

**LORENZO VITRAL
JÂNIA RAMOS**

GRAMATICALIZAÇÃO
uma abordagem formal

FALE
FACULDADE DE LETRAS
FALE

tb
tempo brasileiro

**GRAMATICALIZAÇÃO:
uma abordagem formal**

CIP Brasil. Catalogação-na-Fonte
Sindicato Nacional dos Editores de Livros, RJ

V829g Vitral, Lorenzo

Gramaticalização: uma abordagem formal / Lorenzo Vitral, Jânia Ramos. – Rio de Janeiro: Tempo Brasileiro; Belo Horizonte, MG: Faculdade de Letras FALE/UFMG, 2006.

Inclui bibliografia
ISBN 85-282-0138-4

1. Língua portuguesa – Sintaxe. 2. Língua portuguesa – Gramática. 3. Língua portuguesa – Gramática gerativa. I. Ramos, Jânia M. (Jânia Marins). II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

06-3467

CDD 469.5
CDU 811.134.3'36

LORENZO VITRAL
JÂNIA RAMOS

GRAMATICALIZAÇÃO:
uma abordagem formal



© 2005 *Lorenzo Vitral e Jânia Ramos*

Revisão

LILIAN DE OLIVEIRA

VANESSA BATISTA DE OLIVEIRA

Diagramação e Capa

JUNIA CAMARINHA DA SILVA

Todos os direitos reservados e protegidos.

Proibida a duplicação ou reprodução desta obra ou partes da mesma,
sob quaisquer meios, sem autorização expressa dos editores.

Faculdade de Letras FALE/UFMG

Av. Antônio Carlos, 6627 – Pampulha

CEP: 31.270-901 – Belo Horizonte – MG

Fone: (31) 3499-5106 – Fax: (31) 3499-5120

EDIÇÕES TEMPO BRASILEIRO LTDA.

Rua Gago Coutinho, 61 – Laranjeiras

CEP: 22.221-070 – CP 16.099

Telefax: (21) 2205-5949

Rio de Janeiro – RJ – Brasil

e-mail: tb@tempobrasileiro.com.br

2006

LISTA DE ABREVIACÕES

C: Complementizer

*coord.: Coordenação

CP: Complementizer Phrase

D: Determiner

DP: Determiner Phrase

E1: A letra "E" indica entrevista e o número ao lado, a localiza no *corpus*.

I: Inflection

INFL: Inflection

IP: Inflectional Phrase

LF: Logical Form

N: Noun

Neg: Negation

NegP: Negative Phrase

NP: Nominal Phrase

OV: Objeto-Verbo

P: Preposition

PB: Português brasileiro

PF: Phonetic Form

Q: Quantifier

SC: Small Clause

SS: S-Structure

T: Tense

TP: Tense Phrase

V: Verb

V2: Verb Second

VP: Verbal Phrase

X: Núcleo de um sintagma

XP: X Phrase

ÍNDICE

APRESENTAÇÃO	9
INTRODUÇÃO	13
CAPÍTULO 1	
Definições, pressupostos e proposta	19
CAPÍTULO 2	
A forma CÊ e a noção de gramaticalização	29
CAPÍTULO 3	
Gramaticalização em processo	39
CAPÍTULO 4	
Identificando Clíticos: evidências fonéticas	53
CAPÍTULO 5	
Mais um pronome em processo de cliticização: o par eles/es	71
CAPÍTULO 6	
Gramaticalização de “você”: um caso de perda de conteúdo semântico	81
CAPÍTULO 7	
A interpolação de SE e suas conseqüências para a teoria da cliticização	89
CAPÍTULO 8	
O processo <i>não</i> > <i>num</i> na fala	119

CAPÍTULO 9	
A negação: teoria da checagem e mudança lingüística	133
CAPÍTULO 10	
O percurso de <i>senhor</i>	163
REFERÊNCIAS BIBLIOGRÁFICAS	173

APRESENTAÇÃO

Este livro é um registro de pesquisas realizadas na área de Linguística Teórica, desenvolvidas na Universidade Federal de Minas Gerais, a partir de 1996. Alguns dos textos que o compõem já foram publicados na forma de artigos em revistas nacionais; outros são trabalhos inéditos que só agora, portanto, encontram divulgação. Nosso trabalho recebeu apoio do Conselho de Desenvolvimento Científico e Tecnológico (CNPq), que o financiou na condição de projeto integrado, a partir do ano de 1998. Pudemos contar também com os resultados obtidos pelas monografias, dissertações e teses elaboradas por nossos orientandos que nesses anos tiveram participação neste trabalho enquanto alunos do Programa de Pós-graduação em Estudos Linguísticos da Faculdade de Letras da UFMG. Um deles, Lucas Corrêa, é co-autor do Capítulo 5.

Como é sabido, o tema *Gramaticalização* tem sido retomado nesses últimos anos, no Brasil e no exterior, permitindo a abordagem de diversos fenômenos de linguagem numa perspectiva renovadora e estimulante. Sabe-se, além disso, que a perspectiva teórica desses novos trabalhos é, comumente, classificada como de orientação funcionalista. Na literatura recente, encontra-se igualmente muito fecunda a discussão acerca da questão de saber se a gramaticalização constitui, epistemologicamente, um modelo teórico de pleno direito ou se, *a contrário*, os chamados *Processos de*

Gramaticalização são, tão-somente, epifenômenos que poderiam ser deduzidos de enunciados oriundos de outros quadros teóricos. A questão é, evidentemente, complexa e envolve tomada de decisões, que, seguramente, ultrapassam o campo da Linguística, estritamente considerada.

Nossa posição nesse debate pode ser assim sintetizada, com base nos resultados obtidos na pesquisa realizada: a análise explícita desses fenômenos pode se valer de enunciados abstratos produzidos no interior de outros quadros teóricos. Tal posicionamento justifica, em parte, o título do livro. A abordagem formal da gramaticalização constitui, assim, uma proposta de como capturar processos de gramaticalização, cuja identificação e explicação pode se valer de princípios descobertos na teoria da gramática, e seu percurso histórico pode ser capturado através de análise quantitativa. O apelo às propostas geradas em outros quadros teóricos não é trivial e tem, evidentemente, conseqüências teóricas, ou mesmo percalços, que devem ser abordados com cuidado. Não se pretende aqui fazer uma mistura de “modelos de análise”, que, como é sabido, possuem objetos de estudo próprios, ou usá-los “à la carte” de acordo com necessidades empíricas com as quais deparamos. Nossa pretensão é construir um objeto de estudo distinto, de maneira a levar em conta as contribuições das perspectivas teóricas mencionadas. Nosso objetivo central não é, portanto, tomar parte do desenvolvimento hodierno dessas teorias, mas tratar explicitamente múltiplos fenômenos envolvidos nos processos de gramaticalização.

Por enquanto, apenas demos início à tarefa de construção do nosso objeto de estudo. Por esta razão, o livro ora publicado apresenta reformulações, o que se deve ao fato de refletir a evolução de nosso trabalho no contexto dos últimos seis anos. Por outro lado, a não homogeneidade do livro, embora possa ser vista como uma deficiência, atesta o vigor da perspectiva esboçada e serve também como um convite para que colegas, com suas bem-vindas

críticas e sugestões, tomem parte da discussão e nos permitam amadurecer nosso trabalho.

Agradecemos, enfim, ao prof. Ataliba Castilho, cuja contribuição científica foi primordial para a renovação dos estudos da gramaticalização no Brasil.

OS AUTORES

INTRODUÇÃO

A noção de gramaticalização foi introduzida por Meillet (1948:131) para designar um certo tipo de fenômeno lingüístico de natureza diacrônica. Trata-se da transição gradual de “palavras principais” para “palavras acessórias” e, enfim, para “palavras gramaticais” em estágios de uma língua. Itens pertencentes às categorias “de conteúdo lexical”, como verbos e adjetivos, passam a fazer parte de categorias “vazias de conteúdo lexical”, como auxiliares e certas preposições; e, em seguida, transformam-se em clíticos e afixos, antes de desaparecerem completamente.

Um exemplo bastante recorrente na bibliografia sobre o tema é o verbo inglês *to go*. Como é sabido, há hoje um verbo *to go* de moção no espaço e outro que funciona como auxiliar. Sabe-se, além disso, que, na história da língua inglesa, as ocorrências desse verbo de moção no espaço são anteriores ao seu emprego como auxiliar. Supõe-se assim que houve gramaticalização, isto é, um verbo de moção no espaço passou a indicar moção no tempo, mais exatamente o futuro.

- (1) a. moção no espaço > moção no tempo
- b. verbo > auxiliar

Nos últimos anos, a retomada dos estudos acerca da gramaticalização tem sido desenvolvida a partir de dois eixos principais.

Um deles, estritamente ligado às idéias originais de Humboldt, enfatiza o papel da evolução das categorias da língua, destacando processos diacrônicos que, no âmbito da oração, reanalisam itens lexicais, transformando-os em itens gramaticais. Um marco deste eixo foi o trabalho de Lehmann (1982), e um caso típico é o apresentado em (1b), acima.

O outro eixo põe em destaque a interação entre sintaxe e discurso. Nesta perspectiva, os padrões sintáticos resultam do próprio uso e, por isso, o componente gramatical de qualquer língua não seria autônomo. Em outras palavras, as exigências comunicativas, geradas no contexto do discurso, determinariam as regularidades observadas no nível da oração. Um marco desse eixo são os trabalhos de Sankoff e Brown (1976) e Givón (1979). Um desdobramento é focalizar o processo cognitivo representado em (1a).

Estas propostas abriram uma nova janela para a pesquisa na área de gramaticalização, podendo ser vistas como um esboço de um novo paradigma científico em Linguística. Vários pesquisadores estão envolvidos nesse empreendimento. No Brasil, podem-se citar Gonçalves (1987); Votre (1993); Castilho (1995, 1997); Martelotta, Votre e Cezario (1996); Bittencourt (1999); Naro e Braga (2000); dentre outros. No exterior, até congressos específicos, como o New Reflections on Grammaticalization, realizado em Am Neuen Palais, em junho de 1999, podem ser citados. Números especiais de periódicos também têm sido publicados – como o da revista *Linguistic Science*, 2000 –, dedicados exclusivamente a esse tema, o que mostra sua relevância no quadro da Linguística. Numa palestra proferida em Freiburg, Hopper (2001:1) apresenta um trabalho em que aponta os três tópicos mais atuais sobre gramaticalização: (i) se a unidirecionalidade possui significância teórica; (ii) como pensar sobre mudanças específicas, conhecidas como gramaticalização, em situações de contato; e (iii) como conciliar a pesquisa sobre gramaticalização com a Linguística

Formal, uma vez que foi no âmbito do funcionalismo que a gramaticalização teve seu crescimento. Os tópicos (i) e (iii) foram selecionados como objeto de estudo no presente volume.

São várias as descrições dos processos de gramaticalização encontradas na literatura. Uma breve síntese aparece abaixo, construída a partir de Lehmann (1982); Hopper (1991); Bybee *et al.* (1994):

- (2) a. As “mudanças” seguem as etapas predeterminadas do ciclo da gramaticalização;
- b. O ciclo é unidirecional e universal;
- c. As “mudanças” decorrem de redução fonética e redução semântica dos itens;
- d. As “mudanças” são graduais;
- e. Os processos de gramaticalização implicam uma revisão da distinção sincronia/diacronia.

As proposições (a) e (b) assumem serem os processos de gramaticalização essencialmente diacrônicos e previsíveis, ou seja, as “mudanças” não são aleatórias, mas tendem, em qualquer língua, a trilhar estágios previstos. Os estágios seriam (3), segundo Hopper e Traugott (1993):

- (3) item lexical > item gramatical > clítico > afixo

Esta hipótese é bastante forte e grande parte da resistência a ela deve-se a seu caráter teleológico. Decorre dela ainda a predição de que não há retorno de um item a uma etapa da gramaticalização já percorrida, isto é, a mudança é unidirecional.

De acordo com a proposição (c), descrição e explicação das “perdas” de forma e conteúdo não são geradas no interior de uma “teoria”. Os autores que trabalham com essa noção buscam, em

outros quadros teóricos, os instrumentos para tratar das reduções fonética e semântica. Sendo assim, permanece aberta a questão de saber que hipóteses teóricas são as mais adequadas.

A proposição (d), que diz respeito à natureza gradual das “mudanças”, assume ser a categorização sintática não discreta e instável, sujeita, portanto, à atuação, consciente ou não, dos falantes num sentido amplo.

A proposição (e) é, na nossa visão, conseqüência da dificuldade teórica de lidar com a proposição (d) e suas conseqüências, isto é, a suposta gradualidade dos processos de gramaticalização vai de encontro à definição da noção de “estado de língua”, que é crucial na caracterização da sincronia. A atribuição de gradualidade parece-nos vir do fato de ser possível, num recorte sincrônico, atribuir a um item que se gramaticaliza duas categorizações diferentes e afêrir índices de freqüência diferentes para cada um deles. Tal linha de análise será desenvolvida neste livro e pode dar conta, igualmente, de explicar dois outros aspectos dos processos de gramaticalização. Em primeiro lugar, a coexistência, num dado estágio de uma língua, de um item com função lexical e com a função gramatical derivada. Este fato, como se verá adiante, é que nos faz repensar a adequação do termo ‘mudança’ para descrever os processos de gramaticalização. Em segundo lugar, o fato de o item que se gramaticaliza aumentar sua freqüência.

Os problemas apontados acima, a nosso ver, permitem supor que não haveria ainda um “modelo teórico” que explique os processos de gramaticalização. Isto é, um modelo que faça deduzir de um conjunto interligado de princípios as características observáveis dos processos de gramaticalização. Nem parece haver ainda instrumentos teóricos específicos capazes de descrever, de forma explícita, esses processos. Este livro é uma tentativa de dar alguns passos para a elaboração de um esboço de uma teoria da gramaticalização.

Tomaremos como objeto de estudo epifenômenos do português brasileiro, como os pronomes *você* e *eles*, a forma de tratamento ‘senhor’ e a partícula de negação *não*. Esses itens apresentam, no momento atual, pelo menos dois tipos de realizações fonéticas, que poderiam ser descritas respectivamente, *grosso modo*, como foneticamente plenas e foneticamente reduzidas. Veja-se o quadro a seguir:

Forma foneticamente plena	Forma foneticamente reduzida
Você	Ocê, Ce
Eles	Es, Eis
Não	Num
Senhor	Sô

Do ponto de vista semântico, há também diferenças que poderiam ser descritas como “certo esvaziamento semântico”.

É nosso propósito, no decorrer deste volume, discutir as propriedades acima como etapas de um processo de gramaticalização. É importante ressaltar que se trata de um processo ainda em curso, conforme veremos nos próximos capítulos.

CAPÍTULO I

DEFINIÇÕES, PRESSUPOSTOS E PROPOSTA

Alguns pontos de reflexão no intuito de mostrar que os fenômenos de gramaticalização podem receber uma abordagem formal são elaborados neste capítulo. Embora ainda não se disponha de um quadro teórico homogêneo para tratar formalmente a gramaticalização, conforme assinalam Hopper (2001) e Newmeyer (2000), formulamos algumas diretrizes para uma abordagem dessa natureza, que são, na verdade, reflexões acerca de propriedades atribuídas aos fenômenos de gramaticalização, normalmente encontradas na literatura dessa área. O que fizemos então foi “interpretar” propriedades da gramaticalização a partir de pressupostos de um modelo formal que, no nosso caso, é o da gramática gerativa, em sua fase mais atual, o Programa Minimalista.

Nos processos de gramaticalização, afirma-se que as alterações da natureza dos itens incidem sobre *três aspectos gramaticais*, que tornam os fenômenos de gramaticalização distintos ou, ontologicamente, identificáveis. Assim, quando um item se gramaticaliza, sofre alterações sintáticas, muda de classe de palavra; semânticas: “esvaziasse” semanticamente ou “perde conteúdo”; e morfofonéticas: ocorre “redução” ou “diminuição” de sílabas, sons e/ou acento.

Como se sabe, essas alterações são previstas por meio de um ciclo que é formulado da seguinte maneira, de acordo com Hopper e Traugott (1993: 7):

- (1) a. item lexical > b. item gramatical > c. clítico >
d. afixo

A alteração da etapa (a) para a (b) diz respeito a uma alteração da classe de palavra do item e é acompanhada por uma alteração relacionada com o significado, ou seja, tornar-se um item gramatical traz certas conseqüências do ponto de vista do conteúdo do item. As etapas (c) e (d) prevêm alterações do item do ponto de vista morfofonético. Cada etapa levaria, como corolário, a modificações da distribuição sintática do item, tornando-a mais restrita, à medida que caminha para a direita no ciclo. Essa distribuição sintática faz prever, pelo menos à primeira vista, que haveria gradualidade entre as etapas do ciclo, em relação ao significado. Isto é, quanto mais à direita, maior o esvaziamento semântico. Outra expectativa seria a de que os processos podem tornar-se concomitantes: quanto maior o esvaziamento semântico, maior a perda de substância fônica. Estas expectativas, explicitadas em Lehmann (1982), entretanto, não se confirmam, conforme veremos nas duas próximas seções, que tratam, respectivamente, da *gradualidade* e da *concomitância*.

1. GRADUALIDADE

É possível um item apresentar estatutos sintáticos distintos, tal como ocorre com a forma verbal *pode* no exemplo abaixo:

- (2) Ele pode comprar um carro novo.

De acordo com a extensa literatura sobre os verbos modais (cf., por exemplo, Pontes (1973), Perini (1977), Lightfoot (1979), Vitral (1987), Vianna (2000)), a oração acima pode ser analisada como um período composto – e neste caso, o item *pode* é analisado como um verbo lexical – ou como um período simples e neste caso, *pode* deverá ser analisado como um tipo de verbo auxiliar. Assim, cabe fazer a

hipótese de que o item *poder* está se gramaticalizando, isto é, passou de verbo a auxiliar. Como este processo atingiu, no máximo as etapas (a) e (b) de (1), é pertinente dizer que as etapas de (1) são etapas *em potência*, isto é, não implicam que, em todo processo de gramaticalização, haja todas as “mudanças” de todos os aspectos gramaticais previstos; ou ainda, que, na ocasião da identificação de um dado processo de gramaticalização num dado recorte de tempo, pode não ter ocorrido ainda todas as etapas previstas em (1).

Por outro lado, a afirmação de que a forma *poder* “passou de verbo a auxiliar” carece de explicitação e requer que sejam respondidas três questões: (i) o que é verbo? (ii) o que é auxiliar? (iii) o que significa “passar a”?

Em primeiro lugar, torna-se necessária uma teoria gramatical que nos forneça critérios consistentes, de modo a permitir demonstrar que um determinado comportamento sintático de um item atesta sua condição de membro pertencente a uma determinada categoria lexical ou gramatical. A Teoria Gerativa – que desenvolveu, como continuadora da tradição gramatical, uma teoria robusta das categorias sintáticas, sobretudo no que se refere à distinção crucial entre *as categorias lexicais e gramaticais* (estas últimas são chamadas de *funcionais* pelo gerativismo) – pode, por isso, servir como quadro teórico para o diagnóstico da natureza lexical ou gramatical dos itens. Para ilustrar este ponto, observemos, por exemplo, os contrastes seguintes:

- (3) a. ? Carlos a podia encontrar
- b.* Carlos a disse encontrar.
- c. Carlos a tinha encontrado.

Consideremos aqui três formas: *podia*, *disse* e *tinha*. Importa-nos saber qual delas é verbo, qual é auxiliar. Para isso podemos contar com critérios sintáticos, no caso, o alçamento do clítico,

isto é, o deslocamento do clítico da posição objeto do verbo que está numa das formas nominais para uma posição junto ao verbo que carrega as marcas de tempo e concordância. Como se sabe, apenas quando se tem uma locução verbal, é admitido o alçamento. Diante disso, o fato de (3b) ser impossível e (3a) aceitável, embora marginal, mostra que *podia* se comporta como um auxiliar ou, pelo menos, tem um estatuto categorial difícil de ser precisado. A Teoria Gerativa pode, portanto, participar desta análise na medida em que dispõe de meios para explicitar o comportamento sintático dos itens, tal como o alçamento, e ainda oferece uma justificativa teórica desse comportamento, por meio, neste caso, da teoria de fronteiras.¹ Esta distingue dois tipos de categorias: aquelas que permitem ou impedem o movimento de constituintes.

Torna-se crucial, na nossa perspectiva, a análise dos dois grupos de categorias sintáticas, isto é, as categorias lexicais e as categorias gramaticais, desenvolvida pela Teoria Gerativa; o que inclui (i) uma descrição das funções dos dois grupos na gramática interna; (ii) o rol das categorias que são lexicais e gramaticais; e (iii) o conjunto de traços que define, primeiro, a distinção gramatical e lexical, e, segundo, caracteriza o conteúdo de cada categoria prevista pela gramática interna.

É preciso agora ter em conta uma restrição. A etapa $a > b$ de (1) tem sido descrita como recategorização. Sob o rótulo da gramaticalização, têm sido tratados dois tipos distintos de processos: aquele em que um item lexical é recategorizado como gramatical e aquele em que um item lexical ou permanece como item da mesma categoria lexical ou passa a funcionar como um item de uma outra categoria também lexical. Este último processo é representável como: *lexical > lexical*. Trata-se de casos, por exemplo, como os seguintes:

¹ Ver Chomsky (1986).

- (4) a. redondo (adjetivo) > redondo (advérbio)²
b. braço (substantivo) > braço de cadeira (substantivo)

No caso de fenômenos como (4), é bastante relevante a discussão acerca do papel de mecanismos cognitivos subjacentes à metáfora e à metonímia e da dicotomia concreto/abstrato na sua criação e descrição semântica. Tais casos são certamente mais bem nomeados de *processos de lexicalização*, devendo, portanto, ser distinguidos dos processos de gramaticalização nos quais, como dissemos, um item pertencente a uma das categorias lexicais passa a funcionar como um item de uma das categorias gramaticais.

Voltemos ao percurso a > b de (1), em que uma categoria lexical é recategorizada como gramatical (ou funcional). O que significa *recategorização* de (a) neste contexto? Do nosso ponto de vista, recategorização constitui um epifenômeno no qual é possível “visualizar” a atuação das operações do sistema computacional, no sentido de Chomsky (1995). Isto é, existem mecanismos que colhem os itens no léxico e os alocam, formando com eles arranjos sintáticos. Ou, dito mais explicitamente, a gramática interna, ou sistema computacional, “enxerga” itens distintos, isto é, pertencentes a categorias distintas e forma arranjos sintáticos com eles por meio de suas operações que os inserem e os movimentam. O sistema computacional não “enxergaria”, portanto, o processo de recategorização. Este, na realidade, seria um epifenômeno captado pelo lingüista quando compara estágios diferentes de uma língua. Para esclarecer o papel do sistema computacional, consideremos ainda os seguintes exemplos:

- (5) a. Eu tenho dez vira-latas.
b. Eu tenho conhecido muita gente boa.

² Comparem-se os exemplos: (i) A terra é redonda; e (ii) A cerveja que desce redondo. Para uma discussão, consultar Reis (1997).

No estágio atual da língua portuguesa, o verbo *tenho*, em (5a), indicando posse, coexiste com o auxiliar *tenho* em (5b). Vamos supor, assim, que se trata de dois itens diferentes que pertencem ao componente lexical. O primeiro deles é categorizado como pertencente à classe dos verbos e o outro à da classe dos auxiliares. Quando o sistema computacional, através de suas operações, insere esses dois itens em arranjos sintáticos, ele “enxerga” os traços categoriais que definem a classe sintática dos dois itens e os aloca de acordo com esses traços. Assim, em (5), *tenho* [lexical] é inserido no ambiente __NP e *tenho* [gramatical] encontra-se no ambiente __VP. Em nossa ótica, portanto, a gradualidade identificada nos trabalhos sobre gramaticalização não tem a ver com a possibilidade de indefinição da classe de um item. Para nós, ou o item é de uma classe ou de outra. Em outras palavras, quando o item é inserido numa estrutura oracional, ele já tem sua classe sintática definida. Diferentemente das outras abordagens, não estamos focalizando aqui as potencialidades do item, mas sua presença em um contexto específico, o que acarreta a necessidade de um estatuto categorial definido, não ambíguo. Deve-se ressaltar aqui que o que dá a impressão de gradualidade categorial do item é o “olhar” externo sobre a língua. Dentro de uma estrutura um item nunca tem estatuto categorial indefinido ou mesmo ambíguo.³

Essa descrição dos fatos permite-nos também esclarecer por que utilizamos, neste texto, o termo ‘mudança’ entre aspas. Como vimos, tanto no caso do verbo modal quanto no caso de *ter*, não se trata de substituição de uma forma por outra, e sim de diferentes estatutos sintáticos atribuídos a uma mesma forma fonética, o que resulta em duas ou mais entradas lexicais distintas. Para explicitar

³ A propósito da reconhecida ambigüidade sintática de certos enunciados, é importante ressaltar que esta resulta da atribuição de mais de uma estrutura pelo falante e não do fato de o item não ter estatuto categorial claro nas diferentes estruturas atribuídas por ele.

essa nossa proposta, as diferentes entradas serão descritas como itens “homônimos”. Contudo, a literatura sobre gramaticalização, que prioriza o aspecto semântico do processo, trata, normalmente, os itens, aos quais estamos nos referindo, como polissêmicos. Entretanto, na nossa visão, embora do ponto de vista diacrônico, seja possível falar de polissemia já que uma acepção de um item é gerada a partir de outra, no que concerne a uma dada sincronia, vale mais a pena considerar que as duas acepções de um item, para um dado falante ou numa gramática particular, se encontram numa relação de homonímia, isto é, como propomos, trata-se de duas entradas lexicais distintas.

Na nossa perspectiva, portanto, a noção de gramaticalização é vista como um artefato descritivo que nos permite identificar e coletar certos tipos de fenômenos lingüísticos de natureza diacrônica. Estamos propondo aqui que a descrição pré-teórica, em termos de perda de substância fônica e esvaziamento de conteúdo, observadas nas línguas, no eixo diacrônico, seja consubstanciada através de metodologia objetiva de base quantitativa. Assim, as generalizações acima poderão ser traduzidas em questões:

- (6) Por que um item lexical perde substância fônica?
- (7) Por que um item perde significado?
- (8) Por que um item lexical passa a item gramatical?
- (9) Por que alguns entram em processo de gramaticalização e outros não?

Desse modo, o “diagnóstico” do processo de gramaticalização teria como suporte recursos objetivos de análise, de forma a minimizar o caráter assistemático das descrições, evitando críticas do

tipo “os estudos de gramaticalização dão impressão de unidirecionalidade porque apresentam lacunas”.⁴ A metodologia de base quantitativa é um caminho promissor, capaz de evitar que etapas de gramaticalização sejam omitidas. Através deste procedimento visamos tornar mais objetiva a nossa análise e também alcançar um grau de detalhamento superior ao das análises correntes.

Identificar a noção de gramaticalização como um artefato descritivo, tal como estamos assumindo aqui, significa sustentar que os estudos sobre gramaticalização formam um corpo de conhecimento no qual respostas para as questões (6-9) não poderão ser respondidas, porque não oferecem um quadro teórico no qual os conceitos de item lexical e item gramatical possam ser satisfatoriamente descritos. Por esta razão, fizemos opção pela Teoria Gerativa como quadro teórico onde buscar descrições adequadas.

2. CONCOMITÂNCIA

Tenhamos em conta a expectativa, decorrente do ciclo da gramaticalização em (1) e explicitada por Lehmann (1982), de que haveria concomitância entre perda de complexidade semântica e perda de substância fonética.

Vários fatos do português brasileiro nos quais a concomitância não se verifica podem, contudo, ser apontados. Conforme veremos, no eixo diacrônico, o item *ter*, que expressava *posse*, passa a também expressar *existência*.⁵ A perda de informação semântica, acrescida do ganho da acepção *existencial*, ocorreu sem a correspondente perda de substância fonética, embora, neste caso, possa ser alegado que o *ter existencial* é uma acepção lexical do verbo *ter*, ou seja, a acepção existencial não define uma categoria gramatical:

⁴ JANDA, Richard D. (2001: 265-340).

⁵ Cf. VIOTTI (1999).

- (10) a. Ele *tem* um apartamento.
b. *Tem* muita gente na sala.

Um outro caso seria o de perda de substância fonética, sem perda de informação semântica, como é o caso de “*tá*” e “*está*”, os dois itens manifestam verbo indicador de estado. Veja-se que “*tá*” é semanticamente idêntico a “*está*”, podendo ambos ser usados tanto numa função lexical, como em (11), como numa função gramatical, como em (12).

- (11) a. Ele *tá* bem.
b. Ele *está* bem.
(12) a. Ele *está* correndo agora.
b. Ele *tá* correndo agora.

Note-se, portanto, que, nesse caso, a redução fonética não é exclusiva do uso do item numa função gramatical.

Há ainda casos em que ocorrem os dois processos, a perda de substância fonética e enfraquecimento semântico simultaneamente, tendo o processo semântico afetado, ao mesmo tempo, as formas que perderam substância fonética e as que não perderam. Vejamos, como exemplo, as formas *você* e *cê*, nos enunciados a seguir.

- (13) a. *Cê* pode visitar sempre esse local. (referência [-específica])
b. Quando *você* chega em Paris (...). (referência [-específica])
c. VOCÊS *cês* aprontam a maior bagunça. (E1) (referência [+específica])

Os enunciados de (13) mostram que tanto a forma reduzida quanto a forma plena do pronome *você* são usadas com dois tipos de referência. Em (a-b) estes itens podem ser substituídos pelo pronome indetermina-

dor *se*, sem que isso afete o conteúdo da sentença. Já em (c) esta substituição não é possível, pois o referente é o próprio interlocutor. O processo manifestado como [+específico] > [-específico] configura o que Lehmann (1982) chama de “esvaziamento semântico” e o que Heine e Reh (1984) descrevem como “perda de complexidade semântica”.

Os dados apresentados em (10-13) mostram que não há concomitância entre os processos semânticos e morfossintáticos, contrariando Lehmann (1982).

3. CONCLUSÃO

O detalhamento e explicitação do ciclo apresentado no início deste capítulo, em (1), são, assim, necessários para uma abordagem formal do que tem sido rotulado como gramaticalização. Além disso, contribuem para uma melhor compreensão dos fenômenos com os quais estamos lidando.

Para efetivar tal detalhamento com objetividade e sistematicidade, a quantificação constitui um procedimento adequado e satisfatório. Para tanto, é necessário identificar as formas foneticamente plenas e reduzidas, compará-las, identificar sua categoria gramatical e diferenças semânticas. E cada uma dessas possibilidades pode ser codificada separadamente, de modo a torná-las explícitas e passíveis de quantificação, utilizando um programa de análise variável. Esse quadro, assim tão favorecedor ao tratamento quantitativo, permite verificar se as etapas em potência, previstas em (1), realmente se efetivaram e em que ritmo, datando seu percurso.⁶ A validade deste procedimento poderá ser comprovada nas análises desenvolvidas nos capítulos seguintes deste livro.

⁶ Uma aplicação da análise propriamente laboviana a dados de gramaticalização encontra-se em Naro e Braga (2000). Para uma crítica quanto à falta de informações sociolinguísticas no tratamento dispensado à gramaticalização, ver Janda (2001).

CAPÍTULO 2

A FORMA CÊ E A NOÇÃO DE GRAMATICALIZAÇÃO¹

1. O PROCESSO DE GRAMATICALIZAÇÃO

Como já dissemos, a noção de gramaticalização, estabelecida por Meillet em 1958 ao retomar idéias de Humboldt desenvolvidas por indo-europeístas (ver Hopper e Traugott (1993); Lehmann (1982); Heine, Hünemeyer e Claudi (1991); Castilho (1995)), tem sido retomada nestes últimos anos e utilizada, com muito proveito, na análise de fenômenos de linguagem.

Esta noção 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”.² A gramaticalização é, então, a mudança de estatuto de um termo da língua: ao perder “significado”, um item lexical passa a ter uma função gramatical.

Hopper e Traugott (1993:2) discutem, entre muitos outros, o exemplo do verbo inglês *to go*. Como vimos, há um verbo *to go* que é um verbo de moção e outro que funciona como auxiliar.

¹ Uma versão desse texto foi publicada pela revista *Estudos da Linguagem*, 5, p. 115-124, 1996, por Vitral, L.

² Kurylowicz (1965), *apud* Lehmann (1982); Castilho (1995: 3).

Historicamente, o verbo auxiliar deriva do verbo de moção. Pode-se dizer, então, que houve gramaticalização: o verbo de moção passou a desempenhar a função de auxiliar.

A gramaticalização não ocorre de maneira abrupta, isto é, a mudança de estatuto de um item lexical tem lugar através de transições graduais que, de acordo com Hopper e Traugott (1993), tendem a suceder, nas línguas, de forma similar. Os autores que trabalham com a noção de gramaticalização parecem concordar com a possibilidade de se destacar as seguintes etapas deste processo:

- (1) item com significado lexical > item gramatical > clítico > afixo flexional

Neste capítulo, pretende-se examinar, através da noção de gramaticalização, o estatuto do item *cê* do português brasileiro atual. Sabe-se que o pronome *você* se origina da forma de tratamento *Vossa Mercê*, tendo havido um estágio intermediário – *Vosmecê* – que foi abandonado (ver Mattoso Câmara Jr. (1979:94); Gonçalves (1987)). Os seguintes estágios teriam sido manifestados:

- (2) *Vossa Mercê* > *Vosmecê* > *Você*

Parece pertinente supor que estes estágios manifestam um processo de gramaticalização, já que uma forma de tratamento, contendo propriedades conotativas, isto é, “favor vosso” ou “respeito a alguém hierarquicamente superior”, passou a desempenhar a função de pronome, perdendo estas propriedades e mantendo apenas um conjunto de traços pronominais (chamados de traços-phi por Chomsky (1981:330)). Pode-se pensar, assim, de acordo com a terminologia utilizada no gerativismo, que o

constituente *Vossa Mercê*, composto do nome *mercê* e do determinante *vossa*, teria passado a dispor, na sua estrutura, apenas do determinante *você*.

Este processo, no entanto, parece estar em curso. O português atual dispõe ainda das formas *cê* e *ocê* (que são bastante usuais no falar mineiro). A distribuição destas três formas – *você*, *ocê* e *cê* – não é idêntica. E é com base no comportamento sintático delas que vamos propor estar havendo um processo de cliticização.

2. A DISTRIBUIÇÃO DAS FORMAS VOCÊ, OCÊ E CÊ:

Observem-se os seguintes exemplos (de Dutra (1991)):

Posição sujeito:

- (3) (V)*ocê* pediu pra sair.
Cê pediu pra sair.

Posição objeto:

- (4) Eu amo (v)*ocê*.
Eu amo **cê*.

Pós-verbal:

- (5) a. Eu fui o culpado.
(V)*ocê* foi o culpado.
Cê foi o culpado.
b. Fui eu o culpado.
Foi (v)*ocê* o culpado.
Foi * *cê* o culpado.

- (6) a. *Aí (v)ocê* chega pra mim ontem e diz...
 Aí cê chega pra mim ontem e diz...
 Aí ele chega pra mim ontem e diz...
 b. *Aí* chega *(v)ocê* pra mim ontem e diz...
 Aí chega * *cê* pra mim ontem e diz...
 Aí chega ele pra mim ontem e diz...

- (7) *Lá* vem ele com mentiras.
 Lá vem *(v)ocê* com mentiras.
 Lá vem * *cê* com mentiras.

Preposto:

- (8) *(V)ocê* ele não viu.
 * *Cê* ele não viu.

Modificado por advérbio.

- (9) *Só (v)ocê* tava mentindo.
 Só * *cê* tava mentindo.

- (10) *Até (v)ocê* podia subir.
 Até * *cê* podia subir.

- (11) – *Quem* vai sair?
 – *(V)ocê*.
 – * *Cê*.

Pode-se acrescentar a este conjunto de dados o ambiente objeto de preposição:

- (12) *Eu* falei pra *você*...
 Eu falei *pr'ocê*...
 Eu falei * *pra cê*...

E o da participação numa coordenação:

- (13) a. Ele e (v)ocê podem votar contra.
b. * Ele e cê podem votar contra.

Os exemplos acima mostram-nos a existência de uma distribuição complementar: a forma *cê*, exceto na posição sujeito, não pode ocorrer nos ambientes em que *você* e *ocê* ocorrem.

A questão que se coloca então é saber que hipótese pode explicar o contraste entre essas formas. Tanto *você* e *ocê* quanto *cê* são pronomes, o que significa terem os mesmos traços. Entretanto, alguma diferença há, uma vez que o comportamento sintático de *cê* é distinto.

Vamos propor a hipótese segundo a qual a forma *cê* constitui uma etapa do processo de gramaticalização da forma *Vossa Mercê*. Trata-se da etapa da cliticização apontada em (1). Assim, tem-se:

- (14) item com significado lexical: *Vossa Mercê* > item gramatical: *você* > clítico: *cê* > afixo flexional

A forma *vosmecê* seria considerada um estágio intermediário entre a primeira e a segunda etapas acima, e *ocê* pode ser tomado como uma forma intermediária entre a segunda e a terceira.³

No entanto, como será desenvolvido neste livro, a forma *cê* não é um clítico como os pronomes pessoais átonos das línguas românicas, sobretudo quando se observam suas características

³No estado de Goiás, ocorre o seguinte fenômeno:

- (i) ê besta (sô!)
“você é besta!”

Em (i), o verbo *ser*, flexionado, é realizado pela vogal fechada [e]. Talvez este exemplo ilustre a quarta etapa do processo de gramaticalização apontado em (14), isto é, a etapa da afixação flexional.

fônicas. Vamos considerar a perspectiva de que esta forma se encontra, na realidade, num *processo de cliticização* e que cliticização é um fenômeno variável e heterogêneo. A hipótese de *cê* como uma etapa de (14) atende, por enquanto, à necessidade de tornar inteligível e dar início à discussão acerca do percurso de mudança que nos leva de *Vossa Mercê* a *cê*.

3. A CLITICIZAÇÃO DE CÊ

Vamos agora voltar aos dados (3-13) e examinar como nossa hipótese pode explicá-los.

O contraste entre (3), (5a), (6a), por um lado, e (4), (5b), (6b) e (7), por outro, pode ser tratado se considerarmos que, (i) *cê* é um clítico e que (ii) em frases simples, o português brasileiro falado, diferentemente do que ocorre no português europeu, apresenta o clítico sempre anteposto ao verbo (ver, entre outros, Said Ali (1950); Duarte (1986); Rouveret (1989); Galves (1990); Pagotto (1993); Kato (1994)). Pode-se dizer que “todos os casos de ênclise produzidos pelos aprendizes são produtos do contacto com a escrita ou da intervenção da escola” (Kato (1994: 13)), ou seja, quando usa a ênclise, o falante se encontra em um registro linguístico diferente daquele no qual a forma *cê* é corrente.

A má-formação de uma frase como **Eu vi cê* se deve, então, à impossibilidade de ênclise na gramática do português brasileiro falado. Neste dialeto, entretanto, a seqüência *V + cê* é bem formada no seguinte contexto:

(15) Vou fazer *cê* feliz.

O exemplo (15) confirma, na verdade, que a restrição sobre *cê* diz respeito à impossibilidade de esta forma aparecer na posição objeto. Ora, em (15), o constituinte “*cê* feliz” deve ser analisado como uma minioração (*small clause*) que dispõe de *cê* como sujeito.

Não se discutirá aqui a razão da impossibilidade da ênclise. Alguns dos autores citados no parágrafo anterior têm procurado relacionar esta impossibilidade à posição do verbo na sintaxe visível, isto é, no nível da Forma Fonológica.

Nossa hipótese acarreta três problemas:

(i) Estudos quantitativos mostram que a frequência de clíticos está em queda no português brasileiro falado (ver Tarallo (1983); Kato (1994); Pagotto (1993); Cyrino (1993)).

(ii) Em frases como *Cê foi o culpado*, teríamos um clítico em início de sentença, fenômeno este que é restrito a certos contextos.

(iii) Nossa hipótese prevê a boa-formação de frases tais como *José cê viu*, com próclise. Esta sentença é, entretanto, rejeitada pelos falantes.

Com relação à (i), Nunes (1991) mostra que a inserção do clítico *se* indeterminador estaria em expansão no português brasileiro, sendo o *se* associado a um pronominal vazio foneticamente na posição sujeito. Note-se que *cê*, assim como *se*, pode ser usado como um marcador de indeterminação do sujeito. Além disso, é importante ressaltar que é a frequência de clíticos acusativos que estaria em queda acelerada no português brasileiro falado.

No que concerne à (ii), pode-se supor que a ocorrência da forma *cê* em início de frase pode estar associada a uma tendência de preenchimento da posição sujeito, tendência esta também apontada por estudos quantitativos (ver Tarallo (1983); Duarte (1995)).

Por outro lado, o lugar da negação nos mostra que a estrutura com *cê* pré-verbal não é idêntica à de *se*:

- (16) a. Não se diz muita bobagem por aí.
- b. Cê não viu a Maria no cinema.

A frase (16b) não é incompatível com nossa hipótese da cliticização de *cê* sujeito. Consideraremos que, ao se movimentar da posição sujeito interna ao VP, *cê* se aloca na posição de adjunção a T na qual se encontra a negação. Esta análise corrobora a previsão de Chomsky (1995b:402) de que o clítico tem o estatuto de projeção máxima, quando se encontra em sua posição temática, e também o estatuto de núcleo, quando adjunge a um núcleo funcional. A indistinção entre as posições de adjunção e de especificador, defendida por Kayne (1993), é também adotada em nossa análise: a checagem de traços que se passa entre T e o sujeito é legítima mesmo quando há adjunção do sujeito a T (ver Chomsky (1995a: sec. 10.1)).

Já a estrutura (15a), de acordo com Cinque (1988: 534) para o italiano, contém um *pronominal* na posição de especificador de IP e *si* (=se) é gerado no núcleo I (ou T). Esta partícula parece, então, ter o estatuto de um afixo, o que pode ser a causa da diferença de distribuição entre esta partícula e *cê*. Teríamos, assim, as seguintes estruturas subjacentes:

- (17) a.....[*cê* [não [viu [t [tt.....
 T T T NegP VP
 b.....[não [se [diz [t [tt.....
 T T t NegP VP

A análise proposta acima não é, entretanto, o único recurso de que dispomos. É possível pensar também que *cê* tem características que o distinguem dos demais clíticos, ou seja, a cliticização não é um fenômeno homogêneo como normalmente se pensa, principalmente, no domínio das línguas românicas. Nos capítulos seguintes, esta idéia será discutida em detalhes. Investigaremos, por exemplo, se, do ponto de vista fonético, faz sentido dizer que *cê* é um clítico e se, do ponto de vista estrutural, não seria mais

adequado considerar que se trata de uma projeção máxima e não um núcleo como propusemos em (17).

Por outro lado, havia, nos períodos arcaico e moderno do português, a possibilidade de um clítico se alocar numa posição antes da negação. Por exemplo:

- (18) “...quesem estabelecimento o não quer fazer...”
(*Documentos históricos do Espírito Santo, 1ª metade do século XVIII*; citado por Pagotto (1993: 188))

Este fenômeno, conhecido por *interpolação*, é bastante complexo e sua análise detalhada, que será realizada no Capítulo 4, é crucial na análise do processo de cliticização no qual a forma *cê* está envolvida.

Com relação ao problema (iii), podemos talvez nos contentar, por enquanto, com a especulação de que a estrutura *José cê viu* não ocorre porque a gramática do dialeto no qual *cê* é produtivo não é a mesma em que admite a próclise. De toda maneira, são pontos que vão exigir análise aprofundada.

Vamos agora, finalmente, analisar a pertinência de nossa hipótese em relação aos dados (8-13). A agramaticalidade da presença de *cê* nestes ambientes é, então, explicada pela natureza clítica desta partícula. Ora, nesses mesmos ambientes, não encontramos clíticos no português brasileiro falado. Assim, um clítico não pode aparecer topicalizado como em (8); nem ser modificado por advérbio como em (9) ou (10); não pode ser, sozinho, resposta a uma pergunta como em (11), isto é, um elemento que está se tornando clítico precisa se escorar em algum outro termo ou, dito de outra maneira, ter um “hospedeiro”; não pode ser complemento de preposição, como em (12), ambiente em que só encontramos as formas fortes ou tônicas; e, enfim, não pode ser coordenado com uma forma

tônica como em (13), considerando que só coordenamos constituintes similares.⁴

Concluiremosmos afirmando que, embora tratada ainda de forma superficial, a hipótese central deste capítulo – isto é, de que a forma *cê* se encontra num processo de cliticização, que faz parte de um processo mais amplo de gramaticalização – encontra plausibilidade nos dados analisados, apesar de gerar um certo número de questões, desenvolvidas nos demais capítulos, que mostram, porém, a fecundidade de nossas proposições.

⁴O acento circunflexo em *cê* parece indicar que a vogal é fechada e não a indicação de tonicidade. As propriedades fonéticas desta forma serão analisadas no Capítulo 5.

CAPÍTULO 3

GRAMATICALIZAÇÃO EM PROCESSO¹

Trataremos aqui da freqüência e distribuição de três formas pronominais de segunda pessoa indireta na língua falada no momento atual, no português brasileiro. Assumindo a hipótese de que as formas *você*, *ocê* e *cê* são itens distintos e manifestam etapas de um processo de gramaticalização, tal como proposto no capítulo anterior, é nosso objetivo verificar se esses três itens coexistem no momento atual. Outro objetivo é verificar se estes ocorrem nos mesmos contextos ou se há especialização. O terceiro objetivo é verificar se a perda de substância fonética observada foi concomitante ao esvaziamento semântico.

Na primeira parte deste capítulo será descrito o *corpus* analisado e verificada a ocorrência e freqüência dos itens em análise. Serão também identificados diferentes contextos e será verificado se há co-ocorrência dos três itens nos mesmos contextos. Na segunda parte será observada a distribuição das formas com base no tipo de referência e realização fonética. Para capturar a evolução temporal será feita uma análise com base no tempo aparente.²

¹ Uma versão deste texto foi publicada como capítulo de livro por RAMOS, J. 1997a.

² Este tipo de análise consiste em verificar se há estratificação etária. Se houver, espera-se que a forma mais antiga seja mais freqüente na faixa dos mais idosos e menos freqüente na faixa dos mais jovens. Esse

1. A AMOSTRA

A amostra selecionada para este estudo constitui-se de entrevistas realizadas com 12 informantes de Belo Horizonte, de três faixas etárias. Nesse *corpus* foram identificadas 342 ocorrências dos itens em análise. O valor numérico e a frequência de cada item aparece na tabela abaixo.

Tabela 1
NÚMERO DE PORCENTAGEM DOS ITENS NO *CORPUS*

	Nº	%
Você	112	33
Ocê	41	13
Ce	189	55
Total	342	100

Chama a atenção a co-ocorrência dos três itens. Isso mostra que, no processo de gramaticalização, o surgimento de uma nova forma não implica o desaparecimento das demais. Veja-se na tabela que as frequências mais altas são exatamente a da forma mais plena e a da forma mais reduzida.

Presentes as três formas, consideradas aqui itens lexicais distintos, torna-se importante verificar se cada item, no que diz respeito à distribuição sintática, se diferencia dos demais. Se a resposta for positiva, estaremos diante de um caso de especialização.

perfil constitui um indicador de mudança. Para confirmação do processo de mudança é necessário realizar análise no tempo real.

Tabela 2
DISTRIBUIÇÃO DOS ITENS CONFORME A FUNÇÃO SINTÁTICA

	Cê		Ocê		Você		Total
	Nº	%	Nº	%	Nº	%	
Sujeito	188	58	38	22	96	30	322
Complem. de preposição	–	–	3	17	15	83	18
Complem. de verbo	–	–	1	50	1	50	2
Total	188		42		112		

A posição de sujeito hospeda todos os itens. Já as posições de objeto constituem um *locus* das formas mais plenas, os itens *ocê* e *você*. O que impediria *cê* de ocorrer nestes últimos contextos sintáticos? De acordo com a proposta desenvolvida no Capítulo 1, seria a ausência de ênclise, de modo generalizado no português brasileiro.

Essa distribuição dos itens estaria configurando um processo de especialização? Para responder a tal pergunta, vamos inicialmente definir esse processo e depois voltar à Tabela 2. A especialização consiste na *redução da variedade de escolhas formais disponíveis assim que os significados assumem maior generalidade gramatical* (Bréal, 1991 [1882] *apud* Hopper e Traugott, 1993: 114). Assim definido, o processo de especialização parecer implicar dois fatos: (i) os diferentes itens em análise aqui, por resultarem de um processo de gramaticalização, seriam semanticamente distintos; (ii) a distribuição sintática de cada novo item seria diferenciada, já que a *variedade das escolhas formais sofreria redução*. Assim, a forma mais gramatical teria uma distribuição mais restrita.

Para avaliar a implicação (i), vamos verificar o uso de cada um dos itens em relação a dois traços semânticos: animacidade e tipo de referência. Todas as ocorrências de todos os três itens referem-

se sempre a um interlocutor, seja real seja virtual, porém sempre [+ humano]. A distinção em termos de animacidade, entretanto, não é suficiente para distinguir os itens.

Consideremos, então, o tipo de referência. É inovador na língua o uso da variante *você* com referência indefinida/arbitrária (cf. Duarte, 1995). Como a variante *cê* é a forma mais gramatical, poderá ser esta a forma usada preferencialmente com referência indefinida. A análise do *corpus* mostra, entretanto, que todos os itens ocorrem com ambos os tipos de referência. Com esse tipo de referência é possível a paráfrase com outros recursos de indeterminação, a saber, um pronominal vazio de terceira pessoa do plural (1b), ou clítico *se* (1c), ou mesmo com NPs do tipo “a pessoa” (1d), abaixo.³

- (1) a. *Cê* pode visitar sempre esse local.
- b. *Podem* visitar sempre esse local.
- c. *Pode-se* visitar sempre esse local.
- d. *A pessoa* pode visitar sempre esse local.
- e. *Você* sempre sabe.

Veja-se que em (1c) o clítico nominativo ‘*se*’⁴ tem referência indefinida, tal como ‘*cê*’ em (1a) e (1e). O par (1a)-(1e) enfraquece a hipótese da especialização.

Passemos agora a observar a incidência de focalização no item. Foram encontradas três ocorrências com foco marcado, mais exatamente, ocorrências com acento contrastivo.

- (2) a. E VOCÊS num quer? (E1)
- b. Eu confiar em VOCÊ? (E1)
- c. VOCÊS cês aprontam a maior bagunça. (E1)

³ Sobre o estudo das formas (1a-c) como variantes sociolinguísticas, ver Tarallo e Kato (1986); e Duarte (1995).

⁴ Nunes, J. (1990).

Nenhuma delas apresentou o item *cê* focalizado, e sim *você*. O enunciado (2c) reforça a hipótese de que *cê* seria um clítico: por ser átono, não aparece em posição focalizada. O fato de a variante *ocê* também não ter ocorrido com foco enfraquece, de certo modo, esta evidência, pois impede distinguir estas duas variantes quanto à cliticidade, tomando-se como critério o recebimento, ou não, de focalização.

Vejam agora a distribuição dos itens conforme a idade do informante. Tendo em vista o processo histórico representado em (3) abaixo e o comportamento variável dos itens em análise, é de se esperar que falantes de faixas etárias diferentes apresentem um comportamento que indique mudança em progresso. A tabela a seguir mostra a distribuição das ocorrências segundo a faixa etária do informante.

(3) *Vossa Mercê* > ... > *Vosmecê* > ... > *você* > *ocê* > *cê*

Tabela 3

DISTRIBUIÇÃO DOS ITENS CONFORME A IDADE DO INFORMANTE

	T1		T2		T3		Total	
	Nº	%	Nº	%	Nº	%	Nº	%
<i>Cê</i>	76	61	64	50	49	54	189	55
<i>Ocê</i>	10	8	15	12	16	18	41	12
<i>Você</i>	39	41	48	38	25	28	112	33
Total	125	100	127	100	90	100	342	100

O item preferido pelas três faixas etárias foi *cê*, respectivamente, 54%; 50% e 61%. Este resultado é surpreendente porque esse item, que diacronicamente é inovador, se distribui de modo semelhante

pelas três faixas etárias. Estes resultados levam a concluir que a forma *cê* já está implementada na língua e, com base no tempo aparente, parece não se poder falar em mudança em progresso.

Esta conclusão é surpreendente e, para verificar sua adequação, consideremos o tipo de referência do item, pela seguinte razão. Os itens tendem a sofrer alteração, passando de [+específico] a [-específico]. Até que ponto não estaria havendo, no momento, uma preferência por *você* [+específico] e por *cê* [-específico]?

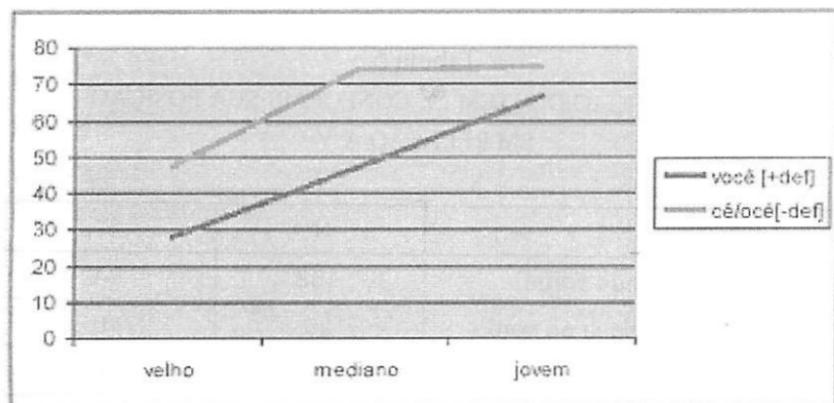
Vejamos então os resultados do cruzamento do tipo de referência e faixa etária. Para isso, vamos agrupar os itens conforme a realização fonética: de um lado, a forma pronominal plena e, de outro, as duas formas pronominais reduzidas. Deste modo, é possível reunir as formas em análise em duas subclasses. Assim, de um lado teríamos *ocê* e *cê*, e de outro *você*. Este recurso metodológico visa a testar a hipótese de que o uso dos itens com referência [-específica] teria favorecido a perda de substância fonética.

Tabela 4
DISTRIBUIÇÃO DOS ITENS, O TIPO DE REFERÊNCIA
E A FAIXA ETÁRIA DO INFORMANTE

	Referência Definida				Referência Indefinida			
	Forma reduzida		Forma plena		Forma reduzida		Forma plena	
	Nº	%	Nº	%	Nº	%	Nº	%
T3	2	33	4	67	75	75	21	25
T2	38	53	34	47	74	74	14	26
T1	79	72	31	28	7	47	08	53
Total	102		86		87		43	

Se interpretarmos os dados da Tabela 4 como tempo aparente, constatamos na língua uma tendência ao uso de *você*, preferencialmente, com referência [+específica] e uso das formas foneticamente reduzidas com referência [-específica]. Este comportamento diferenciado fica claro no gráfico abaixo, e confirma a correlação prevista pela *cline* da gramaticalização.

Gráfico 1
DISTRIBUIÇÃO DOS ITENS COM REFERÊNCIA [-ESPECÍFICA],
CONFORME A FAIXA ETÁRIA



O Gráfico 1 mostra uma tendência crescente ao uso de *você* com referência [+específica].⁵ Embora se possa observar esta tendência, é importante notar que um mesmo item ora é usado com referência [+ específica] ora com referência [- específica], o

⁵ Observa-se um perfil descendente no uso do item *você*. Este processo poderia ser descrito como uma migração do sistema definido para o sistema indefinido. Segundo Laberge e Sankoff (1976) teria ocorrido esse tipo de migração no francês de Montreal em relação ao clítico indefinido "on".

que indica que a perda de conteúdo semântico não é concomitante à perda de substância fônica.

Passemos, agora, a observar a posição dos itens em análise em relação ao verbo, lembrando que os clíticos pronominais no PB ocorrem ou antes ou após o verbo, mas sempre contíguos. Comparem-se:

- (4) a. Eu não o vi.
- b. ?Eu o não vi.
- c. *Eu não o infelizmente vi.

Vejamos agora os resultados.

Tabela 5
DISTRIBUIÇÃO DO ITEM *CÊ* CONFORME SUA POSIÇÃO
EM RELAÇÃO A V

	Nº	%
Posição contígua ao verbo	162	84
Posição não contígua ao verbo	26	16
Total	188	100

A posição favorecida é a contígua, com 84%. Mas há casos de posição não contígua, tendo sido encontradas construções como (5), abaixo. Aqui o item “não” aparece entre o item *cê* e o verbo, o que corresponde a 85% do total de 26 casos de não contigüidade. Os outros itens que ocorrem neste contexto são *já* (7,4%), *só* (3,8%) e *sempre* (3,8%).

- (5) *Cê não paga nada.* (E7)

Esse tipo de ocorrência coloca um problema: se o item *cê* for um clítico, então não seria um clítico de comportamento típico, já que clíticos no PB não ocorrem à esquerda das partículas *não*, *já* e *só*. Comparem-se as orações negativas em (4).

A ordem [cê-não-V], se analisada como [clítico-não-V], parece enfraquecer a hipótese inicial porque, do ponto de vista diacrônico, reconhecer tal ordem no PB atual implica aceitar o reaparecimento de uma ordem que deixou de ser produtiva no século XIX (Morais, 1995: 261).

Uma saída para superar este obstáculo seria propor que *cê* seja um clítico que ocupa a posição de especificador de FLEX,⁶ acompanhando proposta de Kayne (1983), para o francês. Deste modo, *cê* estaria numa posição diferente daquela ocupada por *o* em (4a-b). Assim, o reaparecimento da mesma ordem não implicaria reaparecimento da mesma estrutura. Esta análise permitiria obter um resultado que está em harmonia com duas tendências mostradas por vários estudos diacrônicos sobre o português do Brasil: (i) uso cada vez menos freqüente de clíticos em posição de núcleo de FLEX (no caso, clíticos acusativos (Duarte, 1989)); e (ii) tendência ao preenchimento da posição sujeito (Tarallo, 1983; Duarte, 1995; e outros). Esta solução permite dar conta também dos demais itens que ocorreram no contexto [cê - V].

Feitas estas considerações sobre a contigüidade/não-contigüidade em relação ao verbo, vejamos agora a posição dos itens nas construções de tópico. Foram encontradas 26 construções de tópico, em 88% dos casos o item *cê* ocorre sempre na segunda posição. Um exemplo seria:

⁶ Estou adotando aqui a proposta de Rizzi (1986) para quem o clítico poderia ocupar a posição de especificador da categoria funcional FLEX ou a posição de núcleo dessa categoria, tal como, respectivamente, clítico "il" do francês e "la" do trentino, um dialeto italiano.

(6) *vocês/ cês* aprontam a maior bagunça. (E1)

A forma reduzida neste contexto reforça mais uma vez a hipótese da cliticização, na medida em que funciona como uma retomada do tópico. A posição de tópico pode receber foco, e é ocupada por projeção máxima. O fato de *cê* não ocorrer na posição de tópico constitui mais uma evidência de seu caráter foneticamente reduzido e mais gramatical que o pronome pleno *você*.

Finalmente, consideremos o tipo de oração. Três tipos serão investigados: declarativas (em (7)), interrogativas simples (em (8b; 9b)) e interrogativas 'que-que' (em (8a e 9a)). Os resultados aparecem na Tabela 6.

(7) Maria, *cê* tá muito calma. (E9)

(8) a. De que que *cê* gosta então? (E10)

b. O que *você* gosta então? (E10)

(9) a. Como que *cê* acha que era o namoro? (E11)

b. Como *você* consegue os ursinhos? (E9)

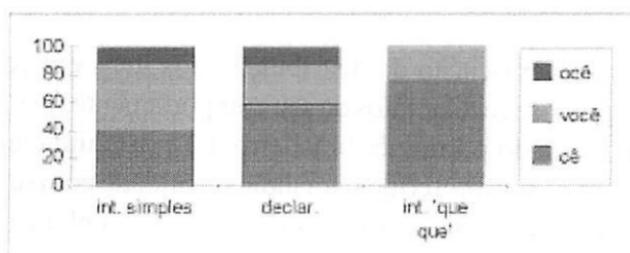
Tabela 6

DISTRIBUIÇÃO DOS ITENS CONFORME O TIPO DE ORAÇÃO

	Você		Oê		Cê		Total	
	Nº	%	Nº	%	Nº	%	Nº	%
Interrog. Simples	42	48	11	12	35	40	88	100
Interrog. 'que-que'	5	24	–	–	16	76	21	100
Declarativas	65	28	30	13	138	59	133	100
Total	112		41		189		342	

Nas interrogativas simples o item preferido é *você*, nas interrogativas ‘que-que’ e nas declarativas, *cê*. O item *ocê* não ocorreu com interrogativas ‘que-que’. Esse comportamento diferenciado fica mais claro no Gráfico 2.

Gráfico 2
DISTRIBUIÇÃO DAS TRÊS VARIANTES POR TIPO DE ORAÇÃO



Estes resultados permitem subdividir esse conjunto de orações em dois grupos, de um lado as declarativas e interrogativas simples, e de outro as interrogativas ‘que-que’. A comparação destes dois grupos mostra uma diferença em relação ao complementizador: apenas as interrogativas ‘que-que’ têm o complementizador preenchido lexicalmente. A questão que se coloca é: qual a relação entre o preenchimento lexical do núcleo e o favorecimento da forma *cê*? Uma resposta para esta questão poderá trazer subsídios para a testagem da hipótese maior discutida neste volume, a saber, a hipótese de que *cê* é clítico. Se *cê* for clítico fonológico, a seqüência [especificador + núcleo lexicalizado + clítico fonológico] não é surpreendente, visto que *cê* teria que ser adjungido ao núcleo do complementizador, na forma fonológica.

Confirmando a argumentação do capítulo precedente, os itens em análise manifestam etapas do processo de gramaticalização. A análise quantitativa mostrou-se também um instrumento eficaz para testar hipóteses gramaticais e para detalhar fases de uma *cline*,

explicitando que as etapas do processo de cliticização não são adequadamente descritas através de (10), mas sim através de (11).

(10 (=4)) você > ocê > cê
referência [+específica] > referência [-específica]

(11) você > ocê > cê você > ocê > cê cê
Referência [+específica] > Referência [-específica] > Referência [-específica]

Através da verificação da distribuição dos itens em contextos sintáticos, pôde-se revelar fatos de seu comportamento no momento atual da história do português brasileiro: (i) existe uma tendência de que o item *cê* seja preferido para indicar entidade com referência [-específica]; (ii) a forma *cê* ocorre adjungida ao complementizador, lexicalmente realizado, em interrogativas, o que evidencia seu caráter de elemento não acentuado no nível fonológico; (iii) o fato de *cê* ser licenciado em posição não contígua ao verbo constitui uma evidência de que seria um clítico fonológico, mas não um clítico sintático. Uma vez que no quadro teórico adotado aqui, em que se desenvolveram as análises de Kayne (*op. cit.*) e Rizzi (1986), estar em uma ou outra posição não afeta o estatuto de clítico de *cê*, e como a definição deste estatuto constitui a hipótese central aqui, por uma questão de limite do presente capítulo, concluímos que *cê* ocupa a posição de especificador de T. Ver Capítulo 7.

Esta questão, relativa à posição e estatuto do item *cê*, não é simples e está longe de ser resolvida. Primeiro porque os itens presentes no contexto [*cê* – V], pelo menos neste *corpus*, são curtos e “leves”, alguns deles reconhecidos como clíticos, como *já* e *não*, o que poderia levar a um amálgama na Forma Fonológica, sendo a ordem definida morfologicamente (Bonet, 1995). Em segundo lugar, o item *cê* ocorre com o item *não*, realizado foneticamente como *num* (ver Capítulo 8), o que reforça o estatuto de clítico

desta última. Em terceiro lugar, o número de casos de não-contigüidade, se suprimidos os possíveis clíticos, soma um total de 7,6%, o que parece indicar certo caráter marginal.

Feitas estas considerações, passemos ao próximo capítulo, em que se buscam evidências de natureza fonética a favor da hipótese de cliticização do item *cê*.

CAPÍTULO 4

IDENTIFICANDO CLÍTICOS: EVIDÊNCIAS FONÉTICAS¹

Como vimos nos Capítulos 1 e 2, a forma pronominal reduzida de 2ª pessoa do singular, isto é, *cê*, redução de *você*, que ocorre em diversos dialetos do português brasileiro, foi analisada como um clítico. Esta análise se deveu à distribuição sintática dessa partícula. Por outro lado, sabe-se que a própria definição de clítico, advinda do grego e que significa “o termo que se escora em outro”, destaca, na realidade, as propriedades sonoras desses elementos, especificamente, o fato de serem átonos. É verdade que a atonicidade dos clíticos condiciona a distribuição sintática deles, mas isso viria como uma consequência de serem átonos, o que seria o traço definitório desses itens.

Essa visão do fenômeno da cliticização, que chamaremos de tradicional, atribui, evidentemente, ao componente fonológico da gramática a capacidade de caracterizar os clíticos e de prever o comportamento sintático deles. Esta visão dos fatos pode ser vista como conflitante com a hipótese central da Teoria Gerativa acerca da autonomia do componente sintático da gramática.

Existe uma análise de Brandão de Carvalho (1986) a respeito dos fenômenos da ênclise e da próclise, analisados, comparativamente,

¹ Este capítulo foi escrito por Lorenzo Vital.

no português brasileiro e no português europeu, que nos permite aferir, de forma bastante objetiva, a contribuição dos componentes sintático e fonológico na determinação do fenômeno da cliticização. Sabe-se, para ilustrar este ponto, que o português europeu exige a próclise quando, na oração, ocorrem certas palavras “atratoras” como a negação, um sintagma interrogativo, certos advérbios e outras. Na ausência desse tipo de palavra, ocorre a ênclise nesse dialeto do português. A posição do clítico, antes ou depois do verbo, parece, assim, ser dependente de algum mecanismo sintático, o que tem sido buscado em vários trabalhos na literatura sobre cliticização (cf. Rouveret (1989); Martins (1994)). Determinada, portanto, a posição sintática do clítico em relação ao verbo, ele pode, de acordo com Brandão de Carvalho, sofrer a ênclise fonológica, no caso do português europeu, ou a próclise fonológica, em se tratando do dialeto brasileiro, em relação ao constituinte à sua esquerda ou à sua direita. Essas duas possibilidades de cliticização fonológica dependem, na sua proposta, do padrão acentual diferenciado dos dois dialetos, o que é explicado através da Teoria Métrica. Esta análise é um exemplo de como pode ser pensada a interação dos dois componentes da gramática na descrição e análise do fenômeno da cliticização.

Já tendo sido atestado o estatuto de *cê* como clítico sob o ponto de vista sintático, neste capítulo, serão analisadas características fonéticas da forma *cê* a fim de verificar se também nesta perspectiva faz sentido analisá-la como clítico. O capítulo contém duas partes.

Na primeira parte, comentamos um experimento que consistiu na análise acústica da fala de cinco informantes que tiveram gravadas suas leituras de pequenos textos, contendo sentenças nas quais aparecem a forma de segunda pessoa *cê* o nome *Zé* e o clítico pronominal de terceira pessoa *se*. Através do programa de análise da fala WINPITCH1.92, foram medidos e comparados os parâmetros da *intensidade* e *duração* dessas três formas com o objetivo de examinar se *cê* aproxima-se de *se*, um clítico

incontestável, ou de *Zé* em relação aos parâmetros mencionados. Na segunda parte deste capítulo, discutimos se a forma *cê*, em lugar de ser classificada como um clítico, não seria mais adequadamente analisada como uma *forma fraca*, como propõem Cardinaletti e Starke (1994) e outros autores para certas formas pronominais reduzidas que ocorrem em várias línguas.

1. A ANÁLISE ACÚSTICA

Foi realizada uma gravação com cinco informantes, sendo três do sexo feminino e dois do sexo masculino, que consistiu na leitura de três pequenos textos – lidos duas vezes por cada informante – dos quais faziam parte as sentenças seguintes, divididas em três grupos:

- A. (1) Como *cê* falou que a gente devia fazer?
(2) Como *Zé* falou que era mesmo?
(3) Como se falou que a gente devia pronunciar?

- B. (4) *Cê* sabe de toda a história.
(5) *Zé* sabe de toda a história.
(6) Se sabe de toda a história.

- C. (7) Quem *cê* disse que viajou junto?
(8) Quem *Zé* disse que viajou junto?
(9) Quem se disse que foi comigo?

Com tais sentenças, pretendeu-se comparar as formas *cê*, *Zé* e *se* que ocorrem, no grupo A, entre dois segmentos átonos; no grupo B, em início de sentença; e, no grupo C, entre dois segmentos tônicos.

A comparação realizada levou em conta os parâmetros de *intensidade* e *duração* das três formas nos três ambientes

mencionados. Nossa intenção foi examinar se a intensidade e a duração da forma *cê*, que, segundo a hipótese apresentada nos Capítulos 1 e 2, é um clítico, são comparáveis às da partícula *se*, que é, sabidamente, um clítico, ou, às do item *Zé*, que é um monossílabo tônico.

Para realizar essa análise, os dados gravados, num total de noventa, foram editados através do programa SOUND DESIGNER e, posteriormente, analisados por meio do programa de análise de fala WINPITCH1.92, que nos permitiu medir a duração (de agora em diante, D) em milésimo de segundos (ms) e a intensidade (de agora em diante, I) em decibéis (dB).

Os resultados obtidos, extraída a média de D e de I das duas leituras dos cinco informantes, são os seguintes, nos três ambientes considerados:

- A. (1) Como cê falou que a gente devia fazer?
I: 36 dB D: 0,137 ms
- (2) Como Zé falou que era mesmo?
I: 41 dB D: 0,170 ms
- (3) Como se falou que a gente devia pronunciar?
I: 33 dB D: 1,137 ms
- (4) Cê sabe de toda a história.
I: 20 dB D: 0,122 ms
- (5) Zé sabe de toda a história.
I: 27 dB D: 0,217 ms
- (6) Se sabe de toda a história.
I: 18 dB D: 0,117 ms

(7) Quem cê disse que viajou junto?

I: 23 dB D: 0,156 ms

(8) Quem Zé disse que saiu de casa?

I: 26 dB D: 0,196 ms

(9) Quem se disse que foi comigo?

I: 20 dB D: 0,134 ms

Extraindo-se a medida dos valores encontrados nos três ambientes, obtêm-se os resultados seguintes:

(10) SE: I: 23,6 dB

D: 0,129 ms

CÊ: I: 26,3 dB

D: 0,138 ms

ZÉ: I: 31,3 dB

D: 0,194 ms

Estes resultados mostram-nos que, em todos os três ambientes, considerando-se os números absolutos, ocorre um tipo de gradação de maneira que a forma *se* é a que registra menos intensidade e menor duração; em seguida, temos a forma *cê* que recebe mais intensidade e maior duração que *se*, mas apresenta valores abaixo dos obtidos pelo monossílabo *Zé*. As médias registradas em (10) evidenciam essa hierarquia com nitidez. Distancia-se deste padrão apenas os valores relativos à duração de *cê* e *se* no ambiente A que, como se pode ver, são idênticos.

A apreciação integral destes resultados deverá levar em conta ainda as seguintes observações. Em primeiro lugar, discute-se na

literatura que diferenças de duração entre dois segmentos são perceptualmente relevantes (cf. Lehiste (1970)). O debate não é conclusivo, o que se deve em parte à diversidade de metodologias e os objetivos dos experimentos realizados. É possível fazer a generalização, no entanto, de que a obtenção de uma diferença de duração entre dois segmentos de até 20 ms não é relevante do ponto de vista da percepção e pode ser considerada não significativa (de acordo com César Reis, comunicação pessoal). Nesse caso, sendo que a diferença de duração entre *se* e *cê* é, levando-se em conta as médias em (10) de apenas 9 ms, pode-se admitir que, apesar da hierarquia encontrada, a duração da forma *cê* é bastante próxima da do clítico *se*. Se observarmos, por outro lado, a duração do monossílabo *Zé*, comparando-a com a de *cê*, notamos que o primeiro é 56 ms mais longo que o segundo. Este dado é significativo, pois nos permite opor, de um lado *cê* e *se*, e, de outro *Zé*, em relação ao parâmetro da duração. O valor alto obtido por *Zé*, comparado com o das outras formas, é ainda mais significativo quando observamos que, segundo Santos (1987); e Stevens e Klatt (1968), que analisam a duração inerente dos sons fricativos, “os sonoros são aproximadamente 50 ms mais curtos que os surdos” (Santos, *op. cit.*: 14).

Em segundo lugar, consideremos a intensidade: a gradação entre as três formas não é proporcional. A diferença entre *se* e *cê* é de 2,7 dB enquanto que aquela entre *cê* e *Zé* é de 5 dB, o que, de novo, nos permite agrupar as duas primeiras de um lado em oposição à última. Essa diferença entre *cê* e *Zé* poderia ser atribuída a um traço intrínseco relacionado à abertura das vogais, como é sugerido na literatura (Di Christo (1996: 476)). A título de ilustração, num experimento relatado por Lehiste (1970: 120), baseado no inglês americano, as diferenças de intensidade intrínsecas encontradas entre /e/ e /ɛ/ foram de 0,07 dB enquanto que as diferenças entre /e/ e /i/ foram de 3,5 dB. Seria evidentemente

mais confiável levar em conta uma descrição similar feita para as vogais do português do Brasil. Contudo, as diferenças entre *cê* e *Zé* parecem ser muito grandes para poder serem atribuídas à intensidade das vogais envolvidas. Parece-nos, assim, haver a atuação de algum mecanismo da gramática que provoca a redução do monossílabo *cê* (cf. o Capítulo 7 para um desenvolvimento desta idéia).

Na verdade, por outro lado, segundo Di Christo (*op. cit.*: 374), o fator que exerce influência mais forte sobre a duração não são os traços intrínsecos, mas o ambiente em que o segmento se encontra. No caso das vogais, é o modo fonatório da consoante subsequente. Ora, esta variável não trouxe nenhuma influência nos valores encontrados, pois, como se pode verificar nos exemplos testados, as consoantes que seguem as vogais dos monossílabos examinados são idênticas. O papel do ambiente na determinação do parâmetro da intensidade pode ser apreciado, contudo, nos nossos resultados da seguinte maneira: note-se que a intensidade média das três formas é diferenciada nos três ambientes considerados. No ambiente A é de 36,6 dB; no ambiente B, 21,6 dB; e, no ambiente C, obtivemos 23 dB. É interessante observar neste caso o valor relativo alto da intensidade registrado no ambiente em que as formas analisadas estão entre segmentos átonos. Isso também é um indício adicional de que traços intrínsecos não parecem ser responsáveis pelos valores encontrados.

Concluimos, assim, que as três formas analisadas podem ser descritas através de uma gradação em relação aos parâmetros da intensidade e duração. Esta gradação, no entanto, não é inteiramente proporcional: cabe, como vimos, agrupar, de um lado, *se* e *cê*, e, de outro, *Zé*.

É possível ainda objetar que as diferenças entre esta última forma e as outras duas podem ser atribuídas também ao fato de *Zé* ser um nome próprio que, como elemento novo na oração, deve, numa

relação de “interface”, ainda bastante mal compreendida, entre componentes da gramática, receber maior destaque do ponto de vista sonoro, o que poderia ser traduzido pelos valores relativamente altos de intensidade e duração encontrados. Apesar de ser exatamente a diferença entre formas pronominais reduzidas e outros tipos de monossílabos que estamos querendo descrever, tentamos atenuar essa eventual dificuldade incluindo o nome *Zé*, nos textos que foram lidos pelos informantes, em orações anteriores às de (2), (5) e (8) nas quais aparecem as ocorrências de *Zé* que foram analisadas. Pretendeu-se com isso reduzir o impacto de se tratar de um elemento novo no texto. A escolha do nome *Zé* para comparação com as formas pronominais deveu-se ainda a dificuldades metodológicas relativas à elaboração de orações, naturais para os informantes, que permitissem a presença das três formas num mesmo contexto sintático e segmental. As intenções neste caso foram evitar segmentos diferenciados subsequentes às formas analisadas que pudessem, como previsto por Di Christo, mencionado acima, influenciar os valores dos parâmetros examinados.

2. CLÍTICO *VERSUS* FORMA FRACA

Num artigo muito influente, Cardinaletti e Starke (1994) distinguem três tipos de pronomes: fortes, fracos e clíticos. A distinção entre essas classes é baseada numa série de propriedades sintáticas, morfológicas, semânticas e fonéticas e pode ser observada em várias línguas. O quadro seguinte sintetiza essas propriedades (cf: Cardinaletti e Starke (*op. cit.*: 22)):

(11)	'morphology	choice	distribution		interpretation
	+ reduced		in PF at SS	*coord.	no range
clitic	1	1	+	+	+
weak	2	2	+	+	+
strong	3	3	-	-	-
	prosody			X-bar	
	reduction rules	no word-stress		X°	
clitic	+	+		+	
weak	+	-		-	
strong	-	-		-	

Deixando de lado a farta exemplificação utilizada pelos autores (cuja verificação fica a cargo do leitor), o quadro acima informa-nos o seguinte: (i) do ponto de vista morfológico, as três formas obedecem a um tipo de gradação de deficiência fonética; (ii) do ponto de vista da escolha ou uso, quando possível, as formas mais deficientes são empregadas preferencialmente; (iii) em relação à distribuição, contrariamente às formas fortes, os clíticos e as formas fracas não podem ocorrer em posições deslocadas tais como as geradas em sentenças clivadas, com posposição do sujeito, topicalizadas e outras, e não podem ser coordenadas nem modificadas; (iv) em relação à interpretação, as formas fortes devem sempre ser associadas a uma *extensão* própria; (v) do ponto de vista da prosódia, os clíticos não têm acento de palavras contrariamente às formas fracas e fortes; clíticos e formas fracas podem ser reestruturados, prosodicamente, como ocorre, por exemplo, em processos fonológicos; e enfim, (vi) do ponto de vista da Teoria X-barra, que faz parte de modelos gerativistas recentes,

os clíticos podem ser analisados como núcleos, isto é, X^o, que se incorporam a outros núcleos, enquanto que tal análise não é possível para as formas fracas e fortes.

Apreciados esses critérios em conjunto, vê-se que a distinção entre, de um lado, formas fortes e, de outro, formas fracas e clíticos, é mais bem caracterizada que a distinção entre clíticos e formas fracas, que nos parece bem menos evidente. Os próprios autores reconhecem que:

All properties separating deficient from strong pronouns uniformly hold of both clitic and weak pronouns. (cf. Cardinaletti e Starke (*op. cit.*: 20))

Certas diferenças distribucionais e prosódicas distinguem, entretanto, as formas fracas e os clíticos. As propriedades distribucionais dizem respeito ao fato de clíticos aparecem sempre ligados aos verbos ou a outras categorias enquanto que as formas que os autores classificam como fracas podem se alocar, aparentemente, em posição de especificador, no sentido da Teoria X-barras. Essa característica dos clíticos pronominais tem sido tratada, como é sabido, através da hipótese de que os clíticos são núcleos, isto é, X^o que se incorporam aos verbos (cf. Kayne (1975, 1991); Rizzi (1982); e outros).

Visando a distinguir as formas fracas dos clíticos, Cardinaletti e Starke propõem que as primeiras são projeções máximas, isto é, XP. Essa hipótese já havia sido adotada por Halpern e Fontana (1992); Fontana (1993), para distinguir dois tipos de clíticos em várias línguas e foi utilizada por nós, como comentaremos adiante, na análise de fatos da história da cliticização em português (cf. Capítulo 7).

Em relação às propriedades prosódicas, o clítico se distingue das formas fracas pelo fato de não poder receber acento de palavra, o que não ocorre com essas últimas.

Há ainda um critério que, segundo os autores, é capaz de distinguir as duas formas e que diz respeito ao uso. Quando as duas formas são possíveis num mesmo contexto, o clítico toma a precedência sobre a forma fraca. Isto é válido também em relação às formas fortes e fracas: são as últimas que são escolhidas. Este fato, aparentemente um efeito do “Avoid Pronoun Principle” (cf. Chomsky (1981)), é analisado pelos autores através de uma restrição, ligada à economia das derivações, que é a seguinte:

- (12) Minimise α , onde α : estrutura, cadeias, “links”, movimento manifesto etc.

Vamos a seguir discutir um pouco mais esses três critérios que distinguem clíticos de formas fracas, começando por esse último e pensando nos exemplos do português do Brasil que discutimos nos capítulos anteriores. Os efeitos do “Avoid Pronoun Principle” (ou de alguma versão deste princípio) – que, intuitivamente, nos parece dizer muito mais acerca de propriedades de performance – não são universais e, como já foi discutido por Nascimento (1984: 25) e Ferreira (2000), não se verificam no português do Brasil. Em relação ao comportamento da forma *cê*, esse princípio não parece ser determinante. *Cê* alterna-se com *você* na posição sujeito e a escolha de uma ou outra forma parece não se deixar determinar exclusivamente por fatores estruturais, mas estar associada a parâmetros de outra natureza como, por exemplo, (i) o registro da língua, isto é, se mais formal ou informal: a forma *você* toma a precedência sobre *cê* em registros mais formais e esta última, quando usada para exprimir indeterminação do sujeito, é mais comum em registros informais enquanto que, nessa mesma acepção, prefere-se empregar o clítico *se* em registros formais; (ii) a faixa etária; e (iii) a referência definida ou indefinida do sujeito: de acordo com a análise desenvolvida no Capítulo 3, quando a

referência é definida, os jovens e os medianos preferem usar a variante *você* (67%); quando é indefinida, preferem *cê* (75%). O inverso observa-se no comportamento dos velhos: quando a referência é indefinida, a preferência é por *você* (53%); quando definida, a preferência é por *cê* (65%) (ver também Alves (1998)). A conclusão é que a escolha das formas pronominais do português do Brasil não parece ser determinada tão-somente por fatores estruturais e que este critério não é decisivo na classificação de *cê* como clítico ou como forma fraca.

Já as propriedades distribucionais e prosódicas mencionadas parecem ser úteis em relação à classificação de *cê*. Assim, o fato de os clíticos serem analisados como X^o, mas não as formas fracas – o que é interno à teoria – tem um papel relevante na descrição de *cê*. Como foi discutido no Capítulo 2, a forma *cê* aparece, sistematicamente, precedendo a negação:

(13) *Cê não vai conseguir esse ingresso!*

O exemplo (13) nos mostra que *cê* não está incorporado diretamente ao verbo já que a negação se interpõe entre os dois constituintes. Por outro lado, como já se analisou em Martins (1994); Fontana (1993); Taylor (1990); Vitral (2001) e outros, o fenômeno da *interpolação* de constituintes entre um clítico e o verbo foi produtivo no português arcaico e moderno e em outras línguas como o grego antigo e o espanhol medieval, deixando de existir na nossa língua no século XIX. Observem-se os exemplos de Martins (1994), (14 a-d) e o apresentado no Capítulo 7:

- (14) a. com entençom de a nuca reuogar (1432).
b. ...sse dona Sancha quitaua desse herdamoto e sse delle saya (1295).
c. ...ca **lho** nō tolhya (1299).

- d. ...E renõciamos (...) a os nõ Alegar (1406).
- e. Enquanto se não conhecem bem instruídas (séc. XVI).

Observe-se que, no português arcaico, além da negação, outros constituintes como sintagmas preposicionados, advérbios e outros podiam se interpor entre o clítico e o verbo.

No Capítulo 3, propomos que a cliticização pode ser analisada como um processo sujeito a estágios específicos. Esta visão dos fatos é compatível com o que se observa na história da cliticização portuguesa já que a interpolação deixou de ser produtiva no século XIX, e também com a distribuição da forma *cê* em orações como (14). Ora, se *cê* é um novo clítico, espera-se que essa forma deva trilhar estágios anteriores da cliticização que admitiam a ocorrência da interpolação. Mas como descrever esse processo de cliticização? Adotando a proposta de Halpern e Fontana (*op. cit.*) e Fontana (1997) de que os clíticos sofrem um processo de mudança que os faz evoluírem de projeções máximas XP para núcleos X. *Cê* sendo então uma projeção máxima, explica-se o fato de ocupar a posição sujeito como em (13). Podemos dizer ainda que categorias XP têm acento de palavra, mas não as categorias X°.

As duas propriedades examinadas, isto é, ser projeção máxima e receber acento de palavra, parecem levar-nos a classificar *cê*, de acordo com a terminologia de Cardinaletti e Starke, como uma forma fraca. Mas se considerarmos todas as outras propriedades arroladas pelos autores, que aparecem no quadro em (11) e também a análise acústica relatada na Seção 1 deste capítulo, há argumentos para classificar *cê* como um clítico. O problema dessas evidências conflitantes reside, talvez, na própria tentativa de Cardinaletti e Starke de estabelecer três classes discretas de pronomes. Ora, a distinção de clítico e forma fraca como duas classes distintas não é tão clara, porque grande parte dos traços que caracterizam uma

suposta classe serve também para caracterizar a outra. Em outras palavras, a interseção entre os conjuntos de traços que definem as duas classes é, aparentemente, muito grande.

Para resolver esse problema, uma opção é, como propomos, analisar a cliticização como um processo. Os próprios Cardinaletti e Starke admitem que há um tipo de hierarquia entre as três formas, deixando evidente um tipo de “ranking in deficiency”. No caso de *cê*, a proposta de que essa forma participa de algum tipo de processo é bastante evidente, pois, como foi analisado no Capítulo 1, *cê* parece ser um estágio de um processo que inclui o pronome *você* e a forma de tratamento *vossa mercê*.

Como vimos, de acordo com a literatura sobre gramaticalização (cf. Heine *et al.* (1991); Hopper e Traugott (1993)), há, nesses processos, uma “perda gradativa de substância fônica e semântica”. Descrever essas “perdas” de forma explícita tem sido um desafio para as análises que lidam com essa noção. Contudo, na análise de Cardinaletti e Starke, há observações sobre a interpretação das formas pronominais analisadas que podem ajudar a tornar mais objetiva nossa visão de fatos.

Segundo os autores, certas propriedades semânticas distinguem as formas fortes das reduzidas. Assim, (i) os pronomes fortes, mas não as formas reduzidas, são capazes de introduzir um referente discursivo não mencionado anteriormente; (ii) apenas as formas reduzidas podem funcionar como expletivo sujeito ou dativo; (iii) os pronomes fortes só designam referentes [+ humano]. Esta última propriedade, contudo, não é verificada em todas as línguas: o inglês e o português do Brasil, por exemplo, são exceções; e, por fim, (iv) somente as formas deficientes participam, como sujeito, de construções impessoais. Os autores propõem uma generalização a fim de captar o essencial das propriedades que acabamos de resumir. De acordo com eles (p. 9), então:

(...) a strong element is incapable of being a semantically vacuous subject, it must be referential. Deficient elements on the other hand do not need to be referential and can be semantic dummies.

Esta generalização é compatível com a perda de “substância semântica” de que falam os autores que trabalham com a noção de gramaticalização. Podemos dizer, assim, que nos processos de cliticização, incluídos no fenômeno mais amplo da gramaticalização, há, paralelamente à redução fonética, uma mudança semântica que implica perda de referencialidade. Essa descrição dos fatos é também adequada em relação ao que se observa acerca do comportamento da forma *cê* (cf. Capítulo 2; Capítulo 6; Alves (1998)). Esta forma é usada para exprimir indeterminação do sujeito, como o mostra o exemplo (15a) a seguir e constata-se igualmente o uso dessa forma como expletivo na posição sujeito, o que pode ser observado em (15b):

(15) a. Na primeira aula [de helicóptero], *cê* fica assim,
meiô *bambinha*, mas depois é ótimo. (ex. 51 de
Alves (1998: 68))

b. Em BH, *cê* tem muita cafeteria boa.

Por outro lado, é evidente que, mesmo analisando a cliticização como um processo, isso não implica que não se possa, como Cardinaletti e Starke o fazem, destacar etapas discretas nesse processo. É preciso observar, no entanto, que os processos de cliticização parecem incluir mais do que as três etapas destacadas por esses dois autores. Num trabalho muito conhecido, Zwicky (1977) (cf. também Klavans (1980)) propõe que existe uma divisão fundamental entre itens que são “simple clitics” e “special clitics”. De acordo com este autor, “simple clitics” são elementos que são

clíticos apenas do ponto de vista fonético, mas não do ponto de vista sintático, enquanto que os “special clitics” são itens que, além de serem clíticos do ponto de vista fonético, necessitam de regras sintáticas específicas para justificar sua distribuição. As formas fracas de Cardinaletti e Starke parecem ser classificadas como “special clitics” devido ao fato de não terem a mesma distribuição das fortes correspondentes. Como exemplos de “simple clitics”, Zwicky propõe as reduções de verbos auxiliares em inglês:

- (16) a. She’s gone = she is/has gone.
b. They’ve all seen this movie before = they have all seen this movie before.

Pensando agora na possibilidade de haver “simple clitics” pronominais na nossa língua, ocorreu-nos que as formas reduzidas de terceira pessoa *el/és/eis*, analisadas por Corrêa (1998), podem ser classificadas como tal. Além de reduzidas, essas formas têm comportamento de clítico, não sendo possível, por exemplo, aparecer coordenadas com um pronome tônico, o que pode ser atribuído ao fato de não podermos coordenar constituintes desiguais. Entretanto, essas formas distinguem-se de *cê*, pois, contrariamente a esta última, podem aparecer na posição objeto e também na posição de complemento de preposição que são posições em que, no português do Brasil, ocorrem os pronomes tônicos. Observe-se o exemplo de Corrêa (*op. cit.*: 1):

- (17) a. aí eu falei pra és que aquilo estava super errado.

De acordo com Corrêa, que trabalhou com um *corpus* de entrevistas, o percentual de ocorrência de formas reduzidas de terceira pessoa na posição sujeito é de 52% enquanto que o percentual de ocorrência dessa forma na posição objeto é de 28 %.

Já no caso de *cê*, segundo os capítulos 3 e 10, o percentual de ocorrência dessa forma na posição sujeito é de 94,4% e na posição objeto é de 0,3%, o que atesta, de maneira clara, a diferença entre as formas reduzidas de terceira e segunda pessoa.

Já que as formas reduzidas de terceira pessoa têm uma distribuição comparável à dos pronomes tônicos, pode-se supor que elas se encaixam na definição de “simple clitics” de Zwicky.

Se quisermos, portanto, distinguir etapas no processo de cliticização, parece necessário diferenciar não apenas três etapas, como fazem Cardinaletti e Starke, mas pelo menos quatro com a inclusão da etapa “simple clitic”. Será preciso levar em conta também a etapa da afixação que, como previsto nos processos de gramaticalização, é o estágio seguinte após a cliticização. Em relação a este último ponto, é bem conhecido o exemplo da formação do futuro sintético românico no qual a forma verbal *habeo* – de um verbo com significado lexical completo presente na perífrase latina *amare habeo*, da qual o futuro deriva – passa a ser um marcador de tempo e pessoa gramaticais nas línguas românicas (cf. Cohen (1988: 44)).

Nossa conclusão é que a forma *cê* pode ser descrita, na perspectiva “classificatória” de Cardinaletti e Starke, como um clítico que, estruturalmente, é uma projeção máxima, ou então como uma forma fraca. Aliás, os termos “clítico” e “forma fraca”, no nosso modo de ver, não têm nenhum estatuto teórico explicativo. Parece-nos, assim, que a perspectiva mais interessante e plena de consequências teóricas é compreender o processo de cliticização no qual as formas pronominais reduzidas estão envolvidas. Alguns passos nesse sentido estão sendo dados nos trabalhos reunidos neste livro.

CAPÍTULO 5

MAIS UM PRONOME EM PROCESSO DE CLITICIZAÇÃO: O PAR ELES/ES¹

O paradigma pronominal do português brasileiro apresenta, para a terceira pessoa, formas que variam em número e gênero, a saber, *ele(s)*, *ela(s)*. Itens que podem ser identificados como redução fonética desses pronomes são documentados na linguagem oral, embora ausentes das gramáticas normativas. São eles: *el*, *éa*, *éas*, *es*, respectivamente.

Neste capítulo examinaremos a redução fonética dos pronomes de terceira pessoa, apresentando dados extraídos de uma amostra de entrevistas com informantes da cidade de Belo Horizonte. Nosso propósito é apresentar evidências de natureza quantitativa de que as categorias de gênero, número e pessoa são identificadas separadamente, funcionando como parâmetros para mostrar como um item gramatical pode se tornar ainda mais gramatical.

1. OS PRONOMES

Corrêa (1998) analisa os seguintes pares de pronomes de terceira pessoa e identifica um processo de perda morfofonológica.

¹ Versão deste texto foi apresentada no Congresso da ABRALIN, Gramado, 1998, por Jânia Ramos e Lucas Teles Correa.

- (1) a. *e el falou que eu não podia mais dar aula.* (E5)
 b. *e ele falou que eu não podia mais dar aula.*
- (2) a. *o albergue 'tava lotado, eis não registravam mais ninguém.* (E2)
 b. *o albergue 'tava lotado, eles não registravam mais ninguém.*
- (3) a. *se não fosse eu, éa tinha matado a colega.* (E9)
 b. *se não fosse eu, ela tinha matado a colega.*

A ocorrência desse fenômeno já tinha sido referida por gramáticos e filólogos. Estes autores identificam as formas fracas como fala popular, o que é relevante para uma abordagem como a que se desenvolverá aqui. Joaquim J. Nunes (1951: 235) afirma que em português *ele* e *el* estão presentes e se diferenciam pela queda de [e] nesse último, o que teria ocorrido *devido à próclise*. Coutinho (1954: 255), ao referir-se à terceira pessoa pronominal, afirma que *no antigo Português e na língua popular de além-mar aparece a forma apocopada e*. Vasconcelos (1959: 49) afirma que *os pronomes tomam outras formas: el, plural eis, le, lhe*.

Uma informação muito relevante é dada por Matos e Silva (1989: 220-221), ao comentar que a forma apocopada *el* “ultrapassa de muito a forma *ele*” no português arcaico e que, nesse período, a forma reduzida pronominal não tinha comportamento clítico, pois *el tem posição acentuada na frase*. A passagem do estatuto de não-clítico para clítico no PB moderno definiria o processo em análise aqui.

2. A DISTRIBUIÇÃO

Posição sujeito:

- (4) a. *eles acham que todas as mulheres.* (E2)
 b. *ês acham que todas as mulheres.* (E2)

Posição objeto:

- (5) a. Nós vimos eles.
a. ‘*Nós vimos es.
b. Nós vimos eles longe.
b. ‘Nós vimos es longe.

Pós-verbal:

- (6) a. Eles foram os culpados
a. ‘Es foi os culpado.
b. Foram eles.
b. ‘?Foram es.
- (7) a. Aí eles chegam para mim e dizem...
es
b. Aí chega es pra mim ontem e diz...
- (8) a. Lá vem eles com mentiras.
b. Lá vem es com mentira.

Preposto:

- (9) a. Eles, Maria não viu.
b. ?Es, Maria num viu.

Resposta à pergunta:

- (10) – Quem vai sair?
– Eles
* es

Objeto de preposição:

- (11) Eu falei pra eles.
... pra es.

Participação em coordenação:

- (12) a. Os meninos e eles podem votar contra.
b. * Os meninos e es podem votar contra.

Com base nos contrastes acima, principalmente (6, 10 e 12), é possível propor que esteja ocorrendo o seguinte processo gramatical:

- (13) eles > es

Em resumo, existe no português brasileiro atual uma contraparte foneticamente reduzida do pronome de terceira pessoa, *eles*. Vejamos, a seguir, alguns comportamentos desse item, que evidenciam seu caráter afixal.

3. O GÊNERO

No português brasileiro, normalmente, os pronomes plenos concordam em gênero com o antecedente.

- (14) Encontrei **as crianças**. **Elas** estavam cansadas.

Com o pronome foneticamente reduzido o gênero do NP antecedente no discurso NPs, quer seja masculino ou feminino, é retomado pelo masculino. E com número plural.

(15) *tem a turma que você fica rezando pra ês te chamá pra dançá.* (E18)

(16) *o povão es fazem assim mesmo.*

Corrêa (1998) analisa duas correlações: a primeira entre a redução pronominal e o gênero do pronomes, e a segunda entre o gênero do pronome e o tipo de referência. A amostra analisada compõe-se de 1.463 dados.

Tabela 1
DISTRIBUIÇÃO DA FORMA REDUZIDA DO PRONOME
CONFORME O GÊNERO DO ANTECEDENTE
(*apud* Corrêa, 1998: 54)

	N	%	P
Masculino	368/917	40	.66
Feminino	39/442	9	.20

Ao contrário do feminino, o gênero masculino favorece as formas reduzidas: a frequência é de 40%, muito superior à do feminino, que é de 9%. A correlação entre gênero do pronome e tipo de referência mostra que o gênero masculino favorece a forma reduzida com referência indefinida. Aqui os traços de gênero [masculino] do pronome pleno não seriam propriamente gênero não-feminino, mas sim uma forma não marcada de gênero (cf. Martin, 1968).

Tabela 2
**DISTRIBUIÇÃO DAS FORMAS REDUZIDAS
 PRONOMINAIS: MASCULINO/SUJEITO, CONFORME O
 TIPO DE REFERÊNCIA.**
 (adaptada de Corrêa 1998: 55)

	N	%	P
Definida	236/629	38	.45
Indefinida	100/162	62	.67

As ocorrências com referência indefinida atingem o índice de 67%. Quando se leva em conta apenas as ocorrências em posição sujeito, tem-se um universo de 161 ocorrências e um percentual de 59%. Nesse último conjunto observa-se um fenômeno muito curioso. Nem sempre há concordância entre sujeito e verbo: o sujeito é plural e o verbo é singular.

Tal “discordância” de número parece estar diretamente relacionada à do gênero e tipo de referência, conforme veremos a seguir.

Em (17b), o pronome pleno favorece o verbo no plural, havendo concordância de número. Com o pronome foneticamente reduzido, entretanto, o verbo fica na terceira pessoa do singular, não havendo concordância de número com o verbo (17a):

- (17) a. **es inventa** um bocado de coisa. (E42)
 b. os próprios industriais, **eles num ampliam**. (E21)

Estes dados não são esporádicos, mas sistemáticos. O resultado da análise quantitativa aparece a seguir.

Tabela 3
**PORCENTAGEM DE PRONOMES NÃO-FORTES,
 CONFORME O NÚMERO DO VERBO**
 (adaptado de Corrêa, 1998, tabela 6)

	Nº	%	Prob.
Sujeito e verbo no plural	82	40	.56
Sujeito plural e verbo no singular	66	65	.74

Estes resultados reforçam o estatuto quase afixal de *es*. Parece que o pronome não forte “compensa” a desinência verbal. Em outras palavras, ele desempenharia a função da desinência.

Do ponto de vista morfossintático, dados do tipo (17a) podem ser descritos como ocorrências em que uma forma gramatical apresenta, ao mesmo tempo, gênero gramatical e referência indefinida.

A análise dos 66 casos desse tipo mostra que o verbo se realiza preferencialmente no presente. Estes resultados levam à seguinte questão: *haveria alguma correlação entre ser morfofonologicamente reduzido e ocorrer em contextos em que Tense é [-passado]*? Para buscar uma resposta, consideremos a proposta de Ritter (1995) sobre a correlação entre realização de pronome [-específico] e tempo presente.

4. A PROPOSTA

Ritter (1995:436) observa que, no hebraico, o presente do indicativo de terceira pessoa licencia sujeitos nulos indefinidos, não específicos, mas não sujeitos nulos específicos. A autora atribui essa restrição ao fato de na estrutura sentencial, no presente, a categoria mais alta é TP; já no passado é DP. Apenas nesta última,

a categoria D poderia checar o traço [+específico] do DP sujeito. O traço D diz respeito à pessoa. Uma evidência seria a ausência de desinência de pessoa no presente no hebraico, apenas as desinências de gênero e número permanecem.

O português brasileiro apresenta certo paralelismo com o hebraico. Conforme se lê nas gramáticas normativas, a chamada terceira do plural, quando o DP sujeito tem referência indefinida, pode ter realização nula.

(18) Bateram na porta. (Alguém indefinido bateu na porta.)

Nos termos de Ritter (1995: 436), o sujeito nulo decorreria da incompatibilidade entre os traços pronominais indefinidos e os traços de pessoa de D, pois o traço pessoa seria sempre definido.

A consulta à Tabela 2 vai mostrar que existe no português brasileiro atual uma tendência de preenchimento do sujeito com o item *es*, mesmo quando este tem referência indefinida.² O que isso significa? Significa que, embora não seja nulo, *es* é incapaz de checar o traço pessoa. Uma evidência disso é o fato de *es* ocorrer na posição de sujeito quando o verbo está no presente, o que significa ser checado por TP, e não por D. Outra evidência é que *es* não retoma anaforicamente seu antecedente, daí não ser também marcado com gênero. *Es* seria marcado apenas com número.

Ora, essa última constatação leva-nos a retomar a Tabela 3. Nela vemos que *es* ocorre com o verbo na terceira pessoa do singular. Se *es* marca o número e se a desinência de terceira “pessoa” também marca número, teríamos aí uma redundância gramatical. Desse

² Duarte (2000: 23) apresenta um levantamento em que 16% das ocorrências de sujeitos indeterminados aparecem preenchidos pelo pronome “eles”. A autora não distingue formas plenas e formas foneticamente reduzidas.

ponto de vista, a ausência de desinência verbal de plural com sujeito *es* poderia ser interpretada como indicação do caráter afixal de *es*, conforme referido acima.

Kato (1999) argumenta que tanto as desinências verbais quanto os pronomes fracos teriam caráter pronominal. Assim, o surgimento de pronomes fracos no português brasileiro estaria compensando a perda da desinência morfológica verbal. Veja-se que os resultados da Tabela 3 constituem uma nova evidência a favor desta hipótese.

Com base nos resultados apresentados neste capítulo é possível também visualizar o percurso de um item – o pronome *eles* – que, ao tornar-se mais gramatical, perdeu os traços de pessoa e gênero. Em vista disso, a escala de definitude, que congrega informações de natureza semântica e sintática, pode ser assim reescrita:

- (19) a. referência definida > referência indefinida
b. +pessoa, +gênero, +número > -pessoa, +gênero,
+número > -pessoa, -gênero, +número.

Passemos agora ao próximo capítulo em que se retoma a escala (19).

CAPÍTULO 6

GRAMATICALIZAÇÃO DE 'VOCÊ': UM CASO DE PERDA DE CONTEÚDO SEMÂNTICO¹

1. INTRODUÇÃO

A maioria dos autores que trabalha com a noção de gramaticalização concorda com a idéia de que esse tipo de processo implica “perda de conteúdo semântico” e também “perda de substância fônica”. Observe-se, por exemplo, o seguinte trecho de Heine e Reh (1984) (*apud* Traugott e Heine (1991)), que definem gramaticalização como “an evolution where by linguistic units lose in semantic complexity... and phonetic substance...”. Visando a descrever este fenômeno de forma mais explícita, Hopper e Traugott (1993) propõem, como já vimos, que o processo de gramaticalização obedece às seguintes etapas:

- (1) item com significado lexical > item gramatical > clítico > afixo

Neste capítulo, pretendemos examinar a natureza da mudança de estatuto de morfemas descrita pelas três primeiras etapas de (1). Buscaremos responder à seguinte pergunta: Como tratar a “perda de

¹ Publicado como artigo na *Revista de Lingüística e Filologia da USP*, 1999, por Vitral, L. e Ramos, J.

conteúdo semântico” alegada pelos autores acima? Vamos considerar o caso da gramaticalização de *você*, que historicamente se inicia com a expressão lexical *vossa mercê* (= favor vosso > pessoa que presta um favor) e chega até o clítico indeterminador *cê*. Em que sentido é possível afirmar que o conteúdo do clítico indeterminador é sem significado enquanto que a capacidade de denotar da expressão *vossa mercê* dispõe de significado? Esta descrição dos fatos, que nos parece intuitivamente adequada, não é, como deixa entender nosso breve resumo, suficientemente explícita.

Com vistas a explicitá-la, faremos inicialmente uma breve retrospectiva de um estudo descritivo das formas *você/ocê/cê* no dialeto mineiro, de modo a introduzir as noções de referencialidade virtual, referencialidade real e não-referencialidade, com o propósito de demonstrar que falar em perda de conteúdo semântico do item durante o processo de gramaticalização é inadequado, visto que nem todo item gramatical é necessariamente menos referencial. Trata-se, portanto, de uma alteração na natureza da referencialidade. Em seguida, utilizando ainda os dados do estudo acima referido, vamos apontar um estágio ainda não contemplado nas descrições correntes do processo de gramaticalização: qual seja, a etapa de formação de expletivos. Por fim, argumentamos a favor de que o ciclo em (1) seja revisto, de modo a evitar mistura de elementos de níveis gramaticais distintos.

2. GRAMATICALIZAÇÃO DO VOCÊ: UMA BREVE RETROSPECTIVA

No Capítulo 2, apresentamos argumentos a favor de que o estatuto gramatical do item *você* sofreu alteração, tendo passado de pronome a clítico (cf. também Capítulo 1). Concomitantemente, a forma *você* passa a concorrer com *cê*. Assim, *você* seria um pronome e *cê* seria um clítico. Teria havido aí, portanto, um processo de cliticização que, nos termos de Hopper e Traugott (1993:132), faria parte de um processo mais geral de gramaticalização.

Num estudo variacionista das formas *você* e *cê*, observamos no Capítulo 3, com base no tempo aparente, certa tendência à especialização no uso destas formas: os jovens tendem a usar *cê* com referência menos definida enquanto que os velhos o usam com referência definida, respectivamente, 72% e 47% (cf. Tabela 4, Capítulo 3; e Capítulo 10). O fato de haver esta tendência à especialização não invalida o tratamento variacionista, visto que não há uso categórico do *cê* quer com referência definida quer com referência indefinida nem no *corpus* geral, respectivamente, 188/342 e 154/342, nem na fala dos informantes, tornada individualmente, e nem num mesmo grupo etário (cf. Tabela 4, Capítulo 3). Por outro lado, o fato de haver tendência à especialização não impede de se identificar um processo de gramaticalização em andamento. Considerando o período de tempo global do processo de gramaticalização, os períodos de especialização indicam muito mais etapas do que interrupção do processo. Portanto, identificar períodos de especialização não significa identificar momentos potenciais de interrupção do processo de mudança lingüística.

O processo de cliticização é descrito por Hopper e Traugott (*op. cit.*:132), apresentando duas propriedades. A primeira é a perda de autonomia lexical, sendo mais contextualmente dependente. Em relação a *você* e *cê*, de fato, o item cliticizado é fonologicamente mais reduzido (duas sílabas > uma sílaba) e sem independência fonológica, o que fica evidenciado, por exemplo, em (2), pela impossibilidade de *cê* aparecer sozinho como resposta a uma pergunta. Outros dados, levando à mesma conclusão, são apresentados em Dutra (1991) (cf. também os capítulos 2 e 3).

(2) Quem vai sair? (ex. 18, de Dutra, artigo inédito).

– (V)ocê!

– *Cê

A segunda propriedade destacada por Hopper e Traugott é uma tendência a “uma significação mais geral” (p. 133). Por significação

mais geral entenda-se um percurso que poderia ser descrito através das seguintes etapas: +def/+ref > -def/+ref > -def/-ref (cf. Croft, 1990: *apud* Hopper e Traugott, 1993:157).

A análise de enunciados com *cê* mostra que este ora ocorre com referência definida, como em (3), ora indefinida, como em (4) (cf. também Alves(1998)):

(3) de que lado *cê* tá? (ex. 1, Ramos, 1997)

(4) *Cê* pode visitar sempre esse local. (ex. 10d, *idem*)
(= Pode-se visitar sempre esse local)

Se for correta a correlação entre perda de substância fonética e perda de conteúdo semântico no ciclo de gramaticalização, é de se esperar que alguma de suas formas seja usada não referencialmente. A candidata preferencial seria a forma clítica.

Tais predições implicam o reconhecimento de que uma nova etapa precisa ser incluída na descrição dos ciclos de gramaticalização: o momento da formação de expletivos. Por definição, os expletivos são itens não referenciais, isto é, sem traços semânticos em sua matriz lexical. Em vista do exposto, propomos que a descrição que aparece em (5a) seja estendida, tal como em (5b), de modo a incluir uma nova etapa, aquela de formação de expletivos. Desse modo teremos a hierarquia semântica (5c).

- (5) a. item lexical > item gramatical
- b. item lexical > item funcional > expletivo
- c. item referencial > item não referencial

3. JUSTIFICANDO A HIERARQUIA (5C)

A suposição de que haveria uma simetria entre perda de substância fonética e perda de conteúdo semântico tem estado

presente na literatura sobre gramaticalização. Com base nessa suposição, fizemos predições sobre um uso não referencial do item *cê*, que seria o estágio mais avançado do processo iniciado com *vossa mercê* no século XV da história do português (Said Ali, 1976).

Para obter um quadro em que se descreva tal processo diacrônico, propusemos (5c). Nesta seção vamos retomar a noção de referencialidade, de modo a distinguir dois tipos de referência e, assim, refinar a hierarquia representada em (5c).

3.1. Tipos de referência: virtual e real

Milner (1982:10) introduz uma distinção entre referência virtual e referência real que corresponde, parcialmente, à distinção fregueana “Bedeutung/Sinn”: “Le segment de réalité associé à une séquence est sa référence actuelle; l’ensemble de conditions caractérisant une unité lexicale est sa référence virtuelle”.

Consideremos, a título de exemplificação, a distinção entre nome comum e pronome pessoal. A todo nome é possível associar, fora do enunciado, uma extensão, a classe de seres designados por este nome. Por outro lado, não se pode associar a um pronome pessoal uma classe bem definida de seres a qual ele se refere. Um pronome que designa qualquer elemento não possui uma *extensão* fora do enunciado. Para analisar o comportamento semântico de um pronome pessoal, é preciso fazer intervir o ato de fala do qual ele participa. Isto quer dizer que um pronome pode denotar, isto é, seleccionar um referente no universo do discurso, mas ele o faz de maneira indireta: a identidade do referente, que não é determinada pelos traços gramaticais do pronome, deve ser recuperada levando-se em conta as indicações do contexto.

Esta descrição nos permite propor a seguinte caracterização, segundo Rouveret (1987): (i) os nomes comuns são termos dotados de referência virtual e de referência real próprias; e (ii) os pronomes

são termos desprovidos de referência virtual, mas dotados, no enunciado em que aparecem, de referência real.

3.2. Refinando a hierarquia (5c)

Com base na distinção entre referência real e referência virtual, é possível refinar a hierarquia (5c). Tem-se, portanto:

(6) referência virtual > referência real > sem referência
[+def > -def]

Se o percurso item lexical > item gramatical for descrito como (6), poderemos descrever a “perda de conteúdo semântico” como perda de traços semânticos que caracterizam a referência virtual (i.e. traços semânticos que permitem ao item selecionar diretamente um referente no universo do discurso).

4. AVALIANDO AS PREDIÇÕES

Na Seção 2, com base na suposição de que haveria harmonia entre perda de substância fonética e perda de substância semântica, fizemos uma predição de que se houvesse uso não referencial dos itens do ciclo de gramaticalização *vossa mercê* > *cê*, o candidato preferencial seria a forma foneticamente mais reduzida, no caso, *cê*. Nesta seção, vamos apontar os resultados da investigação de uso de formas não referenciais no português do Brasil.

4.1. Expletivos

Itens usados não referencialmente são, comumente, identificados como expletivos. Tais itens não contêm traços semânticos. No inglês, tem-se *it* e *there*, como em (7a, b), no francês, tem-se *il*,

como em (8). Há também expletivos em outras línguas, *es* no alemão, *er* no holandês etc.

- (7) a. It seems that Mary is honest.
- b. There are many boys here.

(8) Il pleut maintenant.

Os expletivos, por não conterem traços semânticos, não podem ser interpretados, buscando-se um antecedente no discurso ou na frase, contrariamente aos pronomes pessoais. As posições preferenciais para a ocorrência de expletivos, como *se* sabe, são sentenças existenciais.

Diferentemente do que ocorre em outras línguas, não temos itens expletivos específicos no português. Entretanto, Duarte (1997) mostra que há usos expletivos de formas pronominais no português. Vejam-se os exemplos:

- (9) a. Todas as lojas que você tem aqui nos grandes bairros.
 (ex. 10 de Duarte (1997)).
- b. Em Kioto você tem aquela confusão nas ruas.
- c. Em Buenos Aires você tem confeitarias.

Nos exemplos acima, o item *você* não pode ser interpretado como *possuidor*: ele aparece numa posição não temática e sua presença não pode ser explicada como resultante de movimento a partir de outra posição sintática da sentença. Não pode também ser analisada como vocativo, por não ter recebido entoação marcada.

O uso de *você* em (9) pode representar um problema para a correlação estrita entre perda de substância fonética e perda de informação semântica, na medida em que não é a forma foneticamente mais reduzida, isto é, a clítica, a que está sendo

usada como expletivo. Por outro lado, é possível encontrar também o uso de *cê* em contextos existenciais como (9). Para verificar então se o uso de *você* como expletivo constitui um problema para a suposição corrente na literatura de que haveria correspondência biunívoca entre perda de conteúdo semântico e perda de substância fonética, será preciso investigar, de forma comparativa, a frequência dos usos de *cê* e *você* nestes ambientes.

Podemos concluir que, se aceitarmos que a hierarquia em (6) descreve adequadamente o processo de gramaticalização no que concerne à parte do conteúdo, teremos de aceitar que o expletivo constitui a fase final deste processo, o que nos leva a considerar que a perda de conteúdo semântico descreve, na verdade, o percurso [item lexical > expletivo] e não o percurso [item lexical > item gramatical].

Tais fatos levam-nos a propor uma revisão da hierarquia em (1), com o propósito de desfazer uma mistura de níveis de análise, que reúne informações de natureza distinta: referencialidade e realização morfofonética. Esta tarefa deverá ser buscada na seqüência de nosso trabalho.

CAPÍTULO 7
A INTERPOLAÇÃO DE SE E SUAS CONSEQÜÊNCIAS
PARA A TEORIA DA CLITICIZAÇÃO¹

1. INTRODUÇÃO

Como é discutido neste livro, itens do português brasileiro, como as formas *cê* e *num*, são identificados e analisados como clíticos. *Cê*, na nossa análise, é o penúltimo estágio de um processo de gramaticalização que obedece às etapas seguintes:

- (1) a. item lexical > b. item gramatical > c. clítico > d. afixo (cf. Hopper; Traugott, 1993)

Como foi mostrado nos capítulos anteriores, esta hipótese fez surgirem várias questões de ordem empírica e teórica que têm sido desenvolvidas por nós e outros autores. Neste capítulo, trataremos da seguinte questão: se *cê* é um clítico, ou se encontra num processo de cliticização, ocorreria, em orações como (2), o fenômeno da *interpolação*, isto é, a não adjacência estrita do clítico e do verbo devido à presença de constituintes, no caso, a negação, entre os dois termos:

¹ Uma versão desse trabalho foi publicada na *Revista da Associação Brasileira de Linguística* 1 (2), 2002, por Vitral, L.

(2) Cê não pode fazer isso!

Ora, o fenômeno da interpolação, como no exemplo seguinte, do século XVIII, que existia, de forma ampla, no português arcaico, e, de forma restrita, no português moderno, deixou de ser produtivo a partir do século XIX:

(3) "...que sem estabelecimento o não quer fazer..."
(*Documentos Históricos do Espírito Santo*, citado por Pagotto (1993: 188)).

O uso da interpolação de *não*, que, no português europeu contemporâneo, segundo Cunha e Cintra (1984: 315), é uma opção limitada à língua escrita de alguns escritores, é, seguramente, um recurso não mais disponível no português brasileiro. *Como explicar, portanto, a ocorrência de interpolação na totalidade dos casos, já que orações como as seguintes, em que cê aparece adjacente ao verbo, não existem em português?*

- (4) a. Djalma disse que (*já) cê (já) tinha saído quando ele chegou.
b. Porque (*não) cê (não) exigiu uma outra lasanha?

Se compararmos as orações em (4) com as seguintes orações nas quais aparece o clítico *lhe*, destaca-se o contraste:

- (5) a. Djalma disse que (já) lhe (*já) tinha enviado o pacote.
b. Porque (não) lhe (*não) disse a verdade?

O contraste entre (4) e (5) mostra que *cê* se comporta de maneira oposta à de um clítico pronominal, no que concerne à interpolação.

A fim de tratar desta questão, faremos um estudo do fenômeno da interpolação no português arcaico e moderno focalizando a

trajetória da forma clítica de terceira pessoa, *se* e, em seguida, compararemos o comportamento deste clítico nesses períodos com o da forma *cê* no português brasileiro de nossos dias. A hipótese que iremos investigar pode ser colocada da seguinte maneira:

- (6) *A cliticização deve ser vista como um processo diacrônico, previsto através de estágios discretos que deverão ser percorridos por cada processo específico de cliticização.*

Nossa hipótese inverte os termos do problema colocado, isto é, a caracterização da interpolação nos contextos em que *cê* aparece não invalida atribuir-lhe o estatuto de clítico. Pelo contrário, a interpolação de *cê* passa a ser um fenômeno esperado tendo em vista que *cê* – um novo clítico – deverá trilhar estágios iniciais da cliticização que admitem a ocorrência de interpolação.

Este capítulo está organizado da seguinte maneira: na Seção 2, são apresentados os corpora utilizados no trabalho e são discutidas questões a respeito de sua natureza e da validade da comparação entre dados do português europeu e do português brasileiro. Na Seção 3, são discutidas as razões que nos levaram ao cotejamento da partícula *se* com a forma *cê*. Alguns aspectos do trabalho de Martins (1994), sobre a interpolação no português europeu, são apresentados na Seção 4. Na Seção 5, encontra-se a análise dos dados que realizamos. Nas seções 6 e 7, é elaborada a análise de *se* como um clítico que é, estruturalmente, uma projeção máxima, no sentido da Teoria X-barras, e apresentadas algumas diretrizes por meio das quais pode ser pensado o desenvolvimento posterior de uma abordagem formal da gramaticalização. Nestas seções, utilizamos noções oriundas do quadro da Gramática Gerativa, como será feito no Capítulo 9, mas nosso objetivo não é tomar parte do desenvolvimento das propriedades das operações do sistema computacional, no sentido de Chomsky (1995),

e sim vir a *modelar um conjunto de proposições que analise, de maneira explícita, o papel das propriedades da gramática interna no que é descrito como processos de gramaticalização*. Esta nossa escolha, que é programática, tem as seguintes conseqüências: (i) algumas das hipóteses propostas não estão suficientemente desenvolvidas, ou seja, no estágio atual do nosso trabalho, é inevitável correr alguns riscos em termos de objetividade; (ii) nosso objeto de estudo força-nos a fazer certas opções teóricas dentre o vasto conjunto de propostas atualmente em discussão no âmbito da Gramática Gerativa, ou mesmo, a modificar noções gerativistas a fim de adequá-las aos nossos propósitos.

Na Seção 8, aspectos da trajetória de *se* no que diz respeito à sua origem e interpretação são comentados, levando-nos a propor que a evolução desta partícula na língua pode também ser analisada como um processo de gramaticalização. Na Seção 9, comparam-se características das formas *se* e *cê* que corroboram a hipótese formulada acima e explicita-se a razão da possibilidade do fenômeno exemplificado em (2). Nossas conclusões são apresentadas na Seção 10.

2. OS CORPORA

Como dissemos, analisaremos o percurso da forma *se* no português arcaico e moderno com atenção ao comportamento deste clítico em relação à interpolação. Para o período arcaico que, no nosso caso, corresponde aos séculos XIII, XIV e XV, valemo-nos do *corpus* de documentos notariais editados por Ana Maria Martins (1994). Trata-se de 210 documentos de caráter jurídico, que estão guardados no Arquivo Nacional da Torre do Tombo em Portugal. Os documentos transcritos foram produzidos ao longo de três séculos, entre meados do século XIII e XVI.

As vantagens do uso desse tipo de documento não literário, segundo Martins, dizem respeito à possibilidade de controlar, possivelmente

com mais sucesso, variáveis tais como diferenciação dialetal e estilística; e datar com precisão as formas encontradas. É preciso considerar também “a falta de textos originais e, para a maior parte da Idade Média, a inexistência de prosa de autor” (Martins (*op. cit.*: 2)). Além disso, não corresponde à realidade a impressão de que textos jurídicos são inadequados para estudos em sintaxe diacrônica por serem pobres discursivamente e cheios de fórmulas sintáticas “cristalizadas”. Como nosso próprio trabalho mostrará, esse tipo de texto encontra-se, como qualquer outro, sujeito, evidentemente, à história, permitindo-nos a observação de mudanças na língua.

Em relação ao período clássico, do qual consideramos os séculos XVI, XVII e XVIII, utilizamos cartas e textos literários extraídos do Corpus Anotado do Português Histórico Tycho Brahe, coordenado por Charlotte Galves (cf. Galves; Brito (1999)). Os autores e textos desse *corpus* que foram utilizados são os seguintes:

Século XVI: SOUSA, Frei Luís de. *A vida de Frei Bartolomeu dos Mártires*.

LOBO, Francisco Rodrigues. *Côrte na aldeia e noites de inverno*.

Século XVII: CHAGAS, Antônio das. *Cartas espirituais*.

MELO, D. Francisco Manuel de. *Tácito*.

Século XVIII: GARÇÃO, Correia. *Obras completas*.

AIRES, Matias. *Reflexões sobre a vaidade dos homens ou discursos moraes*.

Devido à disparidade de tamanho entre o *corpus* de documentos notariais (de agora em diante, *corpus* N) e o *corpus* Tycho Brahe (de agora em diante, *corpus* TB) – este já contém, parcialmente, 100.000 palavras –, procedemos da seguinte maneira: extraímos do *corpus* N todas as ocorrências da forma *se*, que totalizaram 190 ocorrências; em seguida, do *corpus* TB, retiramos dos textos de dois autores de

cada século, como explicitamos acima, as primeiras 32 ocorrências da forma *se*, o que totalizou 192 ocorrências. Além de compatibilizar os corpora, o uso de amostra do segundo *corpus* pareceu-nos, como é justificado adiante, plenamente adequado aos nossos objetivos.

Nossas escolhas em relação aos corpora colocam-nos algumas questões. A primeira delas diz respeito à validade da comparação de textos de gêneros diferentes, como os documentos notariais do *corpus* N e os textos literários do *corpus* TB. Como mostraremos na Seção 5, tal comparação não nos parece constituir problema: observando-se os períodos citados, há um decréscimo progressivo do fenômeno da interpolação, que também se verifica, portanto, nos documentos notariais. O cotejamento adicional de documentos notariais do século XVI com textos literários daquela época mostrou, como veremos, que a tendência ao decréscimo da interpolação independe do gênero dos textos.

Outras questões dizem respeito à legitimidade do confronto entre dados do português europeu e do português brasileiro e à definição da moldura teórica dentro da qual essa comparação pode ocorrer. Como se sabe, tais questões são recorrentes nos trabalhos de Lingüística Histórica que levam em conta o português brasileiro. Supõe-se, muitas vezes, uma certa continuidade entre o português europeu e o brasileiro, mas esta continuidade não pode deixar de implicar a possibilidade de escolhas pelo português brasileiro que não foram previstas ou trilhadas pelo português europeu. No entanto, como veremos, a análise da trajetória do clítico *se* comprova que houve prosseguimento no português brasileiro de tendências renovadoras apontadas no português europeu.

3. SE VERSUS CÊ: JUSTIFICATIVAS

O clítico *se* mostra-se o item ideal na tarefa de descrever o percurso da cliticização portuguesa com o objetivo de compará-lo com a

distribuição da forma *cê*. Duas razões podem ser dadas. A primeira delas é que, diferentemente dos demais clíticos, *se* pode associar-se, assim como *cê*, à posição do sujeito, estando relacionado com o caso nominativo, no sentido da Teoria do Caso (cf. Chomsky (1995)) – este não é o caso, evidentemente, do *se* reflexivo. É verdade, por outro lado, que *se* e *cê* não parecem ocupar a mesma posição estrutural. Este fato não parece ser problema porque, como mostraremos, não é possível falar de homogeneização absoluta em relação aos clíticos. As características dos clíticos nas línguas românicas modernas, como o francês e o espanhol, têm sido vistas muito mais como particularidades dessas línguas do que como definidoras de uma categoria universal de clíticos. Inúmeros trabalhos recentes têm mostrado existir certa variação no comportamento de clíticos quando se comparam línguas ou mesmo períodos diferentes de uma única língua (consultem-se, por exemplo, Klavans (1982); Zwicky e Pullum (1983); Fontana (1993, 1997); Vincent (1997); e outros). É preciso dizer também que mesmo quando se considera um período de uma língua, encontram-se discrepâncias de comportamento entre elementos reconhecidamente clíticos. Por exemplo, *se*, diferentemente dos demais clíticos portugueses, não pode sofrer elisão com outro clítico, o que sugere que ele tem mais autonomia fonética que os demais. Mas, de toda maneira, a diferença estrutural entre *se* e *cê* deve ser explicada e veremos adiante como enquadrá-la nas nossas hipóteses. A segunda razão, que nos leva a comparar *se* e *cê*, diz respeito ao fato de o percurso histórico de *se* poder ser analisado como um processo de gramaticalização, que, em certo número de aspectos, é comparável com o mesmo tipo de processo no qual a forma *cê* está envolvida.

4. A INTERPOLAÇÃO: MARTINS (1994)

O fenômeno de não adjacência entre clítico e verbo era comum no português arcaico e moderno (cf. Ogando (1980); Martins (*op.*

cit.); Pagotto (1993)). No período arcaico, a interpolação ocorria de forma ampla, isto é, vários tipos de constituintes, como sintagmas nominais, pronomes, negação, advérbios e sintagmas preposicionais, podiam se interpor entre os clíticos e o verbo. Observem-se os exemplos seguintes de Martins (*op. cit.*: 1.9.1) (os clíticos encontram-se sublinhados):

- (7) a. que sse nō deue a estender esse maravedi (1296).
b. e os que o nō outurgare (1268).
c. que se adeãte segue (1355).
d. o quall pardieiro lhe logo enprazarõ (1494).
e. asi como a atá áqui derõ (1295).
f. isto que lhes eu mãdo (1275).
g. como os elles e a dita as Egreia Auyã (1383).
h. que as Ao dicto Monsteiro deuiã Alguas pessoas (1357).

Uma questão pertinente é a da frequência desse fenômeno no período arcaico: é relevante saber se a interpolação foi produtiva ou apenas marginal no português deste período. De acordo com a análise de Martins, o emprego de estruturas interpoladas é mais frequente que o de estruturas não interpoladas. Observem-se os dois quadros abaixo, elaborados por Martins (*op. cit.*: 193), que apresentam a proporção entre a interpolação realizada e a interpolação não realizada, considerados todos os clíticos. Neste último caso, levou-se em conta o clítico adjacente ao verbo em contextos de potencial interpolação. O Quadro II apresenta os dados relativos à interpolação da partícula de negação *não*; o Quadro I mostra-nos a interpolação com outros tipos de constituintes:

Quadro I
INTERPOLAÇÃO ≠ NÃO

	Séc. XIII	Séc. XIV	Séc. XV	Séc. XVI
Interpolação realizada	66,6%	69,1%	57,0%	51,7%
	26/39	78/113	70/123	62/120
Interpolação não realizada	33,3%	30,9%	43,0%	48,3%
	13/39	35/113	52/123	58/120

Quadro II
INTERPOLAÇÃO DE NÃO

	Séc. XIII	Séc. XIV	Séc. XV	Séc. XVI
Interpolação realizada	94,1%	96,8%	90,7%	90,0%
	16/17	30/31	13/14	18/20
Interpolação não realizada	5,9%	3,2%	9,3%	10,0%
	1/17	1/31	1/14	2/20

Esses valores mostram-nos que a interpolação era a construção preferida, aproximando-se, no caso da negação, a 100% das ocorrências. E a partir do século XV parece haver o início de um processo de mudança, a frequência da escolha da estrutura sem interpolação começa a aumentar.

No período moderno, de acordo com Martins (*op. cit.*: Seção 4.3), o fenômeno da interpolação começa a sofrer um outro tipo de mudança, que tem a ver com a natureza do constituinte que aparece entre o clítico e o verbo. “A partir do século XVII, a interpolação de *não* continua a registrar-se, enquanto a interpolação de outros constituintes desaparece” (Martins *op. cit.*: 564)). A título de ilustração, aí estão alguns exemplos do período utilizados pela autora:

- (8) a. Porque se não póde servir e amar a um, sem ser inimigo do outro (Vieira, II: 334).
- b. A história, assim como a poesia, eram para os antigos coisas sagradas e religiosas que se não tratavam senão debaixo da invocação dos deuses (Garrett: 78).
- c. e o povo, não ousando voltar-se ainda contra elle já o não defendia porém, sumia-se covardemente (Oliveira Martins: 16)
- d. e eu quero que desistam deste grande mal contra seu gosto, pois o não hão-de fazer por vontade (Vieira, II: 346).

Observe-se a seguir o Quadro III, de Martins (*op. cit.*: 566), que nos permite averiguar as percentagens de ocorrência da interpolação de *não* em autores dos séculos XV-XVI a XIX e indica-nos se, nos textos desses autores, existe ou não interpolação de outros constituintes.

Como o Quadro III mostra, há grande variabilidade, considerando os vários autores, no uso de estruturas com interpolação. Pode-se falar, no entanto, de perda progressiva de interpolação quando confrontamos os períodos arcaico e moderno com o contemporâneo. Neste último, a interpolação deixa de existir como uma opção produtiva da gramática, limitando-se, como Cunha e Cintra propõem, à prosa de alguns escritores portugueses.

Quadro III
INTERPOLAÇÃO ENTRE OS SÉCULOS XVI E XIX

	Interpolação de não	Há interpolação de outros constituintes?
A. de Albuquerque (1462?-1515)	64,9% (24/37)	Sim
Damião de Góis (1502-1574)	100% (14/14)	Sim
F. Mendes Pinto (1510-1583)	82,4% (14/17)	Sim
Luís de Camões (1524?-1579)	56,6% (30/53)	Sim
Diogo do Couto (1542-1616)	81,8% (18/22)	Sim
F. Manuel de Melo (1608-1666)	89,5% (34/38)	Sim
Antônio Vieira (1608-1697)	92,5% (37/40)	Não
L. Antônio Verney (1713-1792)	26,3% (5/19)	Não
A. Garrett (1799-1854)	92,6% (25/27)	Não
Oliveira Martins (1845-1894)	83,3% (5/6)	Não

5. A INTERPOLAÇÃO DO CLÍTICO SE

Nossa análise da interpolação de *se* está de acordo com os resultados de Martins para os clíticos tomados de forma conjunta. É óbvio que concorreu para esta conclusão o fato de parte do *corpus* ser o mesmo. De acordo com a análise realizada, encontraram-se, no *corpus* N, 190 ocorrências de *se*. Como, por exemplo, as construções a seguir:

- (9) a. ...que sse nō deue a estender esse maravedi (*corpus* N, 1296).
 b. ...e que dera daquella agua a qual herdamento uija da que sse mais podesse ajudar (*corpus* N, 1310).
 c. ...segundo se en ella cõtijnha (*corpus* N, 1434).
 d. ... de que sse o dicto Rodrigue anes procurador do dicto Conçelho deu por entregue (*corpus* N, 1355).

Do total de 190 ocorrências, 18,4 %, isto é, 35 acham-se em estruturas interpoladas. A seguir, observem-se os quadros IV, que nos fornece o percentual de interpolação com a negação e com outros constituintes, e V, que nos mostra o percentual de interpolação nas estruturas com a negação:

Quadro IV
 PERCENTUAL DE INTERPOLAÇÃO DE *SE* (*CORPUS* N)

	Com não	Com outros constituintes
Interpolação de <i>se</i>	3,2% / 6	15,2% / 29

Quadro V
 PERCENTUAL DE INTERPOLAÇÃO DE *SE* NAS ESTRUTURAS COM NEGAÇÃO (*CORPUS* N)

Com interpolação	Sem interpolação
75% / 6	25% / 2

As conclusões que podemos tirar dos valores apresentados, apesar da exigüidade dos dados, estão em consonância com as tendências apontadas por Martins, isto é, a interpolação era produtiva, ocorrendo tanto com a negação quanto com outros constituintes e com estes últimos aparecia em maior número.

Já no período moderno, também de acordo com o que encontrou Martins, há indícios de que a gramática do português começa a sofrer mudanças. Das 192 ocorrências de *se* selecionadas do *corpus* TB, apenas 5,2%, isto é, 10 encontram-se interpoladas, sendo que, em todas elas, trata-se da interpolação com a negação. Vejam-se os exemplos:

- (10) a. em o qual se não destruam soberbos (*corpus* TB, Lobo, XVI, p.16).
- b. Enquanto se não conhecem bem instruídas (*corpus* TB, Melo, XVII, p.10)
- c. ...as que se não vêem (*corpus* B, Garção, XVIII, p. 113).

O quadro seguinte mostra-nos o percentual de interpolação de *se* em estruturas negativas que totalizaram 19 ocorrências:

Quadro VI
PERCENTUAL DE INTERPOLAÇÃO DE *SE* EM
ESTRUTURAS NEGATIVAS (*CORPUS* TB)

Com interpolação	Sem interpolação
52,6% / 10	47,4% / 9

Nossa análise comprova, portanto, que houve o desaparecimento das estruturas de interpolação com outros constituintes no período moderno. Esta conclusão está de acordo com os resultados de Namiuti (2001), que analisou a incidência da interpolação portuguesa no período moderno, considerando todo o *corpus* TB. O outro resultado relevante é o aumento da frequência das estruturas negativas sem interpolação: é o que nos mostra a comparação dos quadros V e VI.

Se considerarmos, finalmente, que, no português europeu contemporâneo, a interpolação é marginal ou preciosismo de linguagem e que, no português brasileiro, trata-se de um recurso inexistente, podemos reafirmar que tal fenômeno deixou de existir na nossa língua.

É possível voltar agora à objeção que expusemos anteriormente a respeito da validade da comparação entre *corpus* de gêneros diferentes, isto é, no nosso caso, o uso de documentos notariais e de textos literários. Realizou-se uma análise adicional da interpolação que comparou a incidência deste fenômeno em documentos notariais do século XVI e em textos literários do mesmo século, extraídos dos dois corpora já citados. Nos documentos notariais do século XVI, encontramos interpolação de *se* em 13,9% dos casos. É interessante atentar, então, para o fato de que a incidência desse fenômeno diminuiu se levarmos em conta o percentual de 18,2% apurados nos três séculos anteriores. O uso da interpolação nos documentos notariais é maior, no entanto, do que foi apurado nos textos literários desse mesmo século, nos quais foram encontrados 6,5% de interpolação. Esta diferença pode ser explicada pelo reconhecido caráter conservador dos primeiros, que é também sujeito à história. Outro dado interessante em relação a esses documentos é que dos 13,9% de ocorrência de interpolação, 8,5% são de interpolação com negação; muito próximos, portanto, dos 6,5% encontrados nos textos literários do século XVI, que só apresentam interpolação com negação. A partir dos valores apresentados deduz-se também a diminuição da interpolação com outros constituintes que passou a ser de apenas 5,4%.

Concluimos que o fato de os documentos notariais também sofrerem mudanças na direção que nos é pertinente – tendendo a diminuir a incidência de interpolação elimina, pelo menos para esse fenômeno, a objeção acerca da legitimidade da comparação de textos de gêneros diferentes.

6. A ANÁLISE DA INTERPOLAÇÃO:

CLÍTICOS COMO PROJEÇÕES MÁXIMAS

Examinemos mais de perto agora as ocorrências de interpolação que destacamos em (9). É importante observar que, nesses contextos,

o clítico *se* é o segundo termo da oração, aparecendo após um subordinante ou uma conjunção. O fato de clíticos poderem aparecer, sintaticamente, na segunda posição, e enclíticos, foneticamente, em relação ao primeiro constituinte da oração, tem sido descrito, modernamente, através do rótulo 2P (= second position). Várias línguas exibem este fenômeno, dentre elas, o grego homérico, o espanhol medieval, e línguas modernas como o tcheco, o servo-croata e outras (cf. Taylor (1990, 1996); Fontana (1993, 1997); Rivero (1997); Radanovic-Kocic (1996); Progovac (1996)). Pelo que se vê nos exemplos (9), também o português arcaico é uma língua com clíticos 2P (cf. Martins (*op. cit.*)). De fato, em 100% dos casos de interpolação com outros constituintes que examinamos, o clítico aparece como segundo termo da oração, logo após um subordinante como *que*, ou conjunções como *segundo*, *poes* e *e*. Há, entretanto, exceções no português arcaico, detectadas por Martins, em que os clíticos aparecem, por exemplo, na terceira posição nas orações. Essas exceções, aliás, existem também nas línguas reconhecidamente 2P como o grego analisado por Taylor (1990). Na verdade, a terminologia 2P deve ser vista apenas como um rótulo cômodo, descritivo, abrigando fenômenos que buscam uma abordagem explicativa. Pensar esta questão em profundidade exige uma análise mais ampla de dados do português e está fora do âmbito deste capítulo.

Mas, mesmo assim, parece-nos válido levar em conta a distribuição do clítico *se* nos casos que descrevemos: ele se aloca sempre numa posição entre o complementizador e a oração propriamente dita. Nas análises gerativistas deste fenômeno, supõe-se que o clítico se desloca de uma projeção argumental e ocupa uma posição entre as categorias C e IP (ou TP). Haveria também a possibilidade de considerar que o português arcaico apresenta estruturas OV (isto é, Objeto-Verbo) residuais em orações subordinadas (cf. Parcerro (1999)). Esta segunda análise deve, contudo, ser descartada porque, como mostra o exemplo (9d), o clítico pode aparecer antes do sujeito. Somos obrigados a dizer,

assim, que o fenômeno 2P no português arcaico se deve à aplicação de uma regra de movimento que desloca o clítico. As questões que se colocam, portanto, são (i) em que posição o clítico se aloca e (ii) a que se deve a ocorrência deste fenômeno.

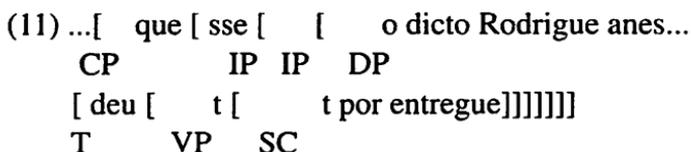
Há, pelo menos, dois tipos de análise para a questão (i) acima. A primeira delas supõe que existe algum tipo de categoria sintática entre C e IP e é esta categoria que abriga o clítico. Rivero (*op. cit.*), por exemplo, para descrever o espanhol medieval e o búlgaro, nomeia esta categoria de WP (= Wackernagel Phrase). Já Martins (*op. cit.*) propõe, baseando-se na análise de Laka (1990), que esta categoria é Σ P (=Sigma Phrase). Na versão desta autora, Σ P é uma categoria abstrata que funciona como expressão sintática das operações de afirmação e de negação proposicionais. É o núcleo de tal categoria que, através do mecanismo da incorporação recebe, de acordo com Martins, o clítico interpolado do português arcaico. A outra possibilidade de análise para o fenômeno 2P foi desenvolvida por Taylor (*op. cit.*), que considerou o grego homérico, e por Fontana (*op. cit.*) para o espanhol medieval. Para esses autores, o clítico 2P encontra-se adjunto à esquerda do domínio oracional IP.

As duas opções de análise descritas acima dão conta de representar a distribuição do clítico interpolado que ocorre após um item da categoria CP. Porém, distingui-las, do ponto de vista empírico, é muito difícil. A escolha por uma ou outra análise parece-nos calcada, na realidade, em razões internas à teoria, que, para serem explicitadas e discutidas, exigiriam um apanhado histórico das noções de movimento por substituição e por adjunção (ver Vitral (1992) para um comentário detalhado sobre esta questão), e também uma discussão acerca da noção de parcimônia e da estrutura de categorias funcionais que queremos admitir na gramática. Em relação a este último ponto, há, por exemplo, o problema de justificar o conteúdo dessa categoria entre C e I. Essa dificuldade é visível na proposta de Martins: sua categoria Sigma é capaz de checar verbos, clíticos, negação e um afixo abstrato

que indica afirmação. Por outro lado, a hipótese de que o clítico se incorpora ao núcleo de uma categoria sintática que é complemento de C parece mais compatível, teoricamente, com a Teoria da Checagem que se encontra atualmente em discussão no âmbito da Gramática Gerativa (cf. Chomsky (1995)). No entanto, esta opção de análise encontrará talvez mais dificuldades para lidar com a não-obrigatoriedade de movimento do clítico. De acordo com Martins (*op. cit.*: 190), por exemplo, “a interpolação não é obrigatória, isto é, há contextos de potencial interpolação em que esta não é actualizada”. Além disso, como vimos, a opção pela não interpolação não era marginal no português daquele período.

É evidente, porém, que a variabilidade do movimento do clítico é também um desafio para a análise baseada no movimento por adjunção ao domínio oracional IP. Essa questão da variabilidade pode encontrar tratamento promissor na hipótese da existência de gramáticas em competição, desenvolvida nos trabalhos de Kroch (1989, 1991); Kroch e Taylor (1997). Apesar das dificuldades apresentadas vamos adotar a opção de analisar o movimento do clítico como uma adjunção, o que se deve também à sua compatibilidade com a análise da gramaticalização que queremos desenvolver.

Consideraremos, assim, que o clítico 2P é deslocado, através da operação Mover, e adjuge ao domínio racional IP. Por meio do diagrama em (11) que representa o exemplo (9d), explicitamos esta análise:



O movimento do clítico não viola a hipótese de Preservação de Estrutura de Emonds (cf. Chomsky (1995: 318)) porque consideraremos, também de acordo com Fontana (*op. cit.*), que os clíticos 2P

não são núcleos, mas projeções máximas. Como tal, eles podem, portanto, adjungir a outra projeção máxima, no caso, à categoria IP. Na análise proposta, importa, sobretudo, constatar que o clítico em (11) se encontra deslocado da posição onde é gerado e o seu lugar de “pouso”. Embora determinar sua posição de origem seja irrelevante neste ponto da nossa análise, pois não nos ocuparemos de condições de localidade, consideramos que uma das opções de análise é dizer que o clítico é oriundo da posição sujeito de uma minioração (*small clause*). Ainda de acordo com Fontana, os clíticos sofrem um processo de mudança que os faz caminhar de projeções máximas para núcleos, ou seja, de XP para X°. Para ele os clíticos são, num primeiro momento, elementos associados à categoria C, isto é, são clíticos-C e se tornam, num segundo momento, clíticos associados a INFL, ou clíticos-I. Tal descrição dos fatos foi também empregada por Rivero (*op. cit.*) na sua análise da cliticização em várias línguas. Observem-se, como ilustração, os exemplos seguintes, extraídos de Rivero (*op. cit.*: 170), do servo-croata, (12a), cujos clíticos são associados a C; do búlgaro, (12b), que apresenta clíticos ligados a I; e do espanhol medieval, (12c e d), que dispõe de um sistema misto, admitindo clíticos associados a C e a I.

- (12) a. Da li nam Olga nesto dovikuje.
Q nos Olga alguma coisa disse?
b. Ako toj burgo go naprovi.
Se ele rápido isso faz.
c. Si lo el rey por bien toviere, mándeme quemar
Se isso o rei considerou por bem, mandar-me
queimar
d. E si él mejor lo faze - dixo el rey.
E se ele melhor o faz - disse o rei.

Pode-se deduzir, de acordo com essas análises, que, como já dissemos, as características dos clíticos pronominais das línguas

românicas modernas podem ser, na realidade, um resultado de um processo diacrônico de cliticização, e não condições universais e a-históricas de uma eventual categoria clítica abstrata. Que as propriedades dos clíticos das línguas românicas modernas não são exclusivas e universais já foi proposto por outros autores em análises de perspectivas diferentes (ver, por exemplo, Zwicky (1977); Klavans (1982)).

7. A ABORDAGEM FORMAL DA GRAMATICALIZAÇÃO: ALGUMAS ESPECULAÇÕES.

A análise através da regra de movimento coloca, evidentemente, a questão de saber por que razão o clítico se desloca. Esta pode encontrar desenvolvimento a partir de certas propostas no âmbito do Programa Minimalista (cf. Chomsky (1995)). Mas, como dissemos na Introdução, o nosso uso dessas noções visa, sobretudo, a demonstrar que propostas oriundas de um quadro formalista podem nos instrumentalizar na descrição e análise dos fenômenos lingüísticos captados através da noção de gramaticalização. É nesta perspectiva, assim, que fizemos certas escolhas teóricas distintas da ortodoxia gerativista, o que nos parece legítimo, já que nossos esforços têm se concentrado na tarefa, que se encontra ainda no seu início, de construir um objeto de estudo específico.

Para tratar, então, da questão da descrição do movimento do clítico, levaremos em conta a idéia de Chomsky de que a regra de movimento pode ser vista como um mecanismo de “atração” (cf. Chomsky (1995: 297)). Vamos rejeitar, no entanto, a proposta de que uma categoria X se move a fim de entrar em relação com uma categoria Y porque X e Y têm traços idênticos que deverão ser checados. Essa proposta, que não nos parece, independentemente de sua formulação, de caráter explicativo real, é tautológica, gera dificuldades, como as apontadas em relação à categoria ΣP , proposta

por Martins (1994), e também não consegue tratar do fenômeno de deslocamentos de clífticos.

A noção de atração de que necessitamos deverá ser, portanto, diferente da proposta por Chomsky. Trata-se da formulação seguinte:

- (13) α atrai β se α é uma categoria gramatical, α comanda β , e β é definido por traços categoriais gramaticais.

Em (13), utilizamos o termo *gramatical* ao invés do termo *funcional*, que é corrente no gerativismo. A razão de tal escolha é simples e sem qualquer efeito no conteúdo de nossas propostas: como nosso objetivo é desenvolver uma abordagem formal da gramaticalização, importamos evitar o termo *funcional*, definitório da perspectiva funcionalista que tem se apropriado, com muito proveito, da noção de gramaticalização na descrição de fenômenos lingüísticos.

A formulação (13) deverá incluir também alguma noção de localidade que consiga prever os resultados desejados, evitando que construções mal formadas sejam geradas, mas não trataremos deste ponto aqui. Além desta condição, precisamos determinar igualmente os possíveis lugares de “pouso” da categoria que sofreu atração. Para lidar com este último ponto, levaremos em conta as seguintes noções de domínio, elaboradas a partir da discussão de Chomsky (1995: 177-179) sobre esse tema:

- (14) o domínio de um núcleo α é o conjunto de nós categoriais contidos em $\text{Max.}(\alpha)$ e a categoria imediatamente dominada pela categoria “irmã” de α ; sendo que,
- (15) $\text{Max.}(\alpha)$ é a menor projeção máxima com o estatuto de categoria plena que domina α .

Vamos propor enfim a noção de *domínio de atração* nos seguintes termos:

(16) o domínio de atração de um núcleo α inclui qualquer posição contida no domínio de α .

De acordo com estas noções, uma categoria pode ser atraída para as posições de especificador de αP , de adjunção a α ou para uma posição dominada imediatamente pela “irmã” de α .

Em (11), de acordo com (16), o clítico, que é um DP, encontra-se, portanto, no domínio de atração de C, isto é, na posição de adjunção a IP que é irmã de C. As outras duas posições lhe são interditas porque a primeira delas, a posição de especificador de CP, já se encontra preenchida por um operador sem matriz fonética que sinaliza a oração subordinada declarativa, e, a segunda posição, a de adjunção a C, não está disponível em razão da Hipótese de Preservação de Estrutura, ou seja, só um núcleo poderia estar adjunto a essa categoria.

Como dissemos, diferentemente das propostas usuais do Programa Minimalista, não consideraremos que o movimento do clítico se deve à checagem de traços. Esta hipótese não conseguiria explicar a distribuição do clítico medieval (e nem a dos clíticos atuais): ora, seria necessário estipular, sem qualquer base empírica e apenas para preservar a teoria da checagem, que a categoria C e os clíticos contêm traços checáveis semelhantes. A hipótese que queremos desenvolver é que as categorias gramaticais têm a propriedade inerente de “atrair” outras categorias e que tal “força de atração” pode ser variável quando comparamos estágios de uma língua ou línguas diferentes. Nesta perspectiva de análise, é interessante notar que os clíticos podem sofrer uma modificação no domínio de atração e isso parece caracterizar o percurso diacrônico da cliticização, isto é, os clíticos podem passar do domínio de atração de C para o domínio de atração de I ou, em outras palavras, a atração de C perde “força” quando se comparam os

períodos arcaico e contemporâneo do português. A fim de ilustrar essa hipótese, consideremos os dois exemplos seguintes, o primeiro do português arcaico e o outro do português brasileiro contemporâneo:

(17) a. Isto que lhes eu mândo (*corpus* N, 1275).

b. O livro que eu lhe mandei.

Vê-se, assim, que, em (17a), o clítico *lhes*, uma projeção máxima, se encontra adjunto a IP e, de acordo com nossas propostas, no domínio de atração da categoria C, que, nesse estágio do português, dispunha de “força de atração” para atrair o clítico. É interessante notar também que, nessa oração, não foi o pronome *eu* que foi atraído por C, o que lança questões acerca do papel das condições de localidade nestes casos. Por outro lado, numa oração contemporânea, como (17b), a categoria C não é mais capaz de atrair o clítico *lhe* que, como núcleo, se encontra adjunto a I e, portanto, no domínio de atração desta categoria. No português moderno, teríamos um período de transição em que o clítico, quando interpolado, é ainda uma projeção máxima, mas se encontra sob o domínio de atração de I. Pode-se supor, além disso, que a atração de I, que, no português contemporâneo, concentra várias entidades, como o verbo, tempo, concordância, clíticos, modo, negação e talvez outras, pode se dever ao fato de I ser o núcleo ou o centro da oração.

Talvez o fenômeno da cliticização não seja o único que possa ser visto na perspectiva da “força de atração” das categorias gramaticais. O fenômeno V2, característica de estágios de muitas línguas, é outro caso que pode ser analisado nessa perspectiva. Neste caso, a idéia a ser desenvolvida é que C deixa de atrair o verbo que passa a ser atraído por I.

É verdade, por outro lado, que as sugestões de análise acima são também estipulações que deverão ser desenvolvidas posteriormente. Há, por exemplo, restrições acerca de que categoria pode atrair que

categoria e se essas restrições são mais bem analisadas por razões estruturais e/ou por razão dos conteúdos das categorias envolvidas. Não vou, porém, tratar desta questão aqui que envolve uma discussão mais ampla das possibilidades formais do conjunto de noções propostas. Deixo perceber, no entanto, a nossa tendência a considerar que as restrições são basicamente estruturais quando se considera a gramática interna e que aspectos de conteúdo, que venham a gerar construções mal-formadas, podem ser tratadas como condições de legitimação nas interfaces LF e PF.

Voltando ainda ao caso dos clíticos no português arcaico, resta um certo número de dificuldades que deverão ser explicitadas numa análise mais completa do fenômeno. A principal delas, no nosso modo de ver, diz respeito ao fato de o clítico interpolado ocorrer em orações subordinadas, majoritariamente; e em orações introduzidas por quantificadores, por advérbios “proclisadores” e por sintagmas focalizados, isto é, em ambientes de anteposição obrigatória de clíticos no português europeu atual (cf. Rouveret (1989); Martins (*op. cit.*: 183-190)).

Pode-se deduzir disso que, no português arcaico, clíticos nunca ocorrem na primeira posição na oração. Esse fenômeno ocorre em várias línguas e foi descrito, no século XIX, através da conhecida lei de Tobler-Mussafia que pode ser enunciada a seguir (cf. Tobler (1875); Mussafia (1898)):

- (18) Pronomes não acentuados não podem ocupar a posição inicial numa oração.

Em (18) temos, na verdade, uma generalização descritiva que, seguramente, deve ser deduzida de um enunciado teórico mais abrangente. As análises da direção da cliticização fonética e dos padrões rítmicos das línguas parecem ser perspectivas promissoras nesse sentido (no que concerne ao português, ver Galves (1999));

Abaurre e Galves (1998); Brandão de Carvalho (1989)). Pode-se observar, no entanto, que, no português brasileiro, a restrição (18), como é sabido, não parece ser atuante:

- (19) a. me passa o sal!
b. se vê muita gente pobre de espírito por aí.

Os dados em (19) mostram que o fato de a forma *cê* poder aparecer na primeira posição na oração não contraria, necessariamente, a nossa hipótese de que esta forma passa por um processo de cliticização.

Outro ponto a ser explorado é saber se a perda de substância fônica, que deverá incluir a perda do acento de palavra, e a perda de estrutura, isto é, a mudança de projeção máxima para núcleo, pode ser vista como uma consequência da atração exercida pelas categorias gramaticais, ou seja, a pergunta que queremos colocar em discussão é a seguinte:

- (20) Pode a atração de itens exercida pelas categorias gramaticais ser vista como causa dos efeitos de redução fonética e semântica descritos através da noção de gramaticalização?

Como vimos, são as categorias gramaticais que atraem itens pertencentes a outras categorias gramaticais (e também pertencentes às categorias lexicais – o que deverá nos levar a reformular (13)), que passam a se alocar, através de movimento ou de inserção, nas posições previstas pelo domínio de atração das categorias gramaticais “atratoras”. A questão em (20) sugere, assim, que seja investigado se é esta “cooptação” de itens, exercida pelas categorias gramaticais, que provoca a redução fonética e a alteração da natureza do significado (ou “esvaziamento semântico”) captadas por meio da noção de processo de gramaticalização.

A questão (20) tem, evidentemente, apenas um valor programático e necessitará de desenvolvimento e explicitação. Vê-se, no entanto, que, na nossa perspectiva, a noção de gramaticalização tem somente um valor heurístico e não é, *per si*, explicativa. As características dos processos de gramaticalização normalmente apontadas são, na realidade, a parte visível da aplicação de mecanismos da gramática interna que operacionalizam ou definem as condições de possibilidade destes processos.

As características da gramaticalização podem, portanto, ser pensadas a partir da determinação de propriedades da gramática interna, no sentido de Chomsky (1995), que se deixam balizar por princípios gramaticais. A unidirecionalidade dos processos de gramaticalização é, por exemplo, altamente esperada, levando-se em conta nossa hipótese sobre o papel da propriedade de atração das categorias gramaticais. Na literatura, tem-se apontado, entretanto, alguns contra-exemplos em relação à unidirecionalidade que, a nosso ver, são escassos e mal compreendidos. De toda maneira, talvez não seja impossível que os chamados contra-exemplos da unidirecionalidade possam ser descritos através de nossa abordagem do papel da atração das categorias gramaticais. É o que uma análise ulterior desta questão poderá mostrar.

Um problema importante é a que atribuir a propriedade de atração das categorias gramaticais. Podemos avançar a respeito desse ponto apenas mais uma especulação: de acordo com o Capítulo 6 (cf., também, Alves (1998)), as categorias gramaticais não possuem *referência virtual* (cf. Milner, 1982)). Vamos considerar que esta característica as impede de obter itens do léxico que sirvam de “corpo” fônico para a expressão de seus conteúdos. Em princípio, a única maneira de elas obterem itens que as expressem é então através da “cooptação” de itens que já funcionam como itens de categorias gramaticais e lexicais. A “força de atração”, inerente às categorias gramaticais, é, assim, o mecanismo da gramática que viabiliza a “cooptação”.

Esta visão do fenômeno coloca-nos, evidentemente, um certo número de questões. Precisaremos determinar, por exemplo, o rol das categorias gramaticais universais que consideraremos e a razão da distribuição entre elas; os itens que podem ser atraídos por quais categorias gramaticais; os tipos de dependência entre categorias gramaticais e lexicais que a gramática admite e outras. Apesar do caráter altamente especulativo das idéias desta seção, o conjunto de propostas é coerente e tem chances de receber um tratamento explícito posteriormente.

8. A TRAJETÓRIA DE SE E A INDETERMINAÇÃO DO SUJEITO

Como vimos, a análise da trajetória do clítico *se* mostrou-nos que, do ponto de vista da forma, este elemento faz sua entrada na língua como uma projeção máxima sob a “órbita” da categoria C e evolui no sentido de tornar-se um núcleo dependente da atração da categoria I. O fenômeno da interpolação e sua extinção podem ser analisados assim como a parte visível dessa mudança.

Vamos nos voltar agora para alguns aspectos da evolução de *se* que dizem respeito à interpretação dessa partícula. O clítico *se*, que, no português atual, participa de um bom número de construções, é originário do pronome reflexivo latino SE, acusativo (e ablativo) que, por sua vez, se vincula à raiz indo-européia *SE- que significa “à parte, separado, para si”, de acordo com Romanelli (1975: 169). *Se*, a partir da construção reflexiva, expandiu-se na língua de maneira a formar, inicialmente, a chamada construção passiva-*se*, com concordância e, posteriormente, a construção conhecida como *se*-impessoal. Observe-se o que diz Naro (1976: 788):

“There can be no doubt about the historical chronology: the *se*-passive, with agreement and agent phrase, precedes the *se*-impersonal, without agreement or agent phrase, by several centuries.”

A construção se-impessoal que, como se sabe, apresenta seu sujeito interpretado como indeterminado, se expande no português seguindo, provavelmente, conforme Nunes (1990: 95), a ordem seguinte:

- (21) a) verbos transitivos diretos usados intransitivamente (ex.: Come-se muito no inverno).
- b) verbos intransitivos (ex.: Trabalha-se bastante neste lugar).
- c) verbos transitivos preposicionados (ex.: Precisa-se de empregados).
- d) verbos de ligação (ex.: É-se feliz quando se é jovem).
- e) verbos ergativos (ex.: Chegou-se tarde à reunião).
- f) verbos em construções passivas perifrásticas (ex.: aqui se é visto por todos).

Este percurso de *se* pode, com segurança, ser identificado como um processo de gramaticalização. Temos, assim, sua origem lexical, que remonta ao indo-europeu, e sua progressiva ampliação, como indeterminador do sujeito, em construções variadas. Este tipo de ampliação de contextos de uso é um dos critérios para “diagnosticar” os processos de gramaticalização (cf. Heine; Hünemeyer; Claudi (*op. cit.*)); e foi identificado na nossa língua, por exemplo, no trabalho de Vianna (2000) sobre os verbos modais. Note-se, porém, que, nesse estágio pelo menos, esse fenômeno de ampliação não implica mudança no sentido laboviano, isto é, uma forma desaparecendo para dar lugar a outra (cf. Vitral (2001); Vianna (2000)). Uma pergunta que se pode fazer é: em que sentido o se-indeterminador é mais gramatical que o se-apassivador? É difícil dar um tratamento explícito à questão. Os autores que trabalham com a noção de gramaticalização insistem que estes processos introduzem uma modificação da natureza dos itens,

levando-os a adquirir um sentido abstrato em detrimento de um sentido mais concreto (sobre este tema, ver o Capítulo 6, e Alves (1998), que levou em conta a noção de *extensão* na abordagem dessa questão). Entretanto, as análises de Nascimento (1984) e Nunes (1990), expressas no quadro gerativista, parecem ser compatíveis com o que precisamos e podem subsidiar uma discussão mais explícita dessa questão. Segundo eles, no caso do se-apassivador, como em *alugam-se casas*, *se* absorve o papel temático do argumento externo e o caso acusativo; já no caso do se-indeterminador, como em *aluga-se casas*, *se* tem somente a função de indeterminar a categoria vazia na posição sujeito e é esta que recebe o papel temático e o caso nominativo reservados ao sujeito. Como se vê, o se-apassivador recebe o papel temático, isto é, é o agente, mas isso não ocorre com o se-indeterminador. Essa descrição gerativista das diferenças entre os dois itens pode ser vista assim como uma caracterização da proposta de que o se-indeterminador é mais gramatical do que o se-apassivador, sendo, portanto, compatível com nossa hipótese de que o percurso de *se* pode ser analisado como um processo de gramaticalização.

9. COMPARANDO SE E CÊ

Nesta seção, veremos como a análise desenvolvida para a trajetória de *se* pode nos auxiliar em relação ao problema, colocado na Introdução, acerca da não adjacência da forma *cê* e o verbo, ilustrada pelo exemplo (3). Aqui, não poderemos comentar as várias conseqüências da nossa hipótese central, proposta no Capítulo 2, de que a forma *cê* se encontra num processo de cliticização, mas os demais capítulos deste livro e os trabalhos citados na bibliografia poderão ser consultados.

Está claro que, na nossa perspectiva, a cliticização é um processo diacrônico, comportando, portanto, etapas que são identificáveis em estágios da evolução das línguas, e que podem, inclusive, ser atribuídas a itens diferentes num determinado estágio de uma língua.

Do ponto de vista estrutural, esta visão dos fatos deixa-se perceber através da hipótese de que os clíticos evoluem de projeções máximas para núcleos. Ora, tal hipótese permite-nos assim explicar a ocorrência da negação entre *cê* e o verbo, ou seja, *cê* deve ser analisado como uma projeção máxima que ocupa a posição de especificador de IP. A análise está representada abaixo:

(22) [*Cê* [não pode [t [t fazer isso]]]
 IP I' NegP VP

Assim como os clíticos medievais, a forma *cê*, que se encontra no estágio inicial de seu processo de cliticização, é uma projeção máxima e *pode*, é claro, ocupar uma posição admitida às projeções máximas. Esta análise explica, enfim, a razão de a negação poder se interpor entre *cê* e o verbo. Já o caso de *se* é diferente, uma vez que já possui uma trajetória de alguns séculos na língua portuguesa. Ele passou, como vimos pelo estágio de projeção máxima, período em que admitia a interpolação, e, atualmente, na condição de núcleo, resultado de seu percurso diacrônico, aparece adjunto ao núcleo I e, portanto, não mais admite a interpolação.

Do ponto de vista do conteúdo, é possível também traçar a seguinte comparação entre *se* e *cê*, que reforça a hipótese de que as duas formas trilham processos de gramaticalização: esta última pode também ser usada como um indeterminador do sujeito em orações como (23) que foram discutidas em Ramos (1997a) e em Alves (1998) (ver também Capítulo 6):

(23) Na primeira aula [de helicóptero], *cê* fica assim, meio bambinha, mas depois é ótimo. (ex. 51 de Alves (*op. cit.*: 68))

É interessante notar que os usos de *se* e *cê* como indeterminadores do sujeito são os empregos mais recentes das duas formas, o que é

previsto pelos pressupostos das análises que lidam com a noção de gramaticalização. Também prevista por esses trabalhos é a expansão dos contextos de uso das formas que se gramaticalizam, o que, como vimos, pode ser verificado em relação às ocorrências das duas partículas.

10. CONCLUSÕES

Nossa discussão mostrou que a cliticização deve ser vista como um processo gradual e heterogêneo que é parte de um processo mais amplo de gramaticalização. Vimos que quando se compara o comportamento de formas reduzidas num determinado estágio de uma língua, em línguas diferentes ou em estágios diferentes de uma língua, constata-se uma certa heterogeneidade das propriedades dessas formas, o que nos forçou a concluir que a cliticização é um processo variável. Parece-nos, portanto, que devemos abandonar a idéia de tratar a noção de clítico como um tipo de categorização de propriedades universais e discretas. É descritivamente mais adequado tentar destacar estágios do processo de cliticização e verificar, para cada forma considerada, em que ponto se encontraria neste processo. O passo adiante é explicar, ou fazer deduzir de propriedades da gramática, o fato de itens, em determinado momento da história da língua, se colocarem num percurso de mudanças que os leva a se reduzirem e a se tornarem clíticos. Na literatura, como se sabe, nada é dito sobre este ponto e nossa proposta do papel da atração das categorias gramaticais pode, com boa vontade, encontrar um desenvolvimento mais objetivo e explícito posteriormente.

CAPÍTULO 8 O PROCESSO *NÃO* > *NUM* NA FALA¹

Estudos variacionistas sobre negação sentencial no PB têm apontado três tipos de construções, conforme a posição e o número de partículas negativas na sentença: uma partícula pré-verbal (1); uma partícula pós-verbal (2); e duas partículas, sendo uma pré-verbal e outra pós-verbal (3), conforme exemplificado abaixo.

(1) e se eu **não** sou formada hoje (E28)

(2) **acredito não** (E30)

(3) que eu **não** concordava com aquilo **não** (E09)

A construção apresentada em (2) tem um caráter certamente marginal na língua, pois sua frequência é de apenas 5,6%² mesmo em dialetos nordestinos.³ Embora essa construção evidencie propriedades

¹ Versão revisada do texto publicado no livro COHEN, M. A.; RAMOS, J. M. (Org.). *Dialeto mineiro e outras falas*. Belo Horizonte, UFMG, 2002. Texto apresentado em uma mesa-redonda no I Congresso Nacional da ABRALIN, 1997, em Maceió.

² Os tipos restantes apresentaram a frequência de, respectivamente, 73,5% e 20,8%, conforme levantamento realizado por Roncarati (1996:103).

³ Embora presente em outros dialetos, essa construção é avaliada pelos informantes como uma especificidade de dialetos nordestinos. Um

sintaticamente relevantes, tais propriedades não serão discutidas no presente trabalho.⁴ Interessam-nos particularmente as construções (1) e (3), referidas daqui para frente, respectivamente, como negativa simples e negativa dupla.⁵

O item **não**, em (1) e (3), tem sido descrito como advérbio, com propriedades de quase-clítico. Para Ilari *et al.* (1991: 131-133), a construção mais habitual da negação em português poderia justificadamente ser caracterizada como uma construção quase-clítica, uma vez que (i) sua posição precede imediatamente o verbo (ou o clítico pronominal quando este está presente); e (ii) o item **não** pode ser reduplicado. Ilari *et al.* concluem que “a idéia de uma classe de ‘quase-clíticos’ pode ser encarada como algo mais do que uma simples conjectura”. O presente trabalho fornece novas evidências a favor do estatuto clítico da negação no português do Brasil.⁶ Documenta a realização do item **não** como **num** no dialeto falado em Belo Horizonte e mostra ser sua ocorrência sistemática. Desse modo, fornece evidências de que é adequada a descrição de **num**, apresentada no *Dicionário Novo Aurélio* (p. 982): *uma variante popular do advérbio não*.

O ESTATUTO DE CLÍTICO

Uma primeira evidência do estatuto clítico vem da própria realização fonológica do item **não**. Vejam-se os exemplos abaixo:

elemento importante na discussão das negativas no PB seria descobrir a razão dessa avaliação. Esse tópico, entretanto, não será desenvolvido aqui.

⁴ Ver Camargos, 2002 e Capítulo 9, neste volume.

⁵ Embora o termo *negativa dupla* tenha sido aplicado a casos de concordância negativa do tipo *Não vejo nada*, Miotto (1991), dentre outros, vou fazer uso dele aqui para ocorrências [neg V neg], já que não temos até agora uma análise satisfatória de [neg V neg] que seja capaz de comprovar que não se trata de um subtipo de concordância.

⁶ No âmbito da Gramática Gerativa, outros autores também sustentam tal posição: Gonçalves (1994); Martins (1997); Miotto (1992).

(4) é que às vezes ele **num** aceita (E09)

(5) ah... **num** teve festa **não** (E09)

Outras evidências podem ser buscadas, com base na distribuição dos constituintes dos enunciados presentes no *corpus*. A partir da suposição de que a forma “enfraquecida” e “reduzida” **num** seria um clítico, foram formuladas predições sobre o comportamento dos dados. Se as predições forem confirmadas, novas evidências serão identificadas.

Passemos, então, às predições: os fatores (a)-(d), abaixo, deverão favorecer a ocorrência de **num**.

- (6) a. posição não final na sentença;
- b. a presença do item pleno “nã”, na mesma sentença;
- c. a posição pré-verbal;
- d. a presença de quantificadores do tipo “ninguém”, “nada”, na mesma sentença;
- e. contigüidade com V.

As predições (a-e) têm como base as propriedades exibidas por clíticos em geral, a saber: (i) nada pode intervir entre clítico e V, exceto outros clíticos; (ii) não pode ser modificado; (iii) não aceita acento contrastivo; (iv) não pode ser coordenado; (v) a ordem clítico-V é fixa; e (vi) não ocorrem sem V (*apud* Kayne, 1975: 84-110). Assim, por não ter independência fonológica o item **num** deverá precisar de um “hospedeiro” ao qual se ligar, geralmente à direita, tal como os clíticos pronominais: “me faz” ao invés de “faz-me”. Este hospedeiro deverá ser preferencialmente um verbo, já que a posição típica da negação sentencial no PB é pré-verbal; o item **não** não será modificado por itens tais como “mesmo”, e não poderá ser coordenado nem colocado em foco. Além disso, por ser “fonologicamente mais fraco”,

poderá ser reforçado por formas plenas, no caso pela forma não reduzida **não** ou por quantificadores como **nada**, **ninguém** etc.

Além das predições (a-e), duas outras puderam ainda ser testadas. Vejamos: orações subordinadas deverão desfavorecer a ocorrência **num**, já que as formas inovadoras são mais freqüentes em orações principais e absolutas (Vennemann (1973); Givón (1974, 1976); Green (1976)). Se a forma **num** é, no eixo do tempo, mais recente que a forma **não**, será possível identificá-la como resultado de gramaticalização, confirmada pela perda de substância fonética. Esta última predição é testável tanto com base no tempo aparente quanto com base no tempo real.

Para realizar uma análise no tempo aparente, selecionamos uma amostra composta de nove entrevistas com falantes residentes em Belo Horizonte, distribuídos em três faixas. Feito o levantamento, encontramos um total de 769 ocorrências, sendo 425 de **não** e 344 de **num**, respectivamente, 55% e 45%. A alta freqüência do item reduzido **num** chama a atenção.

Vejamos, a seguir, os resultados da análise quantitativa da distribuição dos itens. Inicialmente observamos que tipo de material ocorre à esquerda da partícula negativa, se é o sujeito sentencial, se é uma pausa ou se é outro elemento qualquer, tal como um advérbio. Vejam-se as ocorrências abaixo:

(7(=1)) e se eu não sou formada hoje (E28)

(8) /não/ isso eu não acho certo não (E28)

(9) ele não/o irmão defe fez isso (E28)

Em (7) temos negação sentencial. Em (8) temos construções em que a partícula negativa ou inicia a sentença ou corresponde sozinha a um enunciado completo. Em (9) temos a negação de

constituintes menores ou fragmentos de constituintes. Vejamos os resultados na Tabela 1.

Tabela 1
DISTRIBUIÇÃO DOS ITENS,
CONFORME ELEMENTO PRECEDENTE

	Nº	%	Prob.
NP Sujeito	132/260	51	.50
Pausa	109/240	45	.47
Outros	103/269	38	.53
Total	344/769		

O elemento precedente não foi selecionado pelo programa de análise quantitativa,⁷ o que significa que o elemento à esquerda não afeta a realização da partícula negativa. Este resultado constitui indicação de que a forma **num** se liga ao elemento à direita, e não ao elemento à esquerda, pois mesmo após pausa ambos os itens ocorrem.

O segundo fator testado é o tipo de material que aparece à direita do item. Quatro elementos destacam-se: pausa (10), verbo auxiliar (11), verbo não-auxiliar (12) e outros elementos, tais como um sintagma preposicional (13).

(10) isso eu num acho certo **não**/ (E28)

(11) minha mãe **num** vai fazer nada **não** (E15)

⁷O programa utilizado foi o VARBRUL, versão de 1988, desenvolvida por Pintzuk. Trata-se de um programa computacional que realiza manipulações matemáticas sobre um conjunto de dados lingüísticos, calculando frequência e peso relativo. Para maiores detalhes, ver Guy (1998) e Sankoff (1998b).

(12) nem **não** conhecia nada de escola (E15)

(13) **não** na casa da minha vó (E15)

A distribuição dos itens, presentes no *corpus*, conforme o tipo de elemento à direita, aparece na Tabela 2, a seguir.

Tabela 2
FREQUÊNCIA E PROBABILIDADE DO ITEM NUM
SEGUNDO SUA POSIÇÃO NA SENTENÇA

	Nº	%	Prob.
Verbo auxiliar	102/168	61	.76
Verbo não-auxiliar	227/384	58	.75
Pausa	4/180	2	.03
Outros	11/30	37	.51

Este fator foi considerado quantitativamente significativo pelo programa VARBRUL. A probabilidade de **num** é de .76 com formas verbais perifrásticas e .75 com formas verbais não perifrásticas.⁸ Se o item ocorre antes de uma pausa, é amplamente desfavorecida: .03. Isto constitui uma evidência a favor da hipótese de cliticização à direita, uma vez que ausência de pausa constitui uma evidência da não autonomia fonológica do item **num**.

Estes resultados permitem, ainda, definir como variação as ocorrências de **não** e **num** na posição pré-verbal, pois somente aí ambos os itens ocorrem. Veja-se que antes de pausa a probabilidade é quase zero, apenas .03. Os enunciados presentes no *corpus* são:

⁸Esta forma de descrever esse fator me foi sugerida por Marta Scherre, em comunicação pessoal, a quem agradeço.

(14) mas também **num**... num me arrependo por ter sido por provocação (E05)

Veja-se que se trata de pausa com alongamento, o que indica ou hesitação ou tentativa de manter o turno. Estes não seriam enunciados típicos. Assim sendo, parece mais adequado excluir do levantamento tais ocorrências. Em consequência, teremos de modo categórico o item **não** no ambiente [Neg + pausa], sem alongamento.

(15) **não**/ nós não acreditamos em vida após a morte (E04)

Consideremos agora a oração que contém o item em análise, se é, ou não, uma oração subordinada. A frequência e probabilidade do item **num**, apenas quando a negação é sentencial, aparecem na Tabela 3.

Tabela 3
FREQUÊNCIA DO ITEM NUM
EM DOIS TIPOS DE ESTRUTURAS

	Nº	%	Prob.
Subordinada	63/139	45	.41
Principal e absoluta	277/488	57	.58
Total	340/627		

A maior frequência de **num** em orações principais e absolutas do que em subordinadas corrobora a predição de que a oração matriz favoreceria o item reduzido, o que constitui uma evidência a favor da hipótese de mudança.

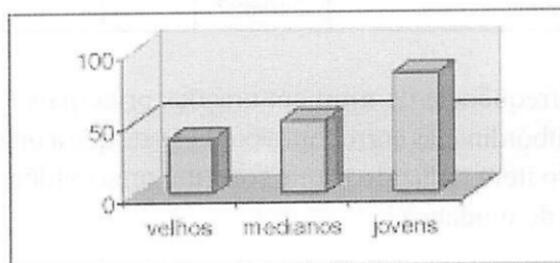
Se a caracterização de **num** como forma inovadora for adequada, a variação em análise deverá apresentar um perfil de mudança em progresso. Para verificar esta predição, consideremos o fator faixa etária. Os informantes de 15 a 29 anos formam o grupo dos jovens; os de 30 a 49 anos formam o grupo dos medianos; e os de mais de 50 anos, o grupo dos velhos.

Tabela 4
DISTRIBUIÇÃO DO ITEM NUM, SEGUNDO A FAIXA ETÁRIA DOS INFORMANTES

	Nº	%	Prob.
Jovens	139/165	84.2	.70
Medianos	120/237	50.6	.48
Velhos	85/225	37.7	.32

Este fator foi selecionado como quantitativamente significativo, confirmando que **num** é forma inovadora. O gráfico abaixo permite visualizar melhor estes resultados.

Gráfico 1
FREQUÊNCIA DO ITEM NUM, SEGUNDO O FATOR FAIXA ETÁRIA DOS INFORMANTES



O gráfico indica implementação do item **num**. Se se confirmar, com base no tempo real e em levantamento mais amplo, que se trata de mudança lingüística, seria importante comparar este processo àquele ocorrido no francês, em que, após o enfraquecimento da partícula negativa, testemunha-se seu desaparecimento. Até que ponto o PB estaria seguindo este caminho? Esta é uma questão que será discutida na próxima seção deste capítulo.

Consideremos por último o fator que avalia se a presença de quantificadores do tipo **nada** e **ninguém** favorece a ocorrência da forma **num**. A expectativa é que a forma plena seja desfavorecida, uma vez que o próprio quantificador já contém em si uma negação.

Dois tipos de ordem foram identificados, sendo o primeiro [Q Neg], somando um total de 6 ocorrências, e [Neg Q], com um total de 58 ocorrências, respectivamente exemplificados em (16) e (17).

- (16) a. barreira **nenhuma num** derruba a gente (E28)
b. **nunca** gostei **não**

- (17) a. tem uma família que **não** tem **ninguém** convertido (E09)
b. e eu **num** tirei **nenhum** tostão (E28)

Em (16a) tem-se NEG pré-verbal e Q em posição sujeito, tendo sido encontrada apenas uma ocorrência. O total de 1 em 64 ocorrências é uma indicação de que estamos aqui diante de um resíduo histórico. Há 5 ocorrências do tipo (16b), alcançando um índice mais alto (7,4%), e configurando uma dupla negação em que o **não** aparece no final do enunciado, antes de pausa, o que desfavorece a forma **num** (cf. Tabela 2).

Vejam os resultados apenas das seqüências [Neg Q], exemplificadas em (17).

Tabela 5
DISTRIBUIÇÃO DOS ITENS EM SENTENÇAS
COM ORDEM [NEG Q]

Realização	Nº	%
Não	21	36
Num	37	64
Total	58	

O item **num** é favorecido, apresentando uma frequência de 64% quando há um NP quantificado à direita.

Até que ponto a presença da “segunda” negação favorece o item **num** pré-verbal? Comparemos a frequência do item **num** pré-verbal em negativas simples e em negativas duplas. Se a “segunda” negação compensar a primeira, teremos um perfil crescente de **num** nas negativas simples e presença categórica de **num** nas negativas duplas. Vejamos a próxima tabela.

Tabela 6
DISTRIBUIÇÃO DO ITEM NUM EM NEGATIVAS
DUPLAS E SIMPLES, CONFORME
A FAIXA ETÁRIA DOS INFORMANTES

	Negativas simples		Negativas Duplas	
	Nº	%	Nº	%
Jovens	98/122	80,3	40/43	93
Medianos	93/187	49,7	26/51	50,9
Velhos	67/184	36,4	18/4	45
Total	258/493		84/134	

A diminuição da frequência de **num** em negativas simples parece evidenciar uma relação de causalidade entre a cliticização e a frequência de [Não V Não]. Entretanto, a presença não categórica de **num** nas negativas duplas enfraquece sobremaneira uma hipótese que assuma tal correlação.⁹

Subjacente à hipótese da correlação causal entre os dois itens estaria a suposição de que ambas fariam parte de um mesmo processo de mudança e cada uma poderia ser apontada como uma etapa de um ciclo, acompanhando a proposta de Jespersen (1917).¹⁰

Entretanto, se analisamos a negativa simples e a negativa dupla como etapas de um mesmo processo de mudança, então é de se esperar que a própria frequência de negativas simples esteja decaindo. É exatamente isso o que o estudo quantitativo mostra. Veja-se a tabela a seguir.

⁹ Salles Filho formula uma hipótese funcional nos seguintes termos: a segunda ocorrência de partícula negativa na mesma sentença decorreria do enfraquecimento da primeira partícula (Thomas (1969); Sales Filho (1980)).

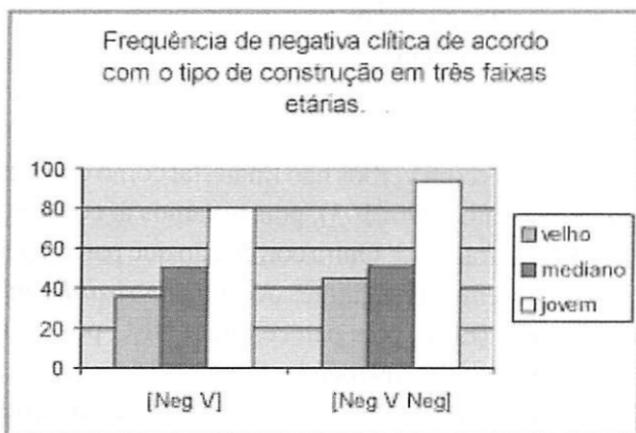
¹⁰ Schwegler (1991) argumenta a favor de que os três tipos de negativas existentes no PB evidenciarão três etapas de um processo diacrônico presente em outras línguas, como no francês, a saber: (i) NEG V > NEG V NEG > V NEG. Um problema para essa análise é haver uma diferença de natureza entre as negativas examinadas acima e as negativas de outras línguas românicas, como o francês: a partícula negativa nesta última segue imediatamente o verbo ao passo que no português a partícula negativa pode não vir imediatamente após o verbo. Vitral (1999) também assume a hipótese do ciclo, adotando um enfoque gerativista.

Tabela 7
**DISTRIBUIÇÃO DAS NEGATIVAS SIMPLES
 E NEGATIVAS DUPLAS, CONFORME
 A FAIXA ETÁRIA DOS INFORMANTES**

	Negativas simples		Negativas Duplas		Total	
	Nº	%	Nº	%		
Jovens	122	74	43	26	165	100
Medianos	187	78,5	51	21,5	238	100
Velhos	184	82,1	40	17,9	224	100
Total	493		134		627	

Como pode ver, as negativas duplas apresentam perfil ascendente, o que contribui para fortalecer a hipótese de que as negativas duplas estejam concorrendo com as negativas simples. Se considerarmos negativas simples e negativas duplas como duas construções, a atuação concomitante da mudança em ambas constituiria uma evidência de que o nódulo funcional NEG teria sofrido variação paramétrica, e a concomitância seria atribuída à Hipótese da Proporção Constante (Kroch, 1989). Veja-se a semelhança do perfil da frequência de **num** nas duas construções no Gráfico 1 anteriormente apresentado.

Gráfico 2



Os dois tipos de construção foram afetados pela cliticização. Se tivermos em conta a hipótese do tempo aparente, identificaremos, no gráfico, três períodos de tempo, respectivamente representados pelas faixas etárias. Na primeira faixa etária, a dos velhos, a cliticização está presente em mais de 35% dos casos. Na segunda há um avanço para 50%. Na faixa mais jovem a frequência avança ainda mais nas duas construções!

Os resultados apresentados aqui permitem mostrar que a alternância **não/num** constitui uma mudança lingüística em progresso. Estes resultados tiveram como base o tempo aparente. Um esquema do referido processo aparece abaixo.

$$(18) [\text{NÃO} > \text{NUM}]$$

$$[\text{NÃO} \vee \text{NÃO} > \text{NUM} \vee \text{NÃO}]$$

A correlação entre negativa simples e negativa dupla como etapas de mudança, já apontada por vários autores, foi também verificada no *corpus* analisado; a novidade aqui é a inserção de ‘num’ nesse processo.

Pode-se argumentar, entretanto, que as conclusões acima careçam de uma análise mais detalhada e a utilização de um *corpus* mais amplo, uma vez que a amostra analisada aqui inclui apenas nove entrevistas. Sem dúvida, esse tema requer maior investigação. Seria, a meu ver, importante, por exemplo, o estudo quantitativo e qualitativo de negativas do tipo V NEG, investigando principalmente frequência de V NEG em sentenças com verbos não finais, tal como em *Ele tomava banho não* (Roncarati, 1996: 103), pois teríamos aí construções que distinguem o PB do francês.¹¹ Outra correlação que parece relevante é aquela estabelecida entre as seguintes ocorrências *sozinho não* (E8) e *bateu no poste não* (E08), pois parece que o NEG pós-verbal nega em bloco algo que foi dito, independentemente de ser verbo, adjetivo ou advérbio.¹² Também seria oportuna uma comparação da frequência de *num* no *corpus* mais amplo do que o examinado aqui, para avaliar se em algum dialeto a cliticização do *não* estaria, de fato, mais avançada e qual seria sua correlação com outras mudanças no português brasileiro.¹³

¹¹ Um trabalho que desenvolve esse tópico de modo bastante interessante é *The Pragmatics of Negation in Brazilian Portuguese*, de Scott A. Schwenter.

¹² Ver Martins (1997) e Camargos (2002).

¹³ Ver Alkmim (2001).

CAPÍTULO 9
A NEGAÇÃO:
TEORIA DA CHECAGEM E MUDANÇA LINGÜÍSTICA¹

Nos últimos anos, têm surgido vários trabalhos que retomam a idéia da necessidade de se levar em conta a influência de aspectos diacrônicos na análise de fatos sincrônicos. Este capítulo, que analisa a distribuição de constituintes negativos em português e outras línguas, pretende ser uma contribuição nesse sentido.

O fenômeno da concordância negativa, existente em algumas línguas, mas inexistente em outras, será levado em consideração e examinado através da Teoria da Checagem (cf. Chomsky (1995)). Vamos nos valer também da noção de gramaticalização que, como se sabe, tem sido usada em estudos de sintaxe funcionalista (cf. Hopper e Traugott (1993)). Esta noção nos servirá, num primeiro momento, apenas como recurso descritivo para sistematizar certas observações de natureza diacrônica. Vamos, entretanto, dar um passo adiante e formular a hipótese de que o sistema computacional, no sentido de Chomsky (1995), reconhece as etapas de mudança lingüística previstas pelo processo de gramaticalização.

¹ Uma versão deste capítulo aparece como artigo na *Revista Delta* 15, 1:57-84, 1999, de autoria de Vital, L.

1. POLARIDADE NEGATIVA E CONCORDÂNCIA NEGATIVA

A discussão das construções negativas do inglês tem privilegiado o contraste entre, de um lado, a distribuição dos itens de polaridade negativa, isto é, *anybody*, *anything*, *anywhere* etc., e de outro, o comportamento dos quantificadores negativos, ou seja, palavras como *nobody*, *nothing* etc. O fenômeno da polaridade negativa é definido levando-se em conta os exemplos seguintes:

(1) a. I didn't see anybody.

b. * I saw anybody.

(2) a. Didn't anybody come?

b. * Anybody come.

Os contrastes acima são, geralmente, atribuídos à necessidade de os itens de polaridade negativa serem c-comandados pela partícula negativa *not* (ver, entre outros, Jackendoff (1972); Ladusaw (1979); Laka (1988); Progovac (1994)). Estes itens podem, entretanto, aparecer em contextos nos quais a partícula negativa está ausente, como, por exemplo, em frases condicionais, interrogativas, comparativas e **construções-too** (cf. Progovac (1994)):

(3) If anyone comes, let me know.

(4) Has anyone come?

Who wants any beans?

(5) John is taller than anyone in his class is.

(6) John is too smart to make any mistakes.

Os dados acima mostram, assim, que os itens de polaridade negativa não são inerentemente negativos, podendo ser licenciados por *not* como em (1a) e (2a) ou nos contextos como (3-6).

Os constituintes nomeados de quantificadores negativos, por outro lado, contrariamente ao que se passa em (1a) e (2a) não precisam ser precedidos da partícula negativa nem podem preceder esta partícula:

- (7) a. Nobody came.
- b. I saw nobody.
- c. * Nobody didn't come.

Em português, a oração que inclui a expressão *qualquer pessoa* parece, quando esta expressão é precedida por *não*, admitir uma paráfrase contendo *ninguém* na posição de objeto:

- (8) José não viu qualquer pessoa lá.

e também, como em inglês, esta expressão não ocorre na posição sujeito, antecedendo a partícula *não*:²

- (9) *Qualquer pessoa não irá à festa.

O fenômeno ilustrado por (8) não é, porém, produtivo em nossa língua. As orações abaixo mostram que, quando se fazem anteceder por *não*, as expressões *qualquer coisa* e *qualquer lugar* não se equivalem a *nada* e a *nenhum lugar*:

² Note-se, entretanto, que a presença de um verbo modal favorece a presença de *qualquer pessoa*:

- (i) a. * Qualquer pessoa foi à festa.
- b. Qualquer pessoa pode ir à festa.

- (10) a. José não comprou qualquer coisa em Miami.
b. Maria não vai viajar para qualquer lugar.

O termo quantificador negativo, por outro lado, tem sido usado em análises do italiano para designar palavras como *nessuno* e *niente*. Estas se comportam como *nobody* e *nothing* quando se alocam na posição sujeito, mas, na posição objeto, exigem a presença de uma partícula negativa (ver Rizzi (1982); Longobardi (1987) e Zanuttini (1989)).

Na nossa língua, os termos que podem ser definidos como quantificadores negativos se comportam da mesma maneira que seus similares italianos, isto é, exigem a presença da partícula *não* quando se alocam numa posição pós-verbal, mas parecem impedir sua ocorrência ao se encontrarem numa posição pré-verbal:

- (11) a. Ninguém (*não) conseguiu passar no vestibular.
b. Manuel não viu ninguém no cinema.
c. * Manuel viu ninguém no cinema.
- (12) a. Nada (*não) foi feito em favor dos pobres.
b. José não comprou nada em Miami.
c. * José comprou nada em Miami.
- (13) a. Nenhum aluno (*não) conseguiu passar no vestibular.
b. Maria não encontrou nenhum aluno na faculdade.
c. * Maria encontrou nenhum aluno na faculdade.

Os contrastes apresentados acima também são observados no comportamento dos advérbios negativos *nunca* e *jamais*:

- (14) a. O Luís nunca/jamais estava satisfeito.
b. * O Luís nunca/jamais não estava satisfeito.

- c. O Luís não estava nunca/jamais satisfeito.
 d. * O Luís estava nunca/jamais satisfeito.

Entretanto, como se verá na Seção 8, há dialetos do português brasileiro que aceitam a co-ocorrência de itens como *ninguém* ou *nunca* e a partícula *não* precedendo o verbo. Por razões de exposição, vou chamar os itens de valor negativo como *nada*, *nunca* ou *ninguém* de itens N.³ O fato, então, de a ocorrência da partícula *não* ser obrigatória quando os itens N se encontram numa posição pós-verbal, isto é, os contrastes em (11-14), é chamado por Zanuttini (1989) e Haegeman e Zanuttini (1991) – que analisam dados do italiano e de outras línguas – de concordância negativa (cf. também Milner (1982), a respeito do francês; e Labov (1972)).

2. A ANÁLISE DE POLLOCK (1989)

A distribuição das palavras de valor negativo do francês é um dos fenômenos analisados por Pollock (1989) num artigo que tem como objetivo estabelecer quais são e como se articulam as categorias da frase. A derivação proposta para orações como (15) se vale da representação, em estrutura-S, exposta em (16):

(15) Jean ne mange pas des andouettes.

(16) [Jean [ne mange [pas [[t [t des...]]]]]]
 TP T' NEGP NEG' AGRP VP

Como se vê, Pollock considera que os elementos *ne* e *pas* são constituintes de uma categoria negativa que funcionam,

³ Como vamos ver adiante, dispõe de valor negativo o item lexical caracterizado inerentemente pelo traço [+NEG]. Nessa perspectiva, os itens de polaridade negativa não têm valor negativo.

respectivamente, como núcleo e especificador. Para obter a ordem desses constituintes em (15), ele supõe que o núcleo *ne* – considerado como um clítico – se move para uma posição adjunta a T e que, a este núcleo, também se encontra adjunto o verbo *mange*. O especificador *pas*, enfim, se mantém na sua posição sendo, portanto, pronunciado depois do verbo.

A análise que Pollock propõe para as ocorrências das expressões negativas *ne...rien*, *ne...personne*, *ne...jamais* e outras difere, em certos aspectos da que foi proposta para *ne...pas*. Supõe-se que *ne...jamais/point/quère/plus* são geradas numa posição adverbial negativa que aparece numa posição inicial do VP; enquanto que *ne...rien* é um NP negativo que ocupa uma posição-A. Nesses casos, *ne* é tratado como núcleo e o segundo termo das expressões negativas como especificador. Por outro lado, com *personne*, a partícula *ne* não forma um constituinte: *personne* é visto como o núcleo de um NP e *ne* é o núcleo de uma categoria negativa sem especificador gerada numa posição à esquerda do VP.

Algumas dificuldades aparecem quando tentamos transpor para o português a análise que acaba de ser resumida. Há, primeiramente, a questão de saber se a partícula *não* pode ser considerada, à semelhança do *ne* francês, como um clítico. Esta questão foi discutida por Zanuttini (1989) a respeito da partícula *non* do italiano. Segundo esta autora, uma vez que existe a ênclise no italiano, o fato de pronomes serem clíticos não os obriga a preceder o verbo. Ora, é claro, que a simples hipótese de que *non* é um clítico não explica por que, em italiano, esta partícula deve preceder o verbo. O argumento, que não é válido para o português brasileiro falado se considerarmos que, quando ocorre neste dialeto, um clítico é sempre proclítico (cf. Kato (1994) parece, entretanto, ser pertinente para o português de Portugal ou mesmo para registros mais formais do dialeto brasileiro nos quais a ênclise aparece. De toda maneira, se levarmos em conta o conjunto de propostas conhecido por Programa Minimalista (cf.

Chomsky (1995)), a hipótese de Pollock é insuficiente na medida em que não nos oferece uma justificativa, baseada em princípios, para o movimento do núcleo da categoria NEGP: na busca constante de adequação explicativa, o movimento de um constituinte, na perspectiva minimalista, é motivado pela necessidade de checagem de traços gramaticais, o que garante uma dependência convergente entre dois constituintes.

No português brasileiro, há, por outro lado, assim como em francês, uma dicotomia entre *não* e *num*. Estas duas partículas, da mesma maneira que *non* e *ne* (cf. Moignet (1965)), se encontram em distribuição complementar nos contextos abaixo:

- (17) a. *não*/* *num* pertinente; *não*/* *num* fiável.
b. *non*/* *ne* pertinent; *non*/* *ne* fiable.
Você comprou o carro?
c. *Não*/* *num*.
Avez-vous acheté la voiture?
d. *Non*/* *ne*.
e. José *não*/*num* é inteligente.
f. Jean **non*/ n'est pas intelligent.
g. *Não*/**num*, o Oto *não* parou de fumar.
h. *Non*/**ne*, Oto n'a pas arrêté la cigarette.
i. Eu *não*/*num* vi a Maria *não*/**num*.

Como se verá na Seção 8, os contrastes acima reforçam a hipótese de que *ne* é um clítico. Poder-se-ia pensar, assim, que *num* é equivalente ao *ne* francês, embora, como o mostra (17e-f), a distribuição complementar entre a partícula portuguesa e *não*, contrariamente ao que se passa em francês, deixe de existir diante do verbo flexionado. Mesmo se *num* for analisado como um clítico, restaria explicar, como já foi mencionado, por que esta partícula deve preceder o verbo.

A eventual transposição da análise de Pollock para o português ainda encontra dificuldades em relação a exemplos como os seguintes:

- (18) a. José nunca/jamais ganhará na loto.
- b. Nunca/jamais o Pedro diria uma coisa dessas.
- c. Pedro nada comprará naquela loja.
- d. NADA, a Maria devolveu ao Pedro.
- e. Nada foi feito por Pedro.
- c. Ninguém sai de casa com um tempo desse.
- f. ? NINGUÉM, o José encontrou na festa.

A comparação dos dados acima com os de (11-14) nos permite fazer as seguintes observações:

- (A) a presença dos itens N em posições pré-verbais torna as orações bem-formadas. Compare, por exemplo, (14d) com (18a);
- (B) os itens N podem ocorrer sem se fazer acompanhar da partícula *não*.

A observação (A) nos leva a concluir que, em orações negativas do português, na ausência da partícula *não*, um item N deve, necessariamente, preceder o verbo. Ora, não há nada no sistema de hipóteses de Pollock que justifique esse fato. Não é possível lançar mão, evidentemente, da idéia de que também os itens N são clíticos.

A observação (B) acima serve como um argumento contra a hipótese da existência da categoria NEGP cujo núcleo seria *não* e cuja função de especificador seria desempenhada pelos itens N. A fim de mantê-la para o português, somos obrigados a considerar que, em orações como (18), NEGP conteria um núcleo *não* não realizado fonologicamente. Esta hipótese, que parece artificial na

forma em que acaba de ser enunciada, torna-se plausível caso consigamos justificar por que a presença de itens N em posições pré-verbais impede a ocorrência da partícula *não*.

O debate a respeito da existência da categoria NEGP pode tomar, na verdade, três possíveis vertentes: a primeira delas nega, simplesmente, a possibilidade de os elementos de valor negativo se constituírem como categoria sintática autônoma. Nesta perspectiva, não haveria NEGP, mas apenas advérbios ou nomes de valor negativo; a segunda, como acabamos de ver, leva em conta a existência de NEGP que dispõe de um núcleo, como o *ne* francês, e constituintes, os itens N, que funcionam como especificador, e a terceira vertente considera que uma partícula como *não* é o núcleo de uma categoria NEG que não dispõe de especificador, ou seja, os itens N não teriam esta função. Essas três possibilidades de análise serão discutidas a seguir.

3. O ESTATUTO NUCLEAR DA PARTÍCULA NEGATIVA

Ao analisar construções do italiano e do francês nas quais as partículas negativas se interpõem entre um clítico e seu vestígio, Kayne (1987) conclui que *non* e *ne* funcionam como núcleo no sentido da teoria X-barra. A má-formação de construções desse tipo se deve, então, a um efeito de Minimalidade (cf. Rizzi (1990)). No sistema de hipóteses de Chomsky (1995) prefere-se dizer que se tem uma violação da Condição do Elo Mínimo (“Minimal Link Condition”).

O fenômeno mencionado acima também ocorre em português:

- (19) a. José não os quer ver.
- b. José não quer vê-los.
- c. *José os quer não ver.

Supõe-se também que, em inglês, a impossibilidade de construções como (20) a seguir é também um efeito de Minimalidade (cf. Chomsky

(1991): *not* impede o movimento do verbo *likes* para INFL na sintaxe encoberta (ou para T, de acordo com Chomsky (1995)):

(20) *John not likes Mary.

Há também certos contrastes entre o comportamento dos advérbios e da partícula *não* que sugerem ser esta partícula não um advérbio usual, mas um núcleo no sentido da teoria X-barras.

Note-se inicialmente que um advérbio não impede o movimento de um clítico:

(21) a. Oto (lhe) tinha (lhe) educadamente pedido para sair do recinto.

b. Pedro não a tinha ainda observado.

A partícula *não*, assim como os auxiliares, permitem a ocorrência de um VP elíptico (ou apagamento de VP). Os advérbios, normalmente, não dispõem dessa propriedade:⁴

(22) a. José não tinha visto a Maria no parque mas o Oto tinha.

b. Udo conseguiu o emprego e o Max não.

c. * Max calmamente sorriu para a criança mas/e Maria calmamente/nervosamente.

Observe-se, enfim, que um advérbio aparece à esquerda ou à direita da categoria modificada enquanto que *não*, que em (24a), como se verá na seção seguinte, encontra-se em relação de

⁴ Assim como *não* e *num*, o advérbio *também* tem a propriedade de ocorrer com VP elíptico, o que talvez se deva ao fato de ele indicar, inerentemente, duplicidade. Por outro lado, o advérbio *Já*, que também admite apagamento de VP, pode ser analisado como um aspectual ou um marcador de frase afirmativa.

dependência com o predicado, só pode ocupar uma posição pré-verbal (ver, no entanto, (17i):⁵

- (23) a. (Provavelmente) Oto vai viajar (provavelmente).
b. Pedro modificou (completamente) seus horários (completamente).

- (24) a. Udo não encontrou (* não) a Maria.
b. * Udo encontrou não a Maria.

O fato de a negação dever aparecer numa posição pré-verbal pode ser analisado através de uma Condição de Adjacência que se faz necessária entre núcleos e seus complementos (ver Williams (1994); Stowell (1981); Ouhalla (1991)). Williams propõe que os núcleos das categorias lexicais atribuem seus papéis temáticos numa direção particular, à esquerda ou à direita. Também as categorias não lexicais, como a negação, que estabelecem uma relação de “functor” com seus complementos se submetem à direcionalidade. Através de um Parâmetro de Direcionalidade Temática, a direção é vista como uma dimensão de variação entre as línguas (ver também Travis (1984)). Seguindo a mesma linha de análise, Ouhalla (1991) considera que a dependência entre categorias funcionais e seus complementos permite tratamento através da noção de Seleção que seria, então, de natureza funcional. Não vamos discutir aqui a possibilidade de a dependência entre categorias funcionais e lexicais poder ser analisada através da noção

⁵ Em alguns dialetos do português brasileiro, é possível, como se sabe, a ocorrência de *não* pós-verbal, principalmente, como resposta a uma pergunta (mais detalhes sobre negação pós-verbal são apresentados na Seção 7):

- (i) – Você viu o José?
– Vi não (* o José)

de Economia (cf. Chomsky (1995)). Esta perspectiva não nos parece promissora devido ao elevado número de virtuais possibilidades combinatórias entre as categorias.

É possível considerar então que a partícula *não* toma uma categoria XP como complemento, sendo, portanto, um núcleo no sentido da teoria X-barra.

4. O ESCOPO DA NEGAÇÃO

A análise das construções negativas deve estabelecer o que exatamente é negado pelos itens de valor negativo. Propostas recentes têm considerado que há uma relação de dependência, que implica numa co-ocorrência, entre as categorias da negação e do tempo. Zanuttini (1989), e Haegeman e Zanuttini (1991) propõem que o núcleo da categoria NEGP seleciona a categoria funcional TP, que é então seu complemento (cf. também Laka (1988)).

Esta hipótese encontra dificuldades, no entanto, quando se observam construções como as seguintes:

- (25) a. A não demonstração do teorema por parte do professor levou o governo a prendê-lo.
- b. José considera o caseiro não condizente com o emprego.

Como se vê, a partícula negativa aparece num constituinte nominal, precedendo um verbal, em (25a), e anteposto em relação ao núcleo adjetival da minioração (*small clause*) em (25b). A negação pode ainda incidir sobre outros constituintes da oração. Em (26), é um quantificador que é negado e, em (27), um advérbio:

- (26) a. Os alunos fizeram não muitos trabalhos.
- b. * Os alunos fizeram não trabalhos.

(27) Ele falou não claramente.

Na tradição gramatical, associa-se normalmente a negação à presença do verbo (Jespersen (1917: 44), por exemplo, diz: “We...always find a strong tendency to attract the negative to the verb”). Vou considerar a hipótese de a negação ter como complemento, no sentido da teoria X-barras uma categoria lexical. O núcleo NEG seleciona, então, um VP em orações como (28) abaixo; um NP em sentenças como (25a); e um advérbio, considerado, portanto, de natureza lexical, em orações como (26a) e (27).⁶

(28) José não ama Maria.

Vamos propor que, a fim de que a negação possa ter sob seu escopo a categoria selecionada, o núcleo desta última deve ser c-comandado pelo núcleo *não* ou por um item N. Esta condição encontra-se destacada abaixo (cf. Vitral (1992)):

(29) Um item de valor negativo deve c-comandar um núcleo lexical na sintaxe visível.

A condição acima é compatível com línguas como o italiano e o espanhol e também com o inglês, como se vê por meio das orações abaixo:

(30) a. Mary will not help him.
b. John can't do this.

⁶O contraste entre (26b) e (25a) parece mostrar que a presença da negação é favorecida quando o núcleo nominal aparece com o seu complemento. Por outro lado, em ambientes contrastivos, é possível a negação do objeto: (i) Os alunos fizeram não trabalhos, mas cópias!

Nossa condição também parece pertinente para explicar o contraste entre orações como (11c), (12c), (13c) e (14d) e orações como (11a-b), (12a-b), (13a-b), (14c). Repito abaixo os exemplos (12b) e (12c):

- (31) a. José nada comprará em Miami.
b. * José comprará nada em Miami.

Em (31a), contrariamente a (31b), o item N *nada* c-comanda o verbo, obedecendo à condição (29).⁷

Há, entretanto, uma diferença de comportamento entre o inglês e o português no que concerne a frases como (31b). Em inglês, elas são bem-formadas:

- (32) a. Mary bought nothing in the shop.
b. John saw nobody at the party.

O problema colocado por (32) à condição (29) parece, no entanto, ser apenas aparente. Considera-se normalmente que, em inglês, o processo de checagem envolvendo o verbo e a categoria T tem lugar na sintaxe encoberta, após “spell-out” (ver Pollock (1989); Chomsky (1995)). Já que a condição (29) é pertinente na sintaxe visível, o verbo, ocupando sua posição de origem no VP, é c-comandado pelos itens N *nothing* e *nobody* em orações como (32). Por outro lado, em português, *nada* em (31b) não pode c-comandar o verbo se admitirmos que o verbo, diferentemente do inglês, alça a T antes de “spell-out” (cf. Vitral (1992)).

⁷O contraste abaixo pode servir de apoio à condição (29) (ver, no entanto, a Seção 8a respeito de (ib)):

- (i) a. O fato de ninguém ter saído não incomodou Maria.
b. * Ninguém não incomodou Maria

A análise que acabamos de propor encontra, entretanto, algumas dificuldades. A primeira delas diz respeito ao estatuto teórico de uma condição como (29) que determina a posição linear de um item a fim de que ele tenha uma categoria (ou o seu conteúdo) sob seu escopo. No modelo minimalista, tem-se procurado determinar a ordem dos constituintes através do processo de checagem. Assim, os itens se alocam no domínio de checagem de uma categoria de forma a checar seus traços formais em relação aos traços formais inerentes da categoria funcional que os abriga. É teoricamente desejável, então, que os fatos previstos por meio da noção de escopo possam ser reduzidos à teoria da checagem. O segundo problema da condição (29) é empírico e diz respeito à sua aparente falta de universalidade. Em algumas línguas, como por exemplo o francês falado e as línguas escandinavas – os exemplos abaixo são, respectivamente, do sueco e do islandês –, a negação aparece numa posição pós-verbal e, nestas línguas, considera-se que o alçamento do verbo se faz antes de “spell-out” (cf. Pollock (1989); Holmberg e Platzack (1988)):

(33) Jean achetera pas le bouquin.

(34) a. Köpte Jan inte boken.

b. Keypti Jón ekki bókina.

comprou J. não o livro.

“João não comprou o livro.”

Em (33) e (34b), admite-se que o verbo ocupa a posição T e, em (34a), é em C que ele se encontra. Nestas posições, os verbos não podem ser c-comandados pelos elementos negativos, o que viola a condição (29).

Na seção seguinte, os dados descritos pela condição (29) e também os problemas que esta condição levanta serão reconsiderados à luz da teoria da checagem.

5. A NEGAÇÃO E A TEORIA DA CHECAGEM

Não é nova a idéia de o escopo da negação ser determinado através de uma configuração especificador-núcleo, envolvendo uma concordância de traços formais negativos (cf. Rizzi (1990); Haegeman e Zanuttini (1991); Zanuttini (1994)). Essa hipótese pode ser aprimorada através da teoria da checagem de Chomsky (1995). Vamos considerar, então, que as construções negativas dispõem de uma categoria funcional negativa inerentemente definida pelo traço [+NEG]. Desde já, é possível perguntar a respeito da localização desta categoria negativa. Na literatura, há duas hipóteses principais: (i) o núcleo NEG toma TP como complemento; e (ii) o núcleo T toma NEGP como complemento. A distribuição das partículas negativas, se compararmos línguas como o inglês, o português e o árabe, favorece a segunda hipótese (cf. Benmamoun (1990) sobre o árabe). Além disso, adotá-la nos evita ter de encontrar uma razão para explicar o fato de o sujeito, de acordo com a primeira hipótese, ocupar a posição de especificador de NEGP.

A estrutura da oração negativa é, então, a seguinte (o uso do esquema X-barra é notacional e irrelevante na nossa discussão):

$$(35) \left[\begin{array}{c} [T \left[\begin{array}{c} [\quad \quad \quad \text{NEG} [\quad \text{DP V DP}]]]] \\ \text{TP} \quad \text{T}' \quad \text{NEGP} \quad \text{NEG}' \quad \quad \text{VP} \end{array} \right. \end{array} \right]$$

Vamos considerar assim que a presença na numeração da partícula *não* ou de um item N, que são caracterizados por um traço formal [+NEG], faz surgir uma categoria também definida pelo traço [+NEG], que tem a natureza Forte, isto é, exige a checagem antes de “spell-out”. A partícula *não* é inserida por meio da operação JUNTAR (MERGE) em NEG de forma a checar o traço [+NEG]. Os itens N tais como *nada* ou *ninguém*, que se alocam na posição objeto ou, funcionando como argumento

externo, na posição de especificador de VP, e *nunca* e *jamais*, que aparecem numa posição adverbial adjunta ao VP, se deslocam através da operação MOVER (MOVE) e podem ocupar a posição de especificador de NEGP ou transitar por esta posição. As duas operações permitem a checagem do traço [+NEG]. Quando o deslocamento dos itens N tem lugar na sintaxe visível, a inserção da partícula não é necessária e ter-se-á a geração de orações como (36):⁸

- (36) a. José nada encontrou na estante.
b. Ninguém conseguirá passar no vestibular.
c. Pedro nunca estava satisfeito.

Quando a operação JUNTAR insere a partícula *não*, o deslocamento dos itens N não precisa ser feito de forma que são derivadas frases como (37):

- (37) a. Maria não encontrou nada na estante.
b. José não ganhará nunca na sena.

A má-formação de orações como (38) abaixo se deve ao fato de o traço Forte [+NEG] não ter sido checado antes de “spell-out”:

- (38) a. * José encontrou ninguém na sala.
b. * Maria ganhará jamais na sena.

⁸ Nossa análise deixa entender que a ausência de *não* na numeração já predetermina o movimento de itens N antes de “spell-out”. A numeração pode não ser, então, apenas um “punhado” de itens do léxico. A formação das frases negativas inglesas, com inserção de *do*, nos leva à mesma conclusão. Permitir a inserção de *não* (em construções em que esta partícula co-ocorre com um item N) e *do*, que estariam assim ausentes da numeração, num ponto posterior da derivação acarreta, por outro lado, dificuldades em relação à Economia.

Vamos admitir então que o traço [+NEG] que define a categoria NEG é interpretável e permanece disponível para o sistema computacional e visível no nível da Forma Lógica mesmo após ter sido checado. Adaptando para as frases negativas do português a análise de Chomsky (1995) para as orações interrogativas do inglês, pode-se dizer que o traço [+NEG] dos itens N também é interpretável e dispensado de checagem. Como vimos, os itens N alçam, então, para o domínio de checagem de NEG, apenas se esta opção é selecionada para eliminar o traço Forte de NEG.⁹

⁹Na literatura, há, entretanto, evidências, a respeito de fenômenos de localidade, que nos levam a concluir que itens N, como *nessuno* ou *personne*, sofrem movimento na Forma Lógica (cf. Rizzi (1982); Longobardi (1991); Moritz e Valois (1994). Dever-se-á supor, assim, que o traço interpretável [+NEG] de itens como *nessuno* se encontra no domínio de checagem do núcleo NEG no nível da Forma Lógica.

⁹A agramaticalidade de frases como (i), contrariamente ao que se passa em inglês, em que um advérbio de modo aparece diante do verbo parece ser um contra-exemplo à hipótese do não movimento do verbo, no português brasileiro, na sintaxe visível (cf. Vitral (1992)):

(i) * Oto completamente modificou seus horários.

No entanto, um advérbio em *mente* que permite, além da interpretação de modo, também a leitura “orientado para o sujeito” (ver Jackendoff (1972)), é admitido entre o sujeito e o verbo:

(ii) Mané calmamente sorriu.

A frase (ii) é ambígua, permitindo as duas interpretações mencionadas.

O fato de a posição preferencial para o advérbio, no português brasileiro, ser entre o sujeito e o verbo pode ser também uma evidência da hipótese do não movimento do verbo antes de “spell-out”. Em relação ao problema colocado pela frase (i), é possível, enfim, fazer a previsão de que, provavelmente, um advérbio de modo “puro”, como *completamente*, é mais frequente em posição final, o que dependerá, evidentemente, de um levantamento estatístico.

A distribuição dos advérbios tem sido, por outro lado, colocada em dúvida como critério para se atestar o movimento do verbo (cf. Chomsky (1995); Costa (1996))

A derivação de frases como (36b) supõe, por outro lado, que *ninguém* transita pela posição de especificador de NEGP antes de se alocar na posição de especificador de TP na qual o traço D de T é checado. Os itens N podem, também, recebendo foco, aparecer numa posição no início da frase, após terem participado da checagem do traço [+NEG]:

(39) NUNCA, o Mané conseguirá ganhar na loto.

Finalmente, somos levados a concluir que, na nossa língua, o movimento do verbo para T não se faz na sintaxe visível.

6. A TEORIA DA CHECAGEM E A NEGAÇÃO PÓS-VERBAL

Como vimos na Seção 4, existem línguas, como o francês falado e as línguas escandinavas, nas quais os itens negativos aparecem numa posição pós-verbal (cf. (33) e (34)).

Em islandês, um advérbio negativo pode ainda aparecer entre o objeto direto e um sintagma preposicionado (cf. Holmberg e Platzack (1988)):

(40) Hann stingur smjörinu aldrei i vasann.
ele coloca manteiga jamais no seu bolso.
“Ele jamais coloca manteiga no seu bolso.”

Orações como (33), (34) e (40) têm desafiado as várias análises da negação que procuram estabelecer um eventual caráter universal do escopo dos itens negativos, levando em conta a relação entre a posição desses itens e a posição das categorias T (ou I) e V.

Considerando a análise proposta na seção anterior, perguntaremos, então, quais são as previsões feitas pela Teoria da Checagem em relação a orações como (33), (34) e (40). Como foi

dito anteriormente, supõe-se que, em orações como (34a), o verbo encontra-se na posição C, o que, aliás, é proposto, normalmente, para as línguas V2. Se, então, NEGP se aloca entre T e o VP, V, ao se mover para C, teria violado a Condição do Elo Mínimo, ao “pular” a posição núcleo de NEGP. Esta conclusão indesejável é evitada se dissermos que itens como *ekki* e *inte* não ocupam a posição núcleo de NEGP. Estes itens podem ser analisados como especificadores de NEGP ou como adverbiais gerados na posição de adjunção a VP. A primeira opção é, aliás, admitida por Pollock (*op. cit.*) para o *pas* francês. Vou adotar a seguinte análise desses fatos: itens como *pas*, *ekki* e *inte*, através da operação JUNTAR, encontram-se adjuntos ao VP antes de “Spell-Out” e se movimentam para o domínio de checagem de NEG na sintaxe encoberta de forma a checar o traço [+NEG], ou seja, nestas línguas, este traço é Fraco.

A boa-formação de orações como (40) parece mostrar que o item negativo *aldrei* não ocupa, antes de “Spell-Out”, uma posição no domínio de checagem de NEG. Também o fato de o francês falado admitir itens negativos na posição objeto mostra que, nesta língua, o processo de checagem exigido por NEG se faz no componente encoberto:

- (41) a. J'ai vu personne dans le cinema.
b. Paul a acheté rien hier.

A análise desenvolvida até aqui permite enfim fazer a seguinte classificação de línguas:

- (42) A. Línguas em que NEG é Forte: português, inglês, italiano.
B. Línguas em que NEG é Fraco: islandês, sueco, francês falado.

7. O CICLO DE JESPERSEN E A NOÇÃO DE GRAMATICALIZAÇÃO

A classificação estabelecida em (42) faz surgir, naturalmente, a questão sobre a razão da variação entre as línguas no que concerne à natureza da categoria negativa. O que determina o estatuto Forte/Fraco de NEG?

A fim de tentar responder a esta questão e também analisar certos fatos que se mantêm rebeldes ao tratamento proposto nas seções 5 e 6, parece-nos ser de grande valia considerar certos fenômenos de natureza diacrônica, que serão descritos através da noção de gramaticalização.

Sabe-se que a posição da negação em relação ao verbo pode também caracterizar estágios diferentes de uma mesma língua. Observando o inglês e o francês, Jespersen (1917, 1971) sistematiza este fato através do seguinte ciclo:

(43) Ciclo de Jespersen

	Inglês	Francês
1º	Ic ne secge	Jeo ne di
2º	I ne seye not	Je ne dis pas
3º	I say not	Je dis pas
4º	I do not say	
5º	I don't say	

O ciclo em (43) reflete a evolução da expressão da negação em outras línguas além do inglês e do francês: o dinamarquês, cuja partícula negativa *ikke* adquiriu sentido negativo historicamente, parece se situar no estágio 3º de (43). O mesmo pode ser dito para o sueco e o islandês analisados na seção anterior (cf; Jespersen (1971: 480)). Acrescente-se ainda ao paradigma inglês de (43) o fato de, em certos contextos, como em *I don't know*, haver uma redução fonética da partícula negativa, o que é analisado por Zwicky e Pullum (1983) como um fenômeno de cliticização de *not*.

O ciclo de Jespersen deixa entender que há uma relação entre a redução fonética da partícula negativa pré-verbal e o aparecimento de itens de valor negativo em posições pós-verbais, isto é, a redução da partícula pré-verbal faz surgirem itens negativos em posições pós-verbais que, por sua vez, adquiriram sentido negativo historicamente. Após o desaparecimento da partícula pré-verbal, os itens negativos pós-verbais passam a ser os constituintes que expressam negação e tendem, eles também, a sofrer redução fonética e se comportarem como clítico. A aparente relação de *causalidade* entre os dois fenômenos é, entretanto, um tema complexo que requer mais discussão.

A descrição do ciclo de Jespersen que acabamos de fazer parece, porém, ser corroborada por dois tipos de fenômenos portugueses.

O primeiro deles diz respeito à possibilidade de se atribuir o estatuto de clítico à partícula negativa *num* que, assim como foi visto em (17), se encontra em distribuição complementar com a forma *não*. Em (17), o fato de *num* não poder aparecer junto ao adjetivo; não poder funcionar como resposta a uma pergunta; nem poder aparecer topicalizado ou numa posição pós-verbal confirma o estatuto clítico dessa partícula: também um clítico pronominal não ocorre nestes ambientes.

As conclusões a que chega Ramos (1997) também reforçam a hipótese do estatuto clítico de *num* (cf. também Capítulo 8): (i) *num* não aparece antes de pausa, o que indica uma contigüidade estrita entre esta forma e o verbo; (ii) *num* liga-se ao constituinte à direita e não ao constituinte à esquerda, o que é compatível com a ausência de ênclise no português brasileiro falado; (iii) *num* não recebe acento contrastivo; (iv) *num* é favorecido quando há um NP quantificado numa posição pós-verbal, o que permite identificar uma correlação entre a presença de reforço e a redução fonética na posição pré-verbal; (v) levando-se em conta a faixa etária dos

falantes, confirma-se que a variante *num* é inovadora;¹⁰ (vi) as orações subordinadas desfavorecem a variante *num*, o que é coerente com seu caráter inovador.

A outra evidência da descrição proposta para o ciclo de Jespersen se refere ao fato, amplamente reconhecido, de certas palavras de valor negativo do português atual, e de outras línguas, terem adquirido, através de sua evolução histórica, a possibilidade de exprimir negação. É o caso, por exemplo, da palavra *nada* que se originou de uma redução da locução *rem natam* “coisa que exista” e, no português arcaico, tinha, como equivalente, somente a forma *rem* (cf. Mattoso Câmara (1979); Nunes (1945: II 10); Ilari (1984)). *Rem* deu origem também ao francês *rien*. Ainda em relação ao francês, podem-se citar os exemplos dos itens negativos *personne* e *pas* que, em estágios anteriores, significavam apenas “pessoa” e “passo” (cf. Traugott (1980); Hock (1991); Schwegler (1988)). O português atual admite ainda a possibilidade de expressões como *um tostão* ou *um pio* em, por exemplo, *Eu não tenho um tostão* ou *Ele não deu um pio*, participarem da negação da proposição.

O processo de cliticização envolvendo as partículas *não* e *num* parece não ser o único em curso no português brasileiro atual. Como vimos no Capítulo 1, a distribuição complementar entre as formas *você* e *cê* mostra ser esta última o resultado de um processo de cliticização que teve como ponto de partida a expressão *Vossa Mercê*. Admitiu-se então que a forma *cê* é o penúltimo estágio de

¹⁰Sabe-se, porém, que a forma *num* encontra-se atestada em textos medievais (cf. Huber (1933)). Será preciso verificar se o *num* medieval se encontra em distribuição complementar com a forma *não* e também comparar a frequência dessas duas formas nesses textos. Talvez a partícula *num* medieval não apresente comportamento de clítico e tenha-se um caso de variação estável que, apenas recentemente, no português do Brasil, de acordo com a análise com base em dados do tempo aparente, parece começar a se decidir a favor da forma *num*.

um processo de gramaticalização que pode ser descrito por meio das seguintes etapas (cf. Hopper e Traugott (1993)):

- (44) a. item lexical > b. item gramatical > c. clítico > d. afixo

As observações feitas a propósito do ciclo de Jespersen parecem indicar que é possível analisar este ciclo de forma a reduzi-lo ao processo de gramaticalização apontado em (44). O desaparecimento das partículas pré-verbais pode ser visto como o estágio Ø, posterior ao estágio (d) de (44), também previsto por Hopper e Traugott (*op. cit.*). Vem ao apoio dessa proposta o fato de itens de natureza lexical adquirirem valor negativo historicamente e se tornarem, assim, itens gramaticais ou, nos termos da Gramática Gerativa, itens funcionais: é o caso, por exemplo, da palavra *nada*. Por outro lado, a mudança em curso envolvendo o par *não/num*, descrita por Ramos (1997), parece incluir apenas as etapas (b) e (c) de (44). Este fato pode ser analisado de duas maneiras: (i) há mais de um ponto de entrada no processo de gramaticalização, isto é, no caso de *não/num*, ele se inicia no estágio (b); ou (ii) o estágio (a) é tão remoto que se torna difícil recuperá-lo: de acordo com Jespersen (1971: 479) a forma *ne* era, em sua origem indo-européia, uma interjeição que exprimia repugnância.

Comparando agora as observações feitas a respeito do ciclo de Jespersen, o processo de gramaticalização e a análise desenvolvida na Seção 7, diremos que a aplicação da operação MERGE e MOVE que, ao distinguirem traços funcionais e lexicais e a natureza Forte/Fraca dos traços funcionais, são responsáveis pelo processo de checagem, leva em conta o estágio de gramaticalização, no que concerne à expressão da negação, em que a língua se encontra.

Vamos detalhar o que acaba de ser dito observando então que, à primeira vista, os itens do estágio (b) de (44) são aqueles que podem participar do processo de checagem por meio da operação MOVE que os aloca na posição de especificador de NEGP. Já os itens do

estágio (c) de (44) são aqueles que participam do processo de checagem através da operação MERGE e ocupam a posição núcleo de NEGP. No entanto, a partícula *não* não é um clítico, mas, de acordo com nossa análise, é inserida na posição núcleo. Para entender a razão dessa dificuldade é preciso analisar mais de perto as etapas do processo de gramaticalização descrito em (44).

Os dois primeiros estágios de (44) são estabelecidos levando-se em conta um critério de tipo semântico: é a “perda de significado” que permite um item lexical se transformar em item gramatical.¹¹ Por outro lado, a mudança envolvendo os estágios (c) e (d) é determinada através de critério morfofonético. Vamos separar estas duas dimensões e propor uma versão estendida do processo de gramaticalização que terá o seguinte esquema:

(45) A. a. Lexical > b. Gramatical

B. a1. p.máxima > b1.p.máxima > b2.núcleo >
b3.clítico > b4. afixo.

Vamos admitir então que o processo de gramaticalização funciona em paralelo levando em conta os eixos A e B de (45) e que, além disso, o componente computacional distingue os estágios previstos no eixo (B). Por razões de espaço, não vamos nos deter aqui a respeito da questão sobre como o componente computacional distingue, por exemplo, os núcleos clíticos dos não clíticos. Em Vitral (em preparação), desenvolve-se a hipótese de o traço [+FOCO] ser o traço formal responsável por essa distinção.¹²

¹¹ Vitral e Ramos (1999) apresentam uma discussão a respeito da natureza da “perda de significado” apontada pelos autores que lidam com a noção de gramaticalização.

¹² Para uma discussão a respeito do confronto entre as perspectivas sintáticas e fonológicas no que concerne à cliticização, é possível consultar, entre outros, Fontana (1993); Rouveret (1989); Brandão de Carvalho (1989).

A forma *não* pode, então, ser classificada como do tipo (b2), assim como os itens *non* e *not*; exemplos de (b3) são as partículas *num*, *ne* e *n't*; *nunca*, *pas*, *ekki*, *inte*, *nada* são ocorrências de (b1).¹³ Vamos observar, por outro lado, que, nas línguas que dispõem de itens do tipo (b2) e (b3), a categoria NEG é Forte, já para aquelas línguas que só dispõem de itens do tipo de (b1), a categoria NEG é caracterizada como Fraca. A descrição acima é visualizada através do seguinte quadro:

(46)

A item lexical	item gramatical			
B proj. máxima	proj. máxima	núcleo		
		núcleo	clítico	afixo
	<i>pas</i> , <i>ekki</i> , nunc	<i>não</i>	<i>num</i>	
		<i>non</i>	<i>ne</i>	
		<i>not</i>	<i>n't</i>	

A análise proposta parece atestar a existência de uma correspondência entre a Teoria da Checagem e o processo de gramaticalização, que pode ser assim resumida: a Teoria da Checagem determina a ordem dos itens na oração valendo-se da natureza desses itens e esta sofre mutações previstas pelo processo de gramaticalização. É preciso dizer, no entanto, que o componente computacional reconhece não o processo de gramaticalização, e sim as etapas por ele previstas, o que o faz gerar produtos diferentes.

Lançar mão do processo de gramaticalização pode ter, enfim, a vantagem de retirar o caráter estipulativo da distinção Forte/Fraco

¹³ Note-se, porém, que do ponto de vista do eixo A, as formas (b1) não são homogêneas: *nada* e *nunca* adquirirão valores semânticos não admitidos, por exemplo, por *pas*, que é apenas um marcador negativo. Trata-se de mais uma razão obrigando-nos a precisar a natureza da “perda de significado” mencionada na literatura sobre gramaticalização (cf. Vitral e Ramos (1999)).

utilizada pela Teoria da Checagem no que concerne ao fenômeno da negação, mas introduz a questão de explicar a natureza desse processo tal qual está descrito em (45).

8. ALGUNS PROBLEMAS E A INTERPRETAÇÃO DA NEGAÇÃO

Nesta seção, vamos fazer algumas especulações de análise a respeito de certos fenômenos que se mostram rebeldes em relação à análise proposta.

No dialeto do português brasileiro examinado, a co-ocorrência de, por exemplo, o item N *ninguém* na posição sujeito e da partícula *não* é agramatical, mas há dialetos nos quais as frases abaixo são bem formadas:

- (47) a. Ninguém num disse isso.
- b. Ele nem num fez isso.

Construções como (47) já existiram no português do século XVI como nos mostram os seguintes exemplos de Said Ali (1966) (ver também Matos e Silva (1990)):

- (48) a. Nem eu nom vos faço prazer.
- b. Nenhum nom lhe soube dizer.

Também em italiano e espanhol, estruturas como (48) eram produtivas (cf. Meillete e Vendryes (1948)), e subsistem, como se sabe, no francês escrito:

- (49) a. Personne ne crois aux politiciens.
- b. Jamais, il ne dirais ça!

Para analisar esse tipo de construção, vamos, inicialmente, fazer algumas considerações a respeito da interpretação das construções negativas.

A interpretação de orações como (50) abaixo é descrita através da expressão em (51) (cf. Rizzi (1982: 124)):

(50) Ninguém quebrou a Janela

(51) Não há x, x = pessoa, tal que x quebrou a Janela.

Deve-se observar que embora a negação em (50) incida sobre a posição sujeito, como o mostra a fórmula em (51), ela tem o efeito de uma negação sentencial. Não foi praticada assim a ação de “quebrar a janela”. É provável que o correspondente formal dessa interpretação seja, como nossa análise mostrou, a possibilidade de o traço [+NEG] ser checado por *ninguém*, que também dispõe desse traço.

Pode-se agora especular que no dialeto que aceita *ninguém num* pré-verbais um dos dois itens não contém o traço negativo inerente capaz de checar [+NEG]. Já que, como vimos, os itens ganham valor negativo historicamente, é possível que, no português do século XVI, *nenhum* não tivesse, ainda, a capacidade de checar o traço [+NEG] (cf. Martins (1997)), enquanto que, nos dialetos atuais que admitem (47), talvez a partícula *num* não seja capaz de checar este traço, se comportando, assim, como um expletivo.¹⁴

Esta análise parece encontrar apoio nos fatos do francês. Como se sabe, o fato de partículas negativas serem expletivas ou redundantes é admitido para a partícula *ne*. A seu respeito, Robert Martin (1972: 21) nos diz que se trata de um morfema que

“...ne modifie pas le signe de l’énoncé et qui peut être supprimé sans dommage autre pour le sens que celui d’une subtile nuance.”

¹⁴ Vitral e Ramos (1999) discutem dados de Duarte (1997) que mostram, de forma independente, o uso da forma *você* como expletivo.

Também segundo Muller (1984), nos contextos em que o *ne* expletivo aparece, são os verbos (ou as preposições) que retêm o valor negativo do enunciado (ver também Stauf (1928)). Pode-se supor, então, que nos contextos como (49), tenha-se, na verdade, mais uma ocorrência de *ne* expletivo, isto é, incapaz de checar o traço [+NEG].

As hipóteses levantadas nesta seção podem ainda levar em conta, o que não será desenvolvido aqui por razões de espaço, o comportamento da categoria NEGP em comparação com a categoria CP. O dialeto português examinado nesta seção, contrariamente àquele descrito na Seção 1, admite, nos termos da teoria X-barra, o preenchimento simultâneo das posições de núcleo e de especificador de NEGP. É interessante observar que este preenchimento duplo, no que concerne à categoria CP, é a regra em várias línguas e levou Chomsky (1995) a propor a equivalência das operações JUNTAR e MOVER no que concerne à Teoria da Checagem. O dialeto português apresentado na Seção 1 mostra, no entanto, que se impõe a escolha de uma das duas operações para a checagem de [+NEG].

A comparação com a categoria CP deve ainda considerar os seguintes dados:

- (52) a. Nunca, ninguém derrubará o governo.
b. Nada jamais foi feito a seu favor.

As orações acima mostram que não há impedimento de preenchimento múltiplo da posição de especificador de NEGP.¹⁵ Ora, também

¹⁵ O preenchimento múltiplo da posição de especificador de NEGP, assim como no caso da posição de especificador de CP, parece porém, sofrer restrições. Observe-se o contraste abaixo:

- (i) a. *Nada ninguém disse a respeito de Oto.
b. Ninguém nada disse a respeito de Oto.

À primeira vista, a má-formação de (ia) é um problema de superioridade (cf. Rizzi (1990)).

neste aspecto, a comparação com a categoria CP é pertinente: como se sabe, em várias línguas, existem construções com múltiplos sintagmas-WH, que foram analisadas através de uma regra de Absorção na posição de especificador de CP (cf. Higginbotham e May (1981); Aoun, Hornstein e Sportiche (1981); Moritz e Valois (1994)). Nos termos de Chomsky (1995), podemos dizer que, em (52), o traço [+NEG], que é interpretável, pode ser checado por um item, não ser apagado e permanecer visível para a checagem de outro item.

O segundo problema que vamos examinar diz respeito às frases seguintes do inglês:

- (53) a. I see nobody at the party.
b. I buy nothing in the shop.

As orações acima colocam dificuldades para a nossa análise pelas seguintes razões: concluímos que o inglês dispõe de um traço [+NEG] Forte que exige a checagem antes de Spell-Out. Ora, as frases acima parecem mostrar o contrário, ou seja, os traços negativos de *nothing* e *nobody*, provavelmente, alçam para o domínio de checagem de NEG na sintaxe encoberta. A suposição que se pode fazer é que as orações em (53) pertencem ao estágio 3 do Ciclo de Jespersen descrito em (43), isto é, é provável, mas não será investigado aqui, que orações como (53) conviviam com frases do tipo de *I see not Mary*. Tratar-se-ia, portanto, de resíduos de uma época em que a categoria NEG do inglês era Fraca. O mesmo pode ser proposto para orações como *He is not happy* que apresentam verbo cópula.

As observações a respeito de (53) nos permitem concluir que o apelo à noção de gramaticalização pode ainda nos ser útil como critério para decidir se um determinado fato lingüístico é de natureza residual e, assim, localizá-lo em relação aos processos de mudança nos quais a língua se encontra.

CAPÍTULO 10

O PERCURSO DE *SENHOR*¹

Por que estudar formas de tratamento? De acordo com Luft (1957: 193), seria de grande interesse o estudo *das fórmulas de saudação, despedida, desculpa, pedido de permissão ou licença, enfim todas as manifestações de urbanidade (...), que numa determinada sociedade [são impostas] ao sujeito falante, como série de conveniências ou mesmo de obrigações, às quais ele não pode se furtar sem a reprovação social, sem a pecha de incivil, de mal educado.*² Por serem formulaicas, tais expressões, à primeira vista, parecem não constituir um objeto adequado para o estudo da gramaticalização. Esta impressão inicial mostra-se, entretanto, enganadora. Tais fórmulas sofrem alterações no decorrer do tempo, tanto em relação ao contexto em que são usadas quanto ao modo como se realizam morfofonologicamente e por isso é possível recompor o caminho por elas percorrido no sistema lingüístico. E, desse modo, capturar a correlação entre estes dois sistemas.

O presente trabalho vai se desenvolver da seguinte maneira. Iremos apresentar dados sobre a história do item *senhor* e

¹ Este texto constitui uma versão de uma palestra apresentada na USP em 2000, durante estágio de Pós-Doutoramento pelo CNPq, por Jânia Ramos.

² Luft, C. P. Tratamento depreciativo. *Revista Brasileira de Filologia*, v. 3 (II), p. 193-207, 1957.

documentar a ocorrência de formas foneticamente reduzidas. Através da aplicação de testes, argumentaremos a favor do caráter pronominal dessas últimas formas. Desse modo, buscamos explicitar as *clines* (1) e (2).

(1) Senhor >... > sô

(2) Senhora >... > só

Vamos mostrar ter havido perda fonética desses itens e alteração tanto no estatuto gramatical quanto em seu conteúdo semântico. As alterações serão descritas como perda de traços, mais exatamente, [pessoa], [número], [número] e [especificidade].

1. SOBRE A HISTÓRIA DO ITEM *SENHOR*

Cintra (1961) recompõe o percurso do item *senhor*. Mostra que essa forma de tratamento é explicitamente recomendada ao rei, *expressa[ndo] as noções de acatamento e reverência [a ele] associadas. Em 1597, o rei Felipe II ordena que alto das cartas, ou papéis que se me escrevem, se ponha: Senhor, sem outra cousa.*³ Tal instrução certamente visava a simplificar os vocativos presentes nas cartas, que na época eram longos e variados. Nessa mesma legislação inclui-se, dentre outras, a forma *Vossa Senhoria*, com que se deveria tratar duques, marqueses, condes e autoridades religiosas (bispos e arcebispos). Em 1739, Dom João V também legisla sobre o tratamento por *Senhoria*, ampliando seu uso em relação à lei filipina. Essa ampliação é interpretada por Cintra (*op.*

³ A Lei de Felipe II está editada nas *Provas de História Genealógica da Casa Real Portuguesa*, de Antônio Caetano de Souza, VI, Lisboa, 1745, p. 287-290, *apud* Corradello, 1997, p. 14, nota 5.

cit.: 28) como uma evidência do *assalto a que esta fórmula tinha estado submetida por camadas cada vez mais vastas da sociedade do tempo*. A extensão a outras camadas sociais foi ainda mais longe. Conforme relata Filinto Elíseo, no século XVIII, o tratamento por *senhor* era já comum na fala do povo em geral.

Esta extensão de uso implicou uma extensão semântica capturada através dos diferentes significados atribuídos ao item *senhor* nos dicionários, neste período de tempo. Etimologicamente, o item *Senhor* vem do latim *senior, senioris*, mais velho, mais idoso, comparativo de superioridade de *senex, velho*.⁴ Adquiriu o significado de proprietário, dono absoluto, aquele que tinha autoridade feudal sobre certas pessoas ou propriedades.⁵ Para Said Ali (p. 97), esse conceito *ter-se-ia desviado para o idealismo quando, por temor e respeito, se começou a chamar senhores aos fidalgos e poderosos em geral cujos atos e abusos as ordenações régias tiveram de limitar e coibir*. No momento em que passa a ser usado correntemente pelo povo, *senhor* já se referia à distância social entre interlocutores, evidenciando relações assimétricas de poder e respeito.

Destas informações se podem depreender dois fatos importantes. O primeiro é que *senhor* e *vossa senhoria* são usados independentemente desde o século XVIII, já que o rei deve ser tratado como *senhor* enquanto o tratamento por *vossa senhoria* é reservado a outros membros da Corte. Teria havido, portanto, uma ruptura entre ter senhoria e ser senhor. O segundo fato é a ocorrência de extensão de sentido do item *senhor*: *homem idoso > homem possuidor de bens e pessoas > homens poderosos > homens de respeito*. Em outras palavras, o objeto de posse inicialmente é concreto e, no segundo momento,

⁴ BUENO, Francisco da S. *Grande dicionário etimológico-prosódico da língua portuguesa*. São Paulo: Saraiva, 1967. p. 3700.

⁵ MORAIS SILVA, Antônio. *Grande dicionário da língua portuguesa*, 10ª edição revista, corrigida e muito aumentada e actualizada. Lisboa: Editorial Confluencia, 1957. p. 50-1.

abstrato. Esse processo de mudança semântica poderia ser descrito como metafórico.⁶ Veja-se, entretanto, que uma descrição estritamente gramatical obscurece a inter-relação entre os sistemas social e lingüístico.

2. O PERCURSO LINGÜÍSTICO

A título de exemplificação, vejamos alguns fragmentos de diálogos.⁷

(3) O padre Senhor, aqui estaes vos? (XIV)⁸

(4) Dá-me cá no coração/ que é vindo o Senhor pai/ com o Senhor seu irmão (XVI)⁹

(5) Deus venha có o senhor pae/ e o traga na sua guia (XIX)¹⁰

(6) Meu senhor, venha cá.

(7) Eu já falei para o Senhor, mas parece que não ouviu.

(8) Os senhores desejam mais café?

⁶“Metaphorical change involves specifying , usually more complex, thing in terms of another not present in the context. (...) metaphor is correlated primarily with solving the problem of representation (...)”. HOPPER, Paul J.; TRAUGOTT, E. C. *Grammaticalization*. Cambridge/Melbourne: Cambridge University Press, 1993. p. 87.

⁷ Para maiores detalhes do uso de senhor em diálogos entre pais e filhos, ver Ramos (2000).

⁸ Fernão Lopes, Crônica de Dom João II, citado por Luz, *op.cit.*: 99)

⁹ Camões, Acto V, cena 1, citado por Cláudio Basto (1931:187).

¹⁰ Almeida Garrett, 1875, citado por Cláudio Basto (1931:186).

No português medieval, usava-se vós no tratamento de filho para pai. Na prosa quatrocentista, surgem a cada passo expressões de tratamento mais respeitosas, facto que a lei assinala (Ver Partidas de Afonso X, f. 30).¹¹ O item *senhor* aparece inicialmente à direita dos nomes pai/padre e mãe/madre (como em 1-2), e depois passa a aparecer à esquerda,¹² (como em 5-6). Com base nesta documentação, podem-se fazer algumas considerações sobre o estatuto gramatical da forma de tratamento *Senhor*. Sua ocorrência inicialmente à direita do nome e posteriormente à esquerda permite supor que seu estatuto sintático seja o de modificador do nome. Já sua ocorrência isolada, em momento posterior, permite atribuir-lhe um estatuto de núcleo nominal. Comparem-se os vocativos (3-5) com (5-8).

Já as formas morfofonologicamente reduzidas adquirem outro estatuto gramatical, mais exatamente o de pronome. Diferentemente dos nomes, os pronomes não podem ser precedidos de determinantes, pois são eles mesmos determinantes. Daí a má-formação de (9b-c) e (10b).

- (9) a. O Senhor João está aqui.
b. ?O Sô João está aqui.
c.*A Sá Maria está aqui.
- (10) a. O que é isso, meu Senhor.
b. * O que é isso, meu sô?
c. ? O que é isso, Senhor?
d. O que é isso, Senhor João?
e. O que é isso, sô?

¹¹ Luz, *op. cit.*: 109.

¹² *Idem.*

Adjetivos e numerais podem ser antepostos ou pospostos a pronomes, mas nunca co-ocorrerem nas duas posições.

(11) *Senhor idoso* não precisa pagar passagem. (nome)

(12) *Sô* esperto do jeito que é sabe a resposta.

(13) *Só* esperta do jeito que é já sabe a resposta.

(14) *Um senhor idoso* não precisa pagar passagem.

(15) * *Um sô* esperto do jeito que é já sabe a resposta.

(16) * *Uma só* esperta do jeito que é já sabe a resposta.

Os pronomes ocupam posições mais fixas que os nomes.

(17) (a) *O senhor* vai sair agora?

(b) *Sô* vai sair agora?

(c) *Só* vai sair agora?

(18) (a) Aqui que eu comprei pro *senhor*.

(b) *Aqui que eu comprei pro *sô*.

(c) *Aqui que eu comprei pra *só*.

(19) (a) Ele tem medo de sair sem *o senhor*?

(b) *Ele tem medo de sair sem *sô*?

(c) *Ele tem medo de sair sem *só*?

(20) (a) Minha Senhora, assente-se aqui./Meu Senhor.

(b) Senhora, assente-se aqui./ Senhor, assente-se aqui.

(c) Sá, senta aqui.

(d) ô, Sá, senta aqui. / Ô, sô, senta aqui.

A comparação das diferentes realizações de *Senhor* mostra ter havido não só perda de substância fônica como também alteração de seu estatuto sintático: conforme vimos nem todas as diferentes realizações conduzem a estruturas bem formadas.

A rejeição de *sá* ao artigo permite explicitar seu caráter mais gramaticalizado do que *senhor* e *sô*. Pode-se observar que estes itens foram morfologizados e por isso não mais carregam o conteúdo semântico de expressão de respeito e distância.¹³

Moura (2004: 29) documenta certa “generalização” quanto ao conteúdo semântico de *sô*. *A forma reduzida sô, na função de vocativo, é atribuída a pessoas, indiferentemente a sexo e idade, ou seja, essa forma é direcionada a velhos, adultos ou crianças de ambos os sexos (...)*. Vejam-se os fragmentos de diálogos ocorridos em Belo Horizonte, no ano de 2003.

(21) A filha disse para a mãe:

– deixe eu ir, *sô* (A.S.A. 5 anos, *apud* Moura, p. 30)

(22) Conversa entre duas jovens:

– não *sô*, assim não. (L.O., 29 anos, *apud* Moura, p. 30)

(23) Fala do pai para o filho de 8 anos de idade.

– Larga esse portão, *sô*. (*apud* Moura, p: 30)

Se, agora, tivermos em conta o tipo de segmento fonológico que sofreu apagamento ou duplicação, poderemos perceber dois

¹³ Confirmando o que assinala Hopper e Traugott: “As grammaticalized forms become increasingly syntactized or morphologized they unquestionably cease over time to carry significant semantic or pragmatic meaning. (...)”. HOPPER, Paul J.; TRAUGOTT, E. C. *Grammaticalization*. Cambridge/Melbourne: Cambridge University Press, 1993. p. 87.

padrões distintos entre o português europeu e o português brasileiro, representados, respectivamente, em (24) e (25), abaixo. Somente em (25) há duplicação e apagamento quase total de [r].

(24) *sinhor, siôr, siô, sôr, sô, s'nhor, nhor.*¹⁴

(25) *sinhô, siô, nhôr, nhozinho, nhonhô, nhô, sô, seu, sá.*¹⁵

No que diz respeito à evolução de sete séculos do item *Senhor*, vemos que a forma inicial ainda co-ocorre com as novas variantes. Com frequência, uma forma de um estágio posterior acarreta a exclusão da forma antiga, tal como no par *vossa mercê/você*. No que diz respeito a *Senhor*, a forma inicial se mantém. Por quê? Estudos sociolinguísticos mostram que a avaliação social é um fator que pode reverter processos de mudanças, cancelando-os ou tornando-os mais lentos ou acelerados. A resistência de *Senhor*, mantendo-se ao lado das formas reduzidas no PB atual, parece recorrer do estatuto de desprestígio dessas últimas. As formas reduzidas são reconhecidas pelos falantes como marcas de dialeto rural e/ou ausência de escolaridade.

3. HIPÓTESES PARA EXPLICAR MUDANÇAS

Documentada a ocorrência de mudanças lingüísticas, importa agora buscar responder à seguinte questão: por que houve tais mudanças?

Inicialmente tenhamos em conta que formas de tratamento como *senhor* e outras tiveram sua entrada no sistema nos contextos lingüísticos em que se usava *vós*, conforme já mencionamos. Em outras palavras, parece ter havido, num primeiro momento, um processo de

¹⁴ Biderman, p. 370.

¹⁵ *Idem.*

variação em que se referia a *senhor(a)* através de pronomes de segunda pessoa e de terceira pessoa, como respectivamente em:

(26) “*Senhores, peço-vos hu~ Dom: que me outorguedes o que vos quero pedir*”.¹⁶

(27) Senhores, venham cá.

Uma explicação para a queda de uso de *vós* é formulada por Cintra (*op. cit.*: 36), em termos de uma tendência geral à simplificação: *sua substituição por um tratamento que conduzia o verbo para a terceira pessoa foi certamente favorecida por uma tendência para simplificar; num sector em que a gramática portuguesa se apresenta particularmente complexa: a flexão verbal, extremamente rica em formas bem diferenciadas*. Há, entretanto, várias evidências contrárias a essa hipótese. A primeira é que fenômeno semelhante teria ocorrido, simultaneamente, no espanhol. E sendo o espanhol uma língua flexionalmente rica, fica difícil aceitar que tal sistema optaria por uma solução isolada de simplificação. Veja-se que o espanhol e mesmo o português europeu são línguas que mantiveram flexão diferenciada em todas as outras pessoas do discurso. Outra evidência contrária à hipótese é que os pronomes de tratamento são inicialmente usados com o verbo na segunda pessoa do plural, como em (27), conforme vimos acima. Prado (1975), ao estudar as formas de tratamento nos séculos XIV, XV e XVI, observa que o uso com o verbo na terceira pessoa ainda não se manifestava nos textos.

Feitas estas considerações, parece-me mais plausível supor que o próprio uso de *vós*, em contraposição ao *tu*, já indicava uma tendência de tratar indiretamente o interlocutor. O uso do plural já atribuía ao interlocutor uma nova dimensão, representava *uma alusão a mais, [daí] ser empregado com gente de alta jerarquia(sic) enquanto “tu”*

¹⁶ Prado, p. 82-83.

*ficou para os casos gerais.*¹⁷ Se entendermos o plural como um distanciamento, o uso da forma de tratamento com verbo na terceira pessoa constituirá apenas um passo a mais na mesma direção. Deu-se andamento a um processo no qual *a pessoa que fala se dirigia ao interlocutor – no caso, o rei – , não diretamente, como se fazia na vigência do pronome “tu” , mas indiretamente, através da qualidade do atributo peculiar.* Dizer que *Vossa Majestade está servido* significava dizer *A Majestade de vós será servido.*¹⁸ De modo análogo, ao dizer-se *Vossa Senhoria*, está-se dirigindo à autoridade de Senhor que o interlocutor possui, uma vez que *Senhoria* refere-se à *autoridade ou qualidade de Senhor.*¹⁹

No momento posterior, a partir do século XIX, em que as diferentes formas indiretas de tratamento vão atingindo as camadas populares, *senhor* apresenta um ponto em comum com *você*: ambas as formas mantêm o tratamento indireto de segunda pessoa, podendo ser usadas sem mencionar o nome do interlocutor. Nesse sentido são ambas pronominais, conforme mostram, de modo bastante claro, os contrastes apresentados em (9-10).²⁰

Discutiu-se o estatuto gramatical da forma *senhor* e das formas *sô*, *sá*, concluindo-se que estas duas últimas adquiriram o estatuto de pronome. Tal estatuto fica evidente quando se aplicam testes morfossintáticos. Com base neles, identificam-se contrastes através dos quais se chega a propriedades formais, conforme vimos aqui.

¹⁷ MACHADO FILHO, Ayres M. *Português fora das gramáticas*. 2ª ed. Rio de Janