comento nesta seção, como dito acima, os resultados da análise quantitativa com vistas a confirmar a hipótese aventada, que consiste em tratar o percurso diacrônico da forma pronominal *se* como um caso de gramaticalização. Ou seja, considerando a distribuição e a frequência desses resultados nos três períodos analisados, confirma-se, a nosso ver, a hipótese
25 de gramaticalização do pronome *se*.

Consideremos, em primeiro lugar, as frequências apuradas na análise quantitativa. Note-se que há aumento da frequência dos tipos do grupo não-reflexivo em detrimento da frequência dos tipos do grupo reflexivo: os primeiros apareceram em 48% das vezes no período arcaico, em 61% das vezes no período moderno e em 77% no período
30 contemporâneo. Inversamente, a frequência dos tipos do grupo reflexivo diminui: obtivemos

52% no período arcaico, 39% no moderno e 23% no contemporâneo. Como veremos adiante, defenderemos a hipótese de que os tipos do grupo não-reflexivo são “mais gramaticais” do que os do grupo reflexivo.

Em relação à frequência relativa, o comportamento dos grupos reflexivo e não-reflexivo é semelhante: ambos apresentam uma parábola, ou seja, aumentam o valor do período arcaico para o moderno, no primeiro 4,88‰ > 6,9‰, e no grupo não-reflexivo 4,52‰ > 10,23‰, mas esses valores decrescem do moderno para o contemporâneo, no primeiro 6,9‰ > 0,92‰, e no grupo não-reflexivo 10,23‰ > 3,15‰.

Com vistas a confirmar a hipótese de gramaticalização de **se**, os valores da frequência relativa acima apresentados são relevantes porque podem indicar que o processo de gramaticalização de **se** foi mais intenso do período arcaico para o moderno, sendo esse último o ápice do processo. Já no período contemporâneo observa-se uma grande queda de frequência em todos os tipos analisados, o que indica uma tendência geral de desaparecimento de diversos tipos, como **Est**, por exemplo, que praticamente inexistente no período contemporâneo, como veremos abaixo. O que se conclui diante dessa queda é que, no contexto geral dos tipos de **se**, as mais antigas estão alcançando o Estágio Zero do processo, ou seja, encontram-se em estágio avançado de apagamento, que consiste em seu apagamento fônico, sintático e semântico, enquanto que as mais recentes ainda apresentam valores significativos, como **Ind**. Outro dado que será comentado mais adiante é a influência dos gêneros na distribuição da frequência: ambientes sintáticos e semânticos diferentes desempenham um papel fundamental nos valores de frequência, e podem, portanto, interferir na aceleração ou retardamento de certos processos de gramaticalização.

As divergências dos dados das frequências podem esclarecer também, a nosso ver, outros aspectos relevantes da análise. Os resultados indicam que o período moderno apresenta um tipo relevante no desenvolvimento do processo de gramaticalização da forma pronominal

se. Note-se que é nesse período que o número de ocorrências do grupo não-reflexivo passa a ser maior que o do grupo reflexivo, respectivamente, 61% e 39%.

No contemporâneo também se pode observar a preponderância do grupo não-reflexivo sobre o reflexivo.

5 Comparativamente, os dados indicam que em termos de frequência relativa, ou seja, considerando-se a totalidade das ocorrências em relação ao número de palavras, a forma pronominal **se** tende a ocorrer menos através dos últimos séculos, mais precisamente a partir período moderno. Ou seja, mesmo havendo liderança do grupo não-reflexivo nos períodos moderno e contemporâneo, os valores de frequência relativa do pronome **se** caem de forma
10 bastante relevante.

 Considerando as observações acima, fica claro então que a relevância do período moderno no processo de gramaticalização da forma **se** quanto ao aumento de frequência absoluta está no fato de que é nesse período que se operam e consolidam, de forma mais contundente, as alterações semânticas no uso da forma pronominal **se**. Analogamente aos
15 termos do variacionismo, embora não se trate aqui de um caso de mudança, já que não se extinguiu nenhuma forma porque não existe nenhuma competição entre os subtipos funcionais do pronome **se**, poderíamos afirmar assim que o período moderno é o período da consolidação, no sistema gramatical da língua portuguesa, das alterações por que passou tal pronome em sua história.

20 Ressalte-se, entretanto, que o período moderno apresenta também outras especificidades que podem influenciar na diferenciação de sua importância nos processos por nós identificados. Trata-se de um momento histórico – o período moderno - de intensas transformações lingüísticas, e de um modo geral, é nele que as línguas nacionais européias começaram a passar por um processo de normalização, o que pode ter influenciado na
25 consolidação de determinadas estruturas como a passiva, e a inserção do pronome **se** nos radicais de alguns verbos, formando os verbos pronominais. Em hipótese, a queda de Est, por

exemplo, poderia também estar relacionada a essa normatização, devido ao fato de esse tipo ser mais característico de textos menos formais, o que o levou a ser estigmatizado na comunidade escrita do período moderno.

Consideremos agora mais especificamente os dados relativos aos valores das frequências absolutas e relativas dos tipos de **se** analisados. Iniciamos pelo tipo **Ref**, que foi tomado como estágio inicial da gramaticalização e, como veremos nas seções seguintes, será considerado o tipo menos gramatical em relação aos demais subtipos analisados. Seu comportamento, em termos de frequência absoluta, é o seguinte, de 28% no arcaico para 21% no moderno e 10,48% no contemporâneo, ou seja, o tipo **Ref** perdeu espaço através dos tempos em termos de frequência. Com relação aos valores das frequências relativas, **Ref** também apresenta tendência de queda ainda maior comparado aos valores da frequência absoluta, 2,6‰ no arcaico, 3,7‰ no moderno, e 0,61‰ no contemporâneo. A queda de frequência de **Ref** poderia estar associada também à intensificação de uso de formas como *a si mesmo, ele mesmo*, etc., que surgiriam como alternativas para o emprego de formas reflexivas.

O tipo **Ref**, entretanto, apesar de ser o mais antigo, ainda apresenta, de acordo com os dados, os valores mais altos de frequência relativa entre o grupo reflexivo no período contemporâneo, o que nos faz perguntar: Como a forma mais antiga pode apresentar uma frequência mais alta que a das suas inovações? Uma resposta hipotética a essa questão, a meu ver, pode ser a atestação do fato de que a forma matriz continua a preexistir, mesmo diante do desaparecimento de suas inovações, e que ela, em hipótese, poderia ser a última extinta, marcando, definitivamente, o fim do processo. Os próprios valores de queda de frequência do grupo não-reflexivo podem indicar esse fenômeno, pois os tipos de tal grupo, a saber, **Ind** e **Pas** estão, como atestado, em processo de desaparecimento também, mais lento porque se formaram mais tardiamente na língua, e que mesmo assim, **Ref** resiste.

Outro dado relevante da análise quantitativa são os valores de frequência de **Est**, de 14% no arcaico para 3% no moderno e 0,24%, o que é praticamente seu desaparecimento. Em termos de frequência relativa, sua queda é ainda maior, de 1,3‰ no arcaico para 0,6‰ no moderno e 0,01‰ no contemporâneo. O tipo **Est** apresenta então, seus valores mais
5 relevantes no período arcaico, e é que, desse modo, mais se aproxima, do ponto de vista sintático e semântico da matriz **Ref**, dentre todas as suas inovações. A nosso ver, o que tais dados indicam é que o tipo **Est**, sendo o mais antigo a se desenvolver da forma reflexiva de **se**, sofreria um processo de apagamento primeiro, ou seja, desapareceria antes dos demais tipos do pronome **se**, o que é confirmado pelos dados.

10 Diferente do percurso do tipo **Est**, o tipo **Pro** apresenta seus valores mais relevantes, em termos das frequências absolutas e relativas, no período moderno, ou seja, o tipo **Pro** se consolida no sistema gramatical da língua portuguesa no período moderno. O tipo **Pro** aumentou do arcaico para o moderno, de 9,9% para 15%, e apresentou queda significativa no contemporâneo, ou seja, 11,90%. Esses resultados, aliados aos valores da
15 frequência relativa nos períodos arcaico, moderno e contemporâneo, respectivamente, 0,93‰>2,65‰>0,48‰ mostram uma tendência de clara diminuição de frequência ao longo dos períodos analisados, sobretudo quando se considera que, em alguns dialetos do português brasileiro como o mineiro, esse tipo praticamente não ocorre, como nos exemplos: *Ele casou, Ele formou, Ele aposentou*. Os resultados da análise quantitativa indicam também que o tipo
20 **Pro** é mais tardio que **Est**, e que, por isso, o primeiro está ainda menos apagado que o segundo, em termos de frequência. Ou seja, sendo **Pro** mais tardio que **Est**, era esperado que o apagamento do primeiro se apresentasse atualmente menor que o de **Est**, o que se confirma pelos dados.

Os gêneros também desempenham um papel fundamental na compreensão da
25 distribuição das frequências do tipo **Pro**. No texto *Hoje em Dia - Crônicas Futebol e Política de Minas* (CFP), tal tipo apresenta frequência relativa de 2,0%. Comparando-se aos valores

de **Pro** em CFP com os de DGC, 0,49%, POB 0,40%, BUL 0,61% e CMP 0,20%, nota-se claramente que o ambiente sintático e semântico de CFP (verbos de movimentos, ações humanas, aspectos da narrativa), favorece o aparecimento de **Pro**, em detrimento dos demais ambientes dos demais textos analisados.

5 Passemos agora a considerar os resultados das frequências relativas dos tipos não-reflexivos da forma pronominal **se**, que apresentaram valores mais significativos a partir do período moderno. Para efeito de discussão desses valores dentro da hipótese de gramaticalização da forma pronominal **se**, consideremos que **Pas** e **Ind** são as inovações mais recentes da forma de origem indo-européia, e mesmo que, apesar do tipo **Pas** ter sido
10 registrado já no latim vulgar, o grupo não-reflexivo é mais tardio que o reflexivo, e como a forma **Ind** praticamente torna-se produtiva somente a partir do período moderno, ela é a mais recente inovação de **Ref**.

 Primeiramente, em relação aos valores de **Pas**, observa-se o aumento da frequência absoluta dos períodos arcaico para o moderno, mas uma relevante queda desse último período
15 para o contemporâneo: 16%²>19%>8%. Os valores das frequências relativas ainda são mais significantes. Nota-se que a frequência do moderno para o contemporâneo cai mais de dez vezes, como indicam os valores do arcaico, moderno e contemporâneo, respectivamente: 1,50%>3,17%>0,31%.

 Considerando, então, nossa hipótese de gramaticalização, é esperada a diminuição
20 das frequências de **Pas** encontradas na comparação do arcaico com o contemporâneo, e isso é o que acontece. Desse modo, como o tipo **Pas** é o primeiro a se desenvolver a partir da forma de origem, seria esperado que ele se sofresse o processo de apagamento primeiro. Tal fato é o que se observa comparando-se os valores de frequência de **Ind** com os de **Pas** no período contemporâneo. Em termos de porcentagem, **Ind** apresenta 34% enquanto **Pas** 7,62%, e em
25 termos de frequência relativa, **Ind** apresenta 1,40% e **Pas** 0,31%, dados que revelam claramente a tendência de desaparecimento de **Pas** e do incremento de frequência de **Ind**.

Em segundo lugar, o aumento da frequência do tipo **Ind**, que de 7% nos períodos arcaico e moderno passa a 35% no contemporâneo. Entretanto, mais elucidativos são os valores das frequências relativas correspondendo a 0,66% > 1,16% > 1,40%, para os períodos arcaico, moderno e contemporâneo respectivamente. A relevância do valor da frequência

5 relativa de **Ind** consiste na constatação de que do período arcaico para o moderno ele dobra, o que indica, de forma diferente dos dados da porcentagem, que já no período moderno o comportamento de **Ind**, em termos de frequência, apresenta relevância, diferente dos 8% do período moderno. O alto valor da frequência absoluta de **Amb** no período moderno, 34%, indica que nele, muito provavelmente, há uma boa parte de **Ind** que não pôde ser

10 objetivamente contada. Observe-se também que, como será apresentado na próxima seção, o tipo **Ind** é o tipo “mais gramatical” de **se**.

A distribuição das frequências de **Ind** está também condicionada ao gênero textual, isso é o que revelam os resultados da análise quantitativa. No texto DGC, que se constitui de uma tese doutoral intitulada *Desenvolvimento de Gráficos de Controle Aplicados*,

15 o tipo **Ind** apresenta frequência absoluta de 45%, quase a metade do total, enquanto que a frequência relativa 5,09% é a maior do período, e uma das maiores por tipo num texto. Considerando ainda que **Amb** nesse texto apresentou 43%, e 5,01%, isso indica que poderiam ainda haver muitas ocorrências de **Ind** em **Amb**. Desse modo, prova-se claramente que no gênero *texto científico*, isto é, em gêneros mais formais, e com uso de expressões impessoais,

20 observa-se um ambiente sintático e semântico propício à ocorrência de **Ind**.

Outro dado relevante é o incremento da frequência do tipo **Amb** de 25% para 34% do arcaico para o moderno, o que pode indicar também que entre essas ocorrências possa haver um considerável número de **Ind**, o que aumentaria sua frequência nesse período, como também poderia ocorrer o contrário. Ou seja, a existência de **Pas** entre os valores de **Amb**,

25 valores esses que deveriam ser debitados dos valores totais de ocorrência das ocorrências do tipo **Ind**.

Finalmente, considerando e comparando-se os valores totais das frequências relativas nos períodos analisados, a hipótese da gramaticalização é também reforçada. O que se constatou foi o aumento considerável, isto é, de 9,40‰ para 17,15‰, o que é esperado já que, nesse tipo de processo, à frequência do uso inicial se incorporam as frequências das inovações. Entretanto, o que não se esperava é decréscimo robusto do emprego do *se* no contemporâneo, que foi de 4,07‰, valor mais de quatro vezes inferior ao do período moderno. O que esse decréscimo aponta, considerando a hipótese da gramaticalização, é que os tipos mais antigos de *se*, principalmente **Est** e **Pro**, e até mesmo **Ind**, encontram-se em processo de apagamento, ou seja, as inovações de *se* caminham para o **Estágio Zero** da gramaticalização, hipótese que pode se sustentar pelo decréscimo robusto das frequências de todos os tipos de *se*. Para poder comprovar esta hipótese, comento, não seção seguinte, os resultados da análise quantitativa das ocorrências de *se* nas entrevistas do *corpus*. Se de fato há uma tendência de apagamento de *se*, este fenômeno é mais saliente na modalidade de fala.

5.2.1 Análise das entrevistas

Início a análise das entrevistas pelo corpus do Projeto: Projeto POBH (POB), que contabiliza somente 9 ocorrências com o pronome *se*, sendo que só foram registradas ocorrências dos tipos **Ref**, **Pro**, e **Ind**. Vejamos a tabela 24 abaixo com os dados do texto:

Tabela 24
Descrição das porcentagens e frequências por tipo em Projeto POBH

POB	Freq Ab.	Porc. %	Freq.‰
Ref	4	44,4	0,80
Est	0	0,0	0,00
Pro	2	22,2	0,40
Pas	0	0,0	0,00
Amb	0	0,0	0,00
Ind	3	33,3	0,60
Total	9	100	1,80

De acordo com a tabela acima observa-se que, em termos percentuais, os tipos do grupo reflexivo predominam em POB, **Ref** 44,4% e **Pro** 22,2%, correspondendo somados a 66,6% do total. Note-se que, como no texto anterior, os tipos do grupo reflexivo apresentam um valor bem superior, em termos de ocorrência, ao das demais tipos o grupo não-reflexivo somados.

Entre as ocorrências do grupo não-reflexivo, só há um valor significativo, que é **Ind** 33,3 %, sendo que as outros tipos não ocorrem.

A frequência global do pronome **se** no texto, em todos os tipos analisados, é de 1,8%. O tipo com maior índice de frequência é **Ref** 0,8%, seguida de **Pro** 0,4%.

Os tipos do grupo reflexivo somados têm frequência total de 1,2%, e os do grupo não-reflexivo chegam a 0,6%, que é o valor de **Ind**.

O próximo texto do período contemporâneo de fala a ser analisado constitui-se de um conjunto de entrevistas transcritas intitulado *Transcrições das gravações realizadas em Campanha, Minas Novas e Paracatu*. As entrevistas contabilizam 32 ocorrências com o pronome **se**, sendo que só não foram registradas ocorrências do tipo **Pas**. Vejamos a tabela 25 abaixo com os dados do texto:

Tabela 25
Descrição das porcentagens e frequências por tipo em
Transcrições das gravações realizadas em Campanha, Minas Novas e Paracatu.

CMP	Freq Ab.	Porc. %	Freq.%
Ref	15	46,9	0,27
Est	1	3,1	0,02
Pro	11	34,4	0,20
Pas	0	0,0	0,00
Amb	3	9,4	0,05
Ind	2	6,3	0,04
Total	32	100,0	0,58

De acordo com a tabela acima, em termos percentuais, há predominância do grupo reflexivo em CMP, principalmente **Ref** 46,9% e **Pro** 34,4%, correspondendo somados a 84,4% do total. Note-se que, assim como nos textos anteriores, os tipos do grupo reflexivo neste apresentam também um valor bem superior, em termos de ocorrência, ao dos tipos do grupo não-reflexivo somados.

Consideradas as ocorrências do grupo não-reflexivo, há somente dois valores significativos, que é **Amb** 9,4 %, e **Ind** 6,3%, sendo que os outros tipos não ocorrem. Os tipos do grupo não-reflexivo somados correspondem a somente 15,7% do total das ocorrências do texto analisado.

A frequência global do pronome **se** no texto, em todos os tipos analisados, é 0,58%. Os tipos com maior índice de frequência são **Ref** 0,27%, e **Pro** 0,20%. Os tipos do grupo reflexivo somados têm frequência total de 0,5%, e os do grupo não-reflexivo chegam também a somente 0,8%.

Um aspecto de grande relevância que pude observar no *corpus* recolhido de entrevistas foi o apagamento do pronome **se** nos verbos pronominais. Trata-se de uma tendência observada no dialeto mineiro que indica a queda de uso do pronome.

A título de ilustração, apresento a seguir as ocorrências de alguns verbos mais freqüentes no corpus e apresento em termos quantitativos, os índices de apagamento do pronome.

20

(25)

- a. logo quand[o] **eu casei** já era/já tava bem velha né?...
- b. nesse tempo fui pa São Paulo fiquei mais uns tempo em São Paulo uns cinco seis ano em São Paulo... aí **eu vim casei**... logo que eu casei... vô cumeçá lá de casa...
- c. dessa época nunca mais representô... agora depois que **eu casei** passei pra [a]qui
- d. é lá que... agora aqui depois que... **eu casei**... ta[va] fazeno... sete...
- e. parece que foi ontem que **eu casei** qu[e] eu cumecei
- f. é... logo que depois que... larguei a iscola () cumecei a namorá () **casei** agora tô aqui... cheia de filho

30

- g. deles?... ia ... num é muito não mais ia... que o Sérgio... no começo... ele... **casô**... e morava em Belo [Ho]rizonte... mais a família da moça em Valadares cabô puxano pra lá né?...
- h. diferente... e te/ o Vicente que **casô**/que é casado com Dador sua irmã... esse aí
- 5 i. que ia **casá** e num casô (como que) é o nome?
- j. **SE casaram** aqui e a gente...
- k. ele chegô a **SE casá** ô não?

Verbo	Ocorrências	Presença	%	Apagamento	%
FORMAR(-SE)	7	1	14%	6	86%

15 (26)

- a. É **formei** com quaren/aposentei com quarenta e um anos de exercício do I Bê Gê É aqui den[tro] de Minas Nova den[tro] da Comarca de Minas Nova abrangeno... Chapada Berilo Francisco Badaró e Turmalina
- 20 b. até/quando a gente **formô** mesmo... (com) uns dizessete anos... dizessete anos
- c. tem três anos que a gente **formô**... eu principalmente porque agora () eu faço aniversário bem/em fevereiro né?
- d. não... ela é formada... **formô**... tem dois anos né?... e ele istuda... tá na sexta série
- 25 e. **a Edinéia**... **formô** recentemente até... fez istágio aqui em Minas Novas né?... na Policlínica...
- f. ês são... **formô** a família aqui sabe?
- g. o terceiro do sigundo grau a gente **SE formô** junto né?...

Verbo	Ocorrências	Presença	%	Apagamento	%
CASAR(-SE)	11	2	18%	9	82%

35 (27)

- a. e **lembrô** mais/de mais alguém? da cidade?
- b. história? eu **lembrei**
- c. aí gente **lembrei** o nome dele... **Manuel Bandera**
- 40 d. nacional **lembrei**... ((risos)) demorô pa caí a ficha
- e. ah é pois é... **lembrei** veio ladrão aqui
- f. Carlota né?... que eu cunversei **lembrei**
- g. “vô levantá vô oiá... cumecei a rezá o terço num **alembrei** num rezei num alembrei do princípio do
- 45 h. ... cê(s) **SE alembra** do/do Vítor?

Verbo	Ocorrências	Presença	%	Apagamento	%
LEMBRAR(-SE)	8	1	12%	7	88%

(28)

- 5 a. padre Foade... agora há uns três anos... ele ficô aqui vinte e cinco anos... ele istudô/**se ordenô**/ele istudô [a]qui... SE ordenô vei[o] pra cá...
- b. o Júlio... é Júlio... ele é padre... ele **ordenô** pa padre já tem/já tem dois ano que ele ordenô
- c. que o único filho meu... esse que **ordenô** pa padre...

Verbo	Ocorrências	Presença	%	Apagamento	¹⁰ %
ORDENAR(-SE)	3	1	33%	2	67%

15 (29)

- a. É formei com quaren/**aposentei** com quarenta e um anos de exercício do I Bê Gê É
- b. e eu ganhei dele que eu **aposentei** com trinta e oito
- 20 c. ele **aposentô** com trinta e quatro anos de serviço... e eu ganhei dele que eu aposentei com trinta e oito
- d. ao fazer o curso de Adiministração Iscolar ela **aposentô** no ano que eu entrei como substituta ()

Verbo	Ocorrências	Presença	%	Apagamento	%
APOSENTAR(-SE)	4	0	0	4	100%

A seguir apresento os resultados das análises acima sistematizados.

30

Tabela 26
Descrição da frequência por verbo
em relação à frequência total

35

Verbo	Ocorrências	%
CASAR	11	34%
FORMAR	7	21%
LEMBRAR	8	24%
ORDENAR	3	9%
APOSENTAR	4	12%
TOTAL	33	100%

40

Note-se que dentre os verbos analisados *casar(-se)* é o que apresenta maior ocorrência no *corpus* de entrevistas, 34%, seguido por *lembrar(-se)* e *formar(-se)*,

respectivamente, 24% e 21%. Vejamos agora os valores de presença e ausência do pronome **se**:

Tabela 27
Descrição da frequência por verbo
em relação à frequência total de cada verbo

Verbo	Ocorrências	Presença	%	Apagamento	%
CASAR	11	2	18%	9	82%
FORMAR	7	1	14%	6	86%
LEMBRAR	8	1	13%	7	88%
ORDENAR	3	1	33%	2	67%
APOSENTAR	4	0	0%	4	100%
TOTAL	33	5	15%	28	85%

Dentre os verbos analisados e listados à direita da tabela, observe-se que o índice médio de frequência do apagamento do pronome **se** é 85% e o de presença apenas 15%. O verbo *ordenar-se* é o que apresenta maior frequência de presença do pronome 33%, seguido de *casar-se* com 18%, o maior índice entre os demais. Já o verbo *aposentar* apresenta o maior índice de apagamento, 100%, ou seja, nenhuma presença de pronome.

O elevado índice de apagamento do pronome **se** na modalidade falada mostra-nos, assim, que podemos estar diante do **Estágio Zero** da gramaticalização, que avança, provavelmente, em alguns dialetos (como o mineiro) mais do que em outros. Este fenômeno explica, então, o decréscimo robusto das ocorrências de **se** no período contemporâneo.

5.3. Graus de Gramaticalidade da Forma Pronominal **se**

Considerando a revisão da literatura a respeito do assunto, que subdivide os tipos funcionais de **se** em voz média-reflexiva e média-passiva, respectivamente, grupos reflexivo e não-reflexivo, e os dados da análise quantitativa, foi elaborado um esquema de compreensão da trajetória diacrônica do pronome **se**. Tal esquema agrupou, a partir de sua origem reflexiva,

de maneira a estabelecer que usos são “mais gramaticais” que outros? Em outras palavras, que elementos teóricos podem assegurar-nos de que o *se* apassivador é “mais gramatical” que o *se-reflexivo*, e o *se-indeterminador* do sujeito é “mais gramatical” que o *se* apassivador?

Diante de tais questionamentos, apresento, nas próximas seções, a descrição
5 sintática dos grupos reflexivo e não-reflexivo, no intuito de apontar qual deles é mais gramatical. Essa descrição se fará por meio do quadro teórico da Gramática Gerativa, nos seus módulos específicos da Teoria da Ligação, que trata das dependências referenciais, e da Teoria Temática, que estuda a organização da predicação e a sua relação com os argumentos dos verbos.

10

5.4 O estatuto gramatical do grupo reflexivo

15

Dentro do quadro gerativo, o módulo da Teoria da Ligação é o responsável pela estudo das dependências referenciais entre DPs que apresentam potencial de referência, a saber, pronomes, anáforas, e expressões referenciais. Essa área é uma das que tem recebido maior atenção dos gerativistas a partir dos anos 60.

20

Compreende-se, assim, por *potencial de referência*, as expressões lingüísticas que designam entidades (pessoas, coisas, idéias, situações, eventos, ações, etc.) do universo discursivo. Desse modo, o estudo das relações de dependência referencial é o estudo das relações entre os DPs.

25

Por *dependência referencial*, entende-se a situação lingüística na qual o valor referencial de um DP é adquirido de um outro DP presente na enunciação, sendo essas expressões, por sua vez, co-referentes, ou seja, referindo-se às mesmas entidades do universo discursivo. Vejamos alguns exemplos:

(31)

- a. *Ele* pensa que o *pai de Luís* é o culpado por tudo.
- b. *João* barbeou-se.

30

No caso específico de *se*, esta forma apresenta um comportamento anafórico, como em (31b) acima, diferente da forma pronominal *ele* em (31a) e a expressão referencial *O pai do Luis* também em (31a).

O que distingue os pronomes das anáforas é que os primeiros apresentam potencial
5 autônomo de referência, enquanto que as anáforas não. Ou seja, uma anáfora necessita de um antecedente num domínio sintático para ser interpretada. No caso do português, as anáforas compreendem, dentre outras formas, os pronomes reflexivos. As anáforas implicam uma relação de concordância de traços gramaticais de pessoa e número, ou seja, o pronome *se* não é empregado em situações que o sujeito não esteja na terceira pessoa do singular. Vejamos os
10 exemplos abaixo:

(32)

- 15 a. * Nós se viu lá.
[1^a.Pl] [3^a.Sg] [3^a.Sg] onde sg= singular, pl= plural
- b. *Paulo* barbeou-*se*.
[3^a.Sg] [3^a.Sg] [3^a.Sg]

20 Como se percebe acima, a agramaticalidade de (32a) justifica-se por tal estrutura não satisfazer as condições de concordância, que são basicamente homogeneidade dos traços semânticos [gênero], [número], e [pessoa] entre sujeito e o pronome *se*, como ocorre em (32b).

25 As anáforas apresentam também um papel temático, fato que se constitui em um aspecto fundamental em nossa investigação. No modelo gerativo, o módulo responsável pelo estudo dos papéis temáticos é a Teoria Temática.

Nesta Teoria, compreende-se que as expressões lingüísticas se estruturam como um sistema de predicados, ou seja, concebe-se um predicador central e um determinado número de argumentos que lhe completam o sentido, convertendo o predicador numa
30 expressão semanticamente completa. Cada predicador pode ser especificado quanto ao

número de argumentos. Por exemplo, o verbo *chegar* é um predicador de um lugar, o verbo *temer* exige dois argumentos, ou seja, é um predicador de dois lugares. Vejamos outros exemplos:

(33)

- 5 a. Os alunos empurram o professor
b. O professor empurrou os alunos.

Nas orações acima, o verbo *empurrar* é um predicador de dois lugares, ou seja, exige um agente e um paciente. Em (28a), o **DP os alunos** é agente e o **DP o professor** é o
10 paciente, e em (28b) a situação se inverte. Os fatores que determinam essa variação são as características semânticas desses DPs e a função gramatical desempenhada por eles.

Sendo o *se* reflexivo uma anáfora, nos termos da *Teoria da Ligação*, ele deve ser analisado como um elemento gramatical, categorizado como um D (isto é, um Determinante (=”determiner”); cf. Chomsky(1995)).

15 Por outro lado, como vimos, o *se-reflexivo* ocorre com um verbo transitivo direto e é interpretado como uma anáfora. Vamos propor, então, que o *se-reflexivo* recebe um papel temático e que, portanto, é um item [+ argumental], já que a anáfora funciona como um dos argumentos internos do predicador. Veja-se a seguinte representação abaixo:

(34)

20 João barbeou-se; onde: + θ = temático
+ θ

Vamos considerar que o *se-reflexivo* no exemplo 34 é um item argumental, clítico, recebe caso acusativo e é categorizado como D, dispondo dos traços semânticos
25 [3ª. pessoa], [singular], e [masculino] já que está co-indexado com seu antecedente *João*. De acordo com Chomsky (1995, p.230) um item do léxico dispõe dos traços formais, semânticos e fonéticos. Consideremos então que o item **se** tem os traços formais de [3ª. pessoa], [singular], e [masculino], e que seus traços semânticos são determinados por um antecedente, se houver. Assim, o *se-reflexivo* é gerado na posição de objeto e por meio da aplicação da

regra de movimento *mover* α , ele é adjungido ao verbo flexionado, deixando no lugar de origem uma categoria vazia (ou uma cópia, se se preferirem versões mais recentes do Programa Minimalista)

Compreendido o *se-reflexivo* como gramatical, passemos agora à caracterização
5 sintática dos tipos **Est** e **Pro**, que a nosso ver, se desenvolveram do *se-reflexivo*.

Foi dito acima que o critério para se agrupar os tipos **Ref**, **Est**, e **Pro** é o envolvimento da sujeito na ação, ou seja, o pronome como paciente da ação do sujeito recuperado pela anáfora. Espera-se, então, considerando o intuito desse trabalho, que os tipos **Est** e **Pro** apresentem aspectos semânticos semelhantes aos de **Ref**. E isso é o que se verifica,
10 como veremos, apenas do ponto de vista morfológico, pois nos âmbitos sintático e semântico eles apresentam diferenças, embora **Est** e **Pro** se aproximem e os dois se distanciam de **Ref**.
Veamos a frase abaixo:

(35)

- 15 a. [Ele [[se foi] embora]]
[3^a.Sg] [3^a.Sg] [3^a.Sg]
- b. [Ele [[foi] [~~ele mesmo~~] [embora]].
[a si mesmo]
- θ
- 20 c. Ele foi embora
[zero]

Observe-se acima que em (35) o tipo **Est** apresenta a concordância dos traços de
25 pessoa, gênero e número, satisfazendo as condições de referência da anáfora, ou seja, de co-indexação com um referente. Os teóricos da literatura tendem a tratar esse tipo de ocorrência como não apresentando nenhum valor anafórico, mas vejamos esse aspecto mais detalhadamente.

As matrizes semânticas dos verbos de movimento como *partir*, *vir*, *ir*, *sair*
30 implicam um sujeito [+animado] que desempenha uma ação espontânea e reflexa, ou seja, alguém que se movimenta, necessariamente, movimenta a si mesmo, ao seu próprio corpo, e o

faz comumente de forma espontânea, o que sugere que a forma pronominal estilística apresenta resquícios da reflexiva, tantos em seus traços de concordância quanto em seus traços semânticos.

Desse modo, a interpretação de (35a) tem semelhança com a das formas reflexivas, 5 apesar do fato de que estas poderiam ser caracterizadas como mais intencionais, enquanto que as estilísticas seriam mais espontâneas. Outro aspecto relevante em (35) é o fato de que as estruturas estilísticas recaem, predominantemente, nos verbos de movimento com sujeito [+animado], e naqueles que a ação verbal implica um necessário envolvimento do sujeito da oração com ele mesmo. O que se conclui então é que, apesar de o *se-estilístico* não ter 10 representação sintática, ou seja, não possuir nenhuma função gramatical comportando-se, assim, como um afixo, ele apresenta um resíduo de leitura anafórica, que é, no entanto, bastante tênue, embora talvez em períodos pretéritos da língua portuguesa tenha sido mais saliente.

O que auxilia essa especulação é que as estruturas estilísticas têm seu momento 15 mais relevante, em termos de frequência no período medieval, o que nos leva a concluir que o *se-reflexivo* tenha *gerado*, primeiramente, essa forma, e posteriormente, a pronominal, que tem seus maiores índices de ocorrência no período moderno. Digamos que o *se-estilístico* é "o filho mais velho do tipo Ref", e por isso, dentre os tipos de *se*, é o que mais se assemelha a Ref. Note-se também que, como veremos adiante, o resíduo da leitura anafórica de **Est** parece 20 ser menos detectável no *se-pronominal*. Vejamos um exemplo:

(36)
[Ele [se esqueceu das] [fotos]]
[3^a.Sg] [3^a.Sg] [3^a.Sg]

25

Apesar de apresentar a concordância dos traços de pessoa, gênero e número, a forma *se* de (36) não pode ser tomada como anáfora do DP *Ele*, já que o verbo *esquecer-se de*

tem como argumento interno, ou seja, como seu objeto, o DP **as fotos**. Desse modo, ao pronome **se** nesse caso não pode ser aferido nenhum papel temático, como acontece em **Ref**. No caso de **Pro**, a forma **se** parece estar fossilizada no radical verbal, comportando-se como um afixo, e, apesar de apresentar os traços de concordância, não tem função sintática e

5 semântica alguma, mas apenas está presente morfologicamente.

Tal ausência de função do pronome é confirmada pelo fato de que nesses tipos de estruturas com verbos pronominais o pronome **se** está, em alguns dialetos do português, em processo de desaparecimento, como é o caso de *Ele casou*, *Ele aposentou*, etc.

O que essa análise indica é que **Pro** é o último estágio da gramaticalização de **Ref**

10 no percurso (a), mas como o primeiro se formou mais tardiamente que **Est**, ainda está num estágio menos intenso de apagamento, enquanto **Est** já praticamente se extinguiu no português contemporâneo.

Mas como então surgiram os verbos pronominais? Provavelmente, do período arcaico para o moderno o **se** passou a ser utilizado em estrutura transitiva com objeto indireto,

15 ou seja, com preposição. Assim, como a nuance estilística de espontaneidade dos verbos de movimento, por exemplo, foi se desgastando com o tempo, o pronome **se** fossilizou-se nos radicais desses tipos de verbos, como no caso de *estender-se de*, *levantar-se de*, etc. depois se expandiu também com verbos psicológicos, como *lembrar-se de*, *esquecer-se de*, etc.

Considerando a idéia do ciclo da gramaticalização, como os verbos de movimento

20 foram os primeiros a se pronominalizar foram eles que se apagaram mais rapidamente. Note-se no corpus que preponderam verbos pronominais com leitura psicológica. Sem função, então, o pronome **se** no radical dos verbos pronominais tende no português contemporâneo a desaparecer.

Passemos à próxima seção que trata mais especificamente do estatuto gramatical

25 do grupo não-reflexivo.

5.5 O estatuto gramatical do grupo não-reflexivo

5

Com vistas a estabelecer o estatuto gramatical das estruturas do *se-indeterminador* e do *se-passivo*, apresento nesta seção uma descrição sintática delas a partir da análise de Raposo e Uriagereka (1999) para o português europeu, que se apóia no modelo gerativo.

10

A idéia fundamental do trabalho de Raposo e Uriagereka (op. cit.) é propor que o *se-indeterminador* é não-argumental [-arg], enquanto que o *se-passivo* é argumental [+arg]. Eles denominam *reflexive-passive*, o *se-passivo*, e de *impersonal-se* o *se-indeterminador*. O primeiro apresenta concordância do verbo com o objeto direto, enquanto o segundo não. Vejamos seus exemplos:

15

(37)

- a) Compraram-se muitas salsichas. (se-passivo)
- b) Compra-se muitas salsichas. (se-indeterminador)

20

Para os autores as duas construções se diferem quanto ao sentido: a) o *se-passivo*, [+arg], apresenta um sujeito indefinido (nos termos do autor, *quase existencial*); b) enquanto que *se-indeterminador*, [-arg], apresenta leitura genérica, (nos termos do autor, *quase universal*). Cinque (1988:546) apresenta para o italiano as diferentes propriedades das interpretações indefinida e genérica, isto é, das passivas e das indeterminadas. Observe-se a

25

tabela abaixo, adaptada de Cinque (op. cit.):

30

(38)

<i>Quase-existencial</i>	<i>Quase-universal</i>
<i>Interpretação</i>	<i>Interpretação</i>
a. Compatível com referência específica de tempo.	a. ' Incompatível com referência específica de tempo.
b. Incompatível com referência genérica.	b. ' Compatível com referência genérica.
c. Incompatível com contextos que dispensam especificidade de referência temporal.	c. ' Compatível com contextos que dispensam especificidade de referência temporal.
d. Compatível com a existência de um sujeito singular que satisfaça sua descrição.	d. ' Incompatível com a existência de um sujeito singular que satisfaça sua descrição.
Restrito a [NP, IP] θ -papel temático em estrutura-D.	Não restrito a [NP, IP] θ -papel temático em estrutura-D.
se-passivo	se-indeterminador

5

Segundo Raposo e Uriagereka²⁰, as duas interpretações são detectáveis no português europeu em:

(39)

- 10 a. Apenas se beberam todas as cervejas, saíram.
 b. ?? Nas convenções bebem-se sempre muitas cervejas
 c. Quando já se leram todas as revistas [num contexto genérico]
 d. Beberam-se muitas cervejas. Parece que foi o João.

15 (40)

- a. ' ?*Apenas se bebeu todas as cervejas, saíram.
 b. ' ?? Nas convenções bebe-se sempre muitas cervejas.
 c. ' Quando já se leu todas as revistas [num contexto genérico]
 d. ' *Bebe-se muitas cervejas. Parece que foi o João.

20

²⁰ (RAPOSO e URIAGEREKA: 1999, p.6)

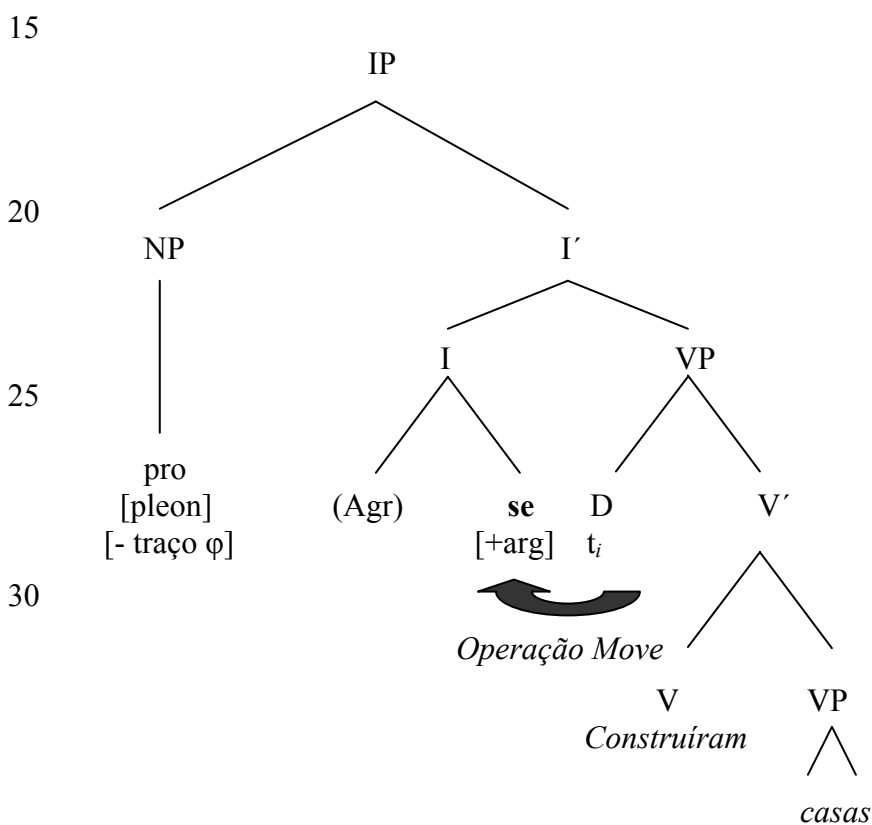
O que se observa nos dados acima é que os contextos das construções com concordância tende a apresentar uma leitura mais indefinida, enquanto que as construções sem concordância tendem a apresentar uma leitura mais genérica.

5 Em termos de confirmação da hipótese de gramaticalização para o percurso de *se*, o que seria relevante, como dito, é a verificação, dentro dos usos de *se*, qual deles seria mais gramatical. Para atestar então o grau de gramaticalidade do *se-apassivador* e do *se-indeterminador* do sujeito consideramos de grande importância a distinção [+/- arg], que, por sua vez, liga-se a propriedades interpretativas das construções em que aparecem conforme

10 o quadro acima.

Passemos então à caracterização sintática do *se-passivo* e do *se-indeterminador* de acordo com a proposta de Raposo e Uriagereka. Vejamos a primeira representação abaixo, que descreve a estrutura com o *se-passivo*, adaptado para o português:

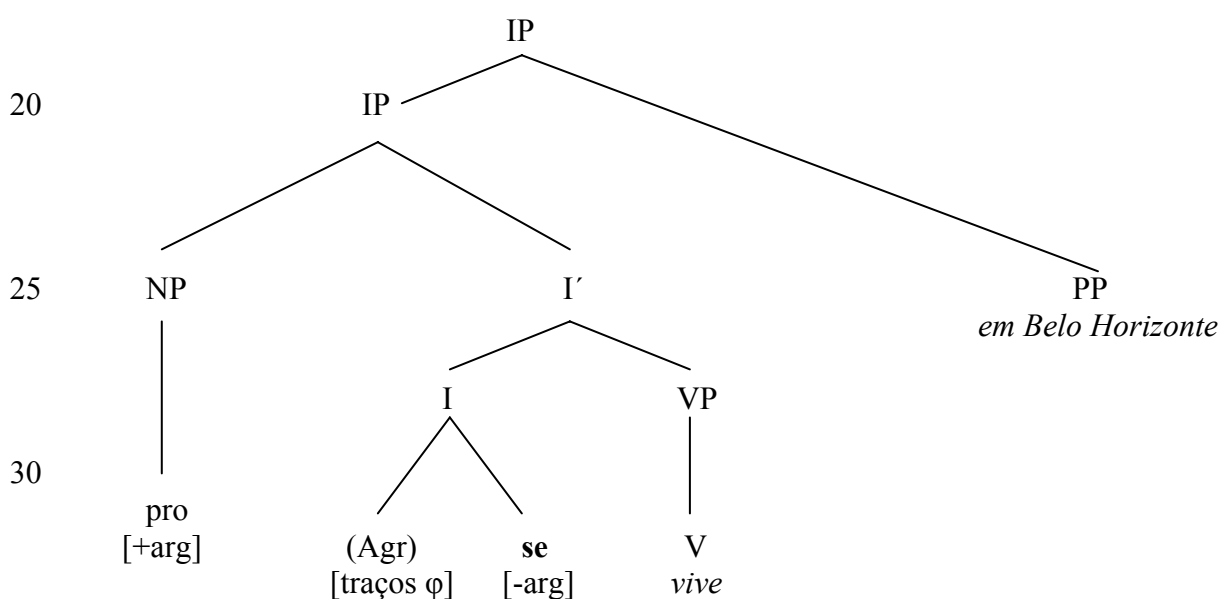
(41) *se-passivo: Construíram-se casas*



De acordo com a representação acima, o *se-passivo* é argumental, e recebe o papel temático externo na posição de sujeito interna ao VP; e recebe o caso acusativo por meio de uma categoria funcional que não especificaremos aqui. Como clítico, ele se move desta
 5 posição e adjunge à categoria I. E, apesar de se tratar de um D, faltam-lhe especificações para os traços semânticos de [pessoa], [gênero] e [número] já que não está co-indexado com um antecedente.

Passemos agora à análise do *se-indeterminador* do sujeito, que é também um D, e não dispõe, como no caso do *se* passivador, dos traços semânticos de [pessoa], [gênero] e
 10 [número]. De forma diferente ao *se-passivo*, o *se-indeterminador* não recebe papel temático, ou seja, é [-arg], o que significa que é inserido diretamente, por meio da operação *Juntar*, na posição de adjunção a I. Desse modo, o *se-indeterminador* do sujeito apresenta uma natureza de *afixo* e não de clítico, já que não surge em I por meio da operação *Mover*. É a seguinte representação:

15
 (42) *se-indeterminador: Vive-se bem em Belo Horizonte*



35

Estamos de acordo aqui com a concepção de que o pronome *se* não pode assumir o papel de argumento externo junto a verbos intransitivos por requerer a existência de acusativo, que não existe nessas construções, e o argumento dessas construções é a categoria vazia que ocupa a posição de especificador de IP.

Passemos agora à última seção em que comento as implicações das teorias apresentadas até agora a respeito do estatuto gramatical do grupo reflexivo e o não-reflexivo com o intuito de fomentar a hipótese de gramaticalização do percurso do pronome *se*.

10

5.6 Considerações finais

Nesta última seção, nosso objetivo é reafirmar a hipótese de gramaticalização para o percurso do pronome *se* a partir da análise teórica apresentada nas seções precedentes. O que se pode concluir dela é foi sintetizado na tabela que se encontra em 43 abaixo.

Considerando os dados dessa tabela, podemos concluir que: a) o *se*-reflexivo e o *se*-apassivador têm em comum o fato de serem clíticos e argumentais; b) o *se* -apassivador e o *se* reflexivo recebem caso acusativo; c) o *se*-pronominal, o *se*-estilístico, e o *se-indeterminador* têm comportamento de afixo, não sendo argumentais; d) o *se*-pronominal, o *se*-estilístico, e o *se-indeterminador* não recebem caso; e) o *se-indeterminador* tem em comum com o *se-apassivador* o fato de não dispor dos traços semânticos de [pessoa], [gênero] e [número]. Por outro lado, todos os tipos de *se* compartilham os traços formais de 3ª pessoa do singular.

O que se observa por meio das análises e das informações apresentadas é que há um *crescente* de “gramaticalidade” dos itens quando se consideram os tipos de *se*. Ora, na literatura sobre gramaticalização, prevê-se que ocorre um gradativo “esvaziamento semântico” e “redução de forma ou fônica” nos itens que sofrem esse tipo de processo, isso é

o que parece ocorrer comparando-se, por exemplo, os traços do tipo **Ref**, **Pas**, e **Ind**. Como **Ref** é a origem e **Pas** o tipo inicial do percurso não-reflexivo de *se*, espera-se que eles apresentem o maior número de traços, e isso é o que ocorre, enquanto que *Est*, *Pro* e *Ind*, apresentam maior apagamento de traços semânticos e sintáticos, comparando-se aos demais.

5 Nessa visão, **Ind** é a forma mais gramaticalizada.

(43)

Tipo \ Aspecto	[gênero]*	[pessoa]*	[número]*	[arg]	[caso]	[presença de agente]	[referência de tempo específica]	Clítico	Afixo
Ref	+	+	+	+	+	+	+	+	-
Est	+	+	+	-	-	+	+	-	+
Pro	+	+	+	-	-	+	+	-	+
Pas	-	-	-	+	+	-	+	+	-
Ind	-	-	-	-	-	-	-	-	+

* [gênero], [pessoa], e [número] são traços semânticos.

10 Considerando que para Kurylowicz a gramaticalização pode ser definida como: *a ampliação dos limites de um morfema, cujo estatuto gramatical avança do léxico para a gramática, ou de um nível menos gramatical para mais gramatical, isto é, de formante derivativo para formante flexional* (op. cit), o percurso do pronome *se* na história constitui-se como um caso de gramaticalização, já que do tipo **Ref** para o **Ind** observa-se um crescente grau de gramaticalização. Vejamos a representação a partir do ciclo da gramaticalização de Hopper e Traugott (1993):

(44) 1°. Item lexical > 2°. item gramatical > 3°.clítico > 4°.afixo.

20

(45) 2°. Raiz gramatical *se- reflexivo *-se >*

→ (a) ⇨ > 4°. (afixo *se-estilístico e se-pronominal*)

25

→ (b) ⇨ > 3°. (clítico *se- passivo*) > 4°. (afixo. *se-indeterminador*)

O ciclo (44) supõe que os processos de gramaticalização ocorrem de forma linear, embora isso não ocorra se observarmos a configuração de (45a-b), que descreve os ciclos de *se*. Em Vitral (1999), lê-se que essa linearidade não é adequada, devido ao fato de que os dois primeiros estágios estão relacionados ao conteúdo dos itens, enquanto os dois últimos se referem à forma deles. Desse modo, é possível haver dois itens que se gramaticalizam, distinguindo-se pela forma e não pelo conteúdo. Por exemplo, os itens *não* e *num* são ambos itens gramaticais, isto é, NEG, mas somente *num* se comporta como um clítico.

De forma semelhante a *não* e *num*, acontece com as ocorrências de *se* do percurso (b) de (45): são itens gramaticais, isto é, D, mas o *se-reflexivo* e o apassivador são clíticos, enquanto o *se* indeterminador é um afixo. No plano do conteúdo, como vimos, os dois primeiros são argumentais e o último é não-argumental e, além disso, a caracterização dos traços semânticos nos três casos é diferente: o primeiro dispõe de especificações para os traços de [pessoa], [número] e [gênero], mas não os dois últimos. Assim, aliadas aos índices de frequência apresentados na seção anterior, essas propriedades caracterizam a expansão de *se* como um processo de gramaticalização.

Considerando os dados de frequência, nota-se, como dito, o aumento da frequência do grupo não-reflexivo em detrimento da frequência dos tipos do grupo reflexivo: os primeiros apareceram em 48,1% das vezes no período arcaico, em 61% das vezes no período moderno e em 77,38% no período contemporâneo. Por outro lado, a frequência dos tipos do grupo reflexivo diminuiu: obtivemos 51,9% no período arcaico, 39% no moderno e 22,62% no contemporâneo. O aumento da frequência ao longo do tempo dos tipos mais gramaticais e o decréscimo das menos gramaticais é um fato relevante na caracterização do processo de expansão de *se* como um processo de gramaticalização.

Em relação à frequência, os dados também confirmam a nossa hipótese de gramaticalização: espera-se ao decorrer do tempo que a frequência se expanda, e que depois decresça, numa distribuição parabólica, e isso é o que acontece, ou seja, aumenta o valor do período arcaico para o moderno, no primeiro 4,88% > 6,9%, e no grupo não-reflexivo 4,52% > 10,23%, mas esses valores decrescem do moderno para o contemporâneo, no primeiro 6,9% > 0,92%, e no grupo não-reflexivo 10,23% > 3,15%.

No intuito de sintetizar as informações acima, elaborei um esquema explicativo em (46) no qual elas podem ser mais bem visualizadas.

Como se pode perceber pelo gráfico, nele estão apresentadas as informações que sintetizam todo o trabalho, e mais especificamente aquelas referentes ao capítulo em questão, que objetiva caracterizar o percurso histórico de *se* na língua portuguesa como um processo de gramaticalização.

Primeiramente, destaca-se no esquema a bipartição do percurso diacrônico do pronome *se* se desenvolvendo em duas linhas gerais: o grupo reflexivo e o grupo não-reflexivo, como já apresentado em (44) e (45). Observe-se também que os tipos se distribuem em termos de relevância da frequência por período, ou seja, a frequência de **Ref** e **Est** no grupo reflexivo é mais alta no período arcaico, mas no período moderno a frequência de **Pro** aumenta, como indica a análise dos dados. Em relação ao grupo não-reflexivo, note-se que a frequência do tipo **Pas** é predominante no período moderno, e o do tipo **Ind** o é no período contemporâneo.

Em relação à curva em forma de parábola, ela indica que a frequência do pronome *se* aumenta muito do período arcaico para o moderno, mas decresce bruscamente do moderno para o contemporâneo.

Na parte superior do gráfico é apresentada a relação entre o período, o estatuto gramatical com maior índice de frequência e o seu nível de gramaticalidade. O que se pode

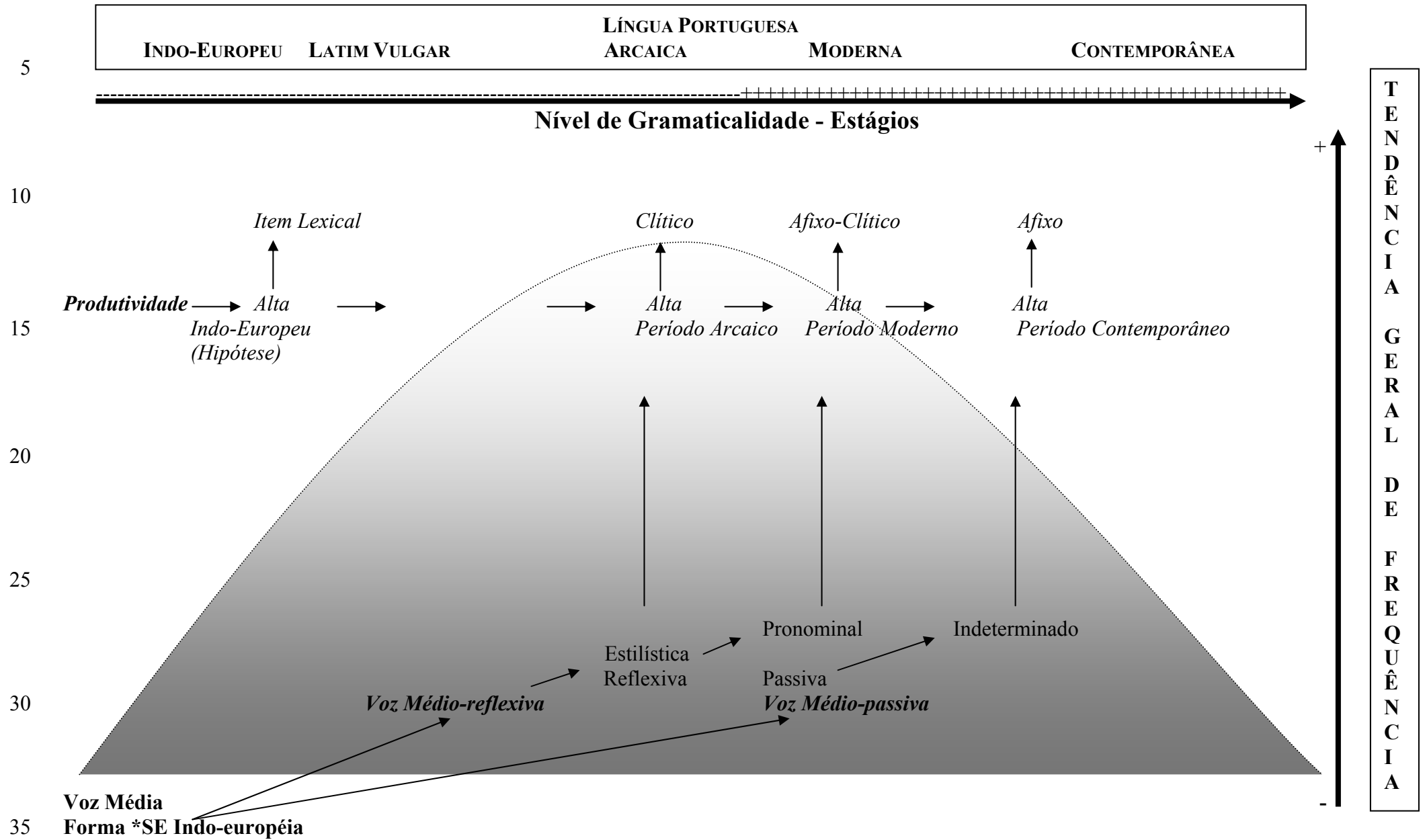
constatar é que os tipos se organizam num *continuum* que vai do menos para o mais gramatical, respectivamente, da esquerda para a direita. Note-se que no período arcaico prevalece a frequência mais alta da forma menos gramatical de *se*, que funciona como um pronome. Já no contemporâneo, o pronome *se* tem sua frequência mais alta em orações em

5 que ele funciona como um afixo, que é mais gramatical. O período moderno é a fase intermediária de mudança geral de tendência clítico > afixo na frequência global dos tipos de **se**.

Considerando então todas as informações apresentadas no gráfico e nos capítulos precedentes, elas nos indicam que o pronome *se* passou por um processo de

10 gramaticalização ao longo da história da língua portuguesa.

ESQUEMA EXPLICATIVO DIACRÔNICO DO PRONOME SE



BIBLIOGRAFIA

1. **AMARAL**, Eduardo Tadeu Roque. *A ausência/presença de artigo definido diante de antropônimos em três localidades de Minas Gerais: Campanha, Minas Novas e Paracatu*. Dissertação (Mestrado em Estudos Lingüísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, 2003.
2. **BECHARA**, E. *Moderna Gramática Portuguesa*. Rio de Janeiro: Lucerna, 1999.
3. **BYBEE**, J. Perkins et Pagliuca. *The evolution of grammar*. Chicago: The University of Chicago Press, 1994.
4. **CÂMARA JR.**, J. Mattoso. *Princípios de Lingüística Geral*. 3. ed. Rio de Janeiro: Acadêmica, 1959.
5. **CÂMARA JR.**, J. Mattoso. *Dicionário de Filologia e Gramática*. Rio de Janeiro: Globo, 1959.
6. **CASTILHO**, Ataliba. A Gramaticalização. *Estudos Lingüísticos e Literários* n°. 19. Salvador: Universidade Federal da Bahia, 25-64 março 1997.
7. **CASTILHO**, Ataliba. Língua Falada e Gramaticalização. *Filologia e Lingüística Portuguesa* n°. 1 São Paulo: Universidade de São Paulo, 107-120, 1997.
8. **CASTILHO**, Ataliba. Para uma sintaxe da repetição: Língua Falada e Gramaticalização. Em *Língua e Literatura*, n°. 23. São Paulo: Universidade de São Paulo, 293-330, 1997.
9. **CAVALCANTE**. S. R. de O. *A Indeterminação do Sujeito na Escrita Padrão: A Imprensa Carioca nos Séculos XIX e XX*. Dissertação de Mestrado. Faculdade de Letras – UFRJ, 1999.
10. **CHOMSKY**, N. *Lectures on Government and Binding*, Dordrecht: Foris Publications, 1986.
11. **CHOMSKY**, N. *The Minimalist Program*. Cambridge: The MIT Press, 1995.
12. **CINQUE**, G. *On Si Constructions and the Theory of Arb*, *Linguistic Inquiry*, 19, 521-581, 1988.
13. **COHEN**, Maria Antonieta A. M. et al. Filologia Bandeirante. *Filologia e Lingüística Portuguesa*. São Paulo, n. 1, p. 79-94, 1997.
14. **COHEN**, M. A. M. A. Gramaticalização e reanálise na Língua Portuguesa. *Caligrama*. 1, pp. 543-52, Belo Horizonte: 1988.
15. **CROFT**, William. *Explaining Language Change: an evolutionary approach*. Pearson Education Limited: Londres, 2000.

16. **CUNHA, C. e CINTRA, L. F. L.** *Nova Gramática do Português Contemporâneo*. Rio de Janeiro: Nova Fronteira, 2001.
17. **DUBOIS**, Jean et alii. *Dictionnaire de Linguistique*. Paris: Librairie Larousse, 1973.
18. **HAWAD, H. F.** *Funções Discursivas da Voz Passiva e da Indeterminação do Sujeito*. Monografia. Inédito, 1999.
19. **HAWAD, H. F.** *Tema, Sujeito e Agente: A voz passiva portuguesa em perspectiva sistêmico-funcional*. PUC-RJ, Rio de Janeiro. Tese. Inédita, 2002.
20. **HEINE, B. B, HÜNNEMEYER & U. CLAUDI**, *Grammaticalization: a conceptual framework*. Chicago: The University of Chicago Press, 1991.
21. **HUBER, J.** *Gramática do Português Antigo*. Trad.. M.M.G. Delille. Lisboa: Fundação Caloust Gulbenkian, 1986
22. **KAYNE, R.** *Connectedness and Binary Branching*. Dordrecht: Foris, 1984.
23. **KROCH, A.** Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change*, vol 1, pp.199-244, 1989.
24. **KURYLOWICZ, J.** *The Inflectional Categories of Indo-European*. Heidelberg: Carl Winter, 1964.
25. **LIGHTFOOT, D.** *Principles of Diachronic Syntax*. Cambridge: MIT Press, 1979.
26. **LIGHTFOOT, D.** *The Development of Language*. Oxford: Blackwell, 1999.
27. **KURY, A. da G.** *Novas Lições de Análise Sintática*. São Paulo: Ática, 1986.
28. **LABOV, W.** *Principles of linguistic Change: internal factors*. Oxford: Blackwell, 1994.
29. **LEHMANN, A.** Grammaticalization: synchronic variation and diachronic change. *Lingua e Stile*, nº 20, p. 303-318. (1982a)
30. **LEHMANN, A.** Thoughts on Grammaticalization: a programatic sketch. *Arbeiten des Kölner Universalien-Projekts*, Köln vol I, 1982.
31. **LUFT, Celso Pedro.** *Dicionário de regência verbal*. Rio de Janeiro: Globo, 1987.
32. **MAROUZEAU, J.** *Lexique de la terminologie linguistique*. Paris: Librairie Orientaliste Paul Geuthner, [s.d.].
33. **MAURER JR, Theodoro H.** *A Propósito da Evolução Semântica do Pronome se em Português*. São Paulo: USP, 1951.

34. **MEGALE**, Heitor. *Filologia bandeirante: estudos*. vol.1. São Paulo: *Humanitas* FFLCH/USP, 2000.
35. **MEILLET**, A. *Linguistique historique et linguistique générale*. Paris: A. H. Champion, 1948.
36. **MEILLET**, A. *Introduction a l'etude comparative des langues indo-europeennes*. Paris: Hachette, 1949.
37. **MEILLET**, A. *Esquisse d'une histoire de la langue latine*. Paris: Hachette, 1948.
38. **MARTELOTTA**, S. VOTRE; CEZARIO M. M. (Org.). *Gramaticalização no Português do Brasil: uma abordagem funcional*. Rio de Janeiro: Tempo Brasileiro, 1996.
39. **MARY**, A. K.; **ROBERTS**, I. (Org.). *Português Brasileiro: uma viagem diacrônica*. 2 ed. Campinas, SP: Editora da Unicamp, 1996.
40. **NARO**, A. *History of Passive and Impersonals*. [S.l.]: Massachussets Institut of Tecnology, 1968.
41. **NARO**, A. The Genesis of the reflexive impersonal in Portuguese: a study in syntactic change as a surface phenomenon. *Language*, v. 52, n. 4, 1976.
42. **MATTOS e SILVA**, Rosa Virgínia. *Estruturas Trecentistas*. Imprensa Nacional, 1989.
43. **NEVES**, M. H. de M. *Gramática de Usos do Português*. São Paulo: UNESP, 2000.
44. **NUNES**, Jairo. *O famigerado se: uma análise diacrônica das construções com se apassivador e indeterminador*. 1990. Dissertação (Mestrado em Lingüística) Unicamp, Campinas.
45. **PAGOTTO**, E. Clítico, mudança e seleção natural, In: **ROBERTS**, I; **KATO**, M. (Org.), *Português Brasileiro. Uma Viagem Diacrônica*, Campinas: Editora da UNICAMP, 1993.
46. **PIMENTA-BUENO**, M. do N. S. A Proposal for a unified treatment of reflexive, reciprocal, intrinsic and impersonal 'se' in Portuguese. *Colloquium on Spanish and Luso-Brazilian Linguistics*. Washington: Georgetown University Press, 1979.
47. **POKORNY**, J. *Indogermanisches Etymologisches Wörterbuch, I Band*. Munique: Francke Verlag Bern, 1969.
48. **PONTES**, E. *Verbos Auxiliares em Português*. Petrópolis: Vozes, 1973.
49. **RAPOSO**, Eduardo Paiva. *Teoria da Gramática: a faculdade da linguagem*. Lisboa: Editorial Caminho SA, 1992.

50. **RAPOSO**, E; **URIAGEREKA**, J. *Indefinite Se*. [S.l]. University of Califórnia/University of Maryland, 1999
51. **ROBERTS**, Ian; **ROUSSEAU**, Ana. *Syntactic Change: A Minimalist Approach to Grammaticalisation*. [S.l], [S.d].
52. **ROMANELLI**, R. L. *O Supletivismo Indo-Europeu na morfologia latina*. Belo Horizonte: Imprensa Universitária, 1975.
53. **SHIBATANI**, M. *Passives and Related Constructions: A Prototype Analysis*. *Language* vol. 61. pp. 821-848, 1985.
54. **VENDRYES**, Joseph. *Le langage, introduction linguistique à l'histoire*. Paris: Éditions Albin Michel, 1968.
55. **VOSSLER**, Karl. *Filosofia del lenguaje: ensayos*. Trad. Amado Alonso y Raimundo Lida, Buenos Ayres: Editorial Losada, AS, 1968.
56. **VIANNA**, Humberto. *A estrutura Modal+Infinitivo em português gramaticalização e modalização*. 2000. Dissertação (Mestrado em Lingüística), UFMG, Belo Horizonte.
57. **VITRAL**, Lorenzo. *Sintaxe formal e gramaticalização: roteiro de uma pesquisa*. (Palestra pronunciada no IEL-Unicamp), [s.l], 2000.
58. **VITRAL**, Lorenzo & Ramos. Gramaticalização de “você”: um processo de perda de informação semântica?. *Filologia e Lingüística Portuguesa*, v.3, p.55-64, Humanitas Publicações, São Paulo: 1999.
59. **VITRAL**, Lorenzo. A Forma Cê e a Noção de Gramaticalização. *Revista de Estudos da Linguagem*, 1965.
60. **VITRAL**, Lorenzo. A negação: teoria da Checagem e mudança lingüística. *D.E.L.T.A.* 1999, vol. 15,1,pp.57-84.
61. **VITRAL**, Lorenzo. A Interpolação de Se e suas Conseqüências para a Teoria da Cliticização., *Revista da ABRALIN*, 2002, v.1,nº 2, p.161-197.
62. **VITRAL**, Lorenzo; **CIRÍACO**, L.; **REIS**, C. Intensidade e Duração de Formas Reduzidas do Português Brasileiro. *Revista de Estudos da Linguagem*, 2004, v.12, nº2, p.143-157.
63. **VITRAL**, Lorenzo; **RAMOS**, J. *Gramaticalização: uma abordagem formal*. A sair pela Tempo Brasileiro, Rio de Janeiro.
64. **VITRAL**, Lorenzo. O Papel da Frequência na Identificação de Processos de Gramaticalização. A sair pela *Revista Scripta*. Belo Horizonte: PUC-MG, 2006.
65. **VITRAL**, Lorenzo. *A Evolução do SE reflexivo em Português na Perspectiva da Gramaticalização*. A sair nos Anais do VI encontro do PHPB, Salvador.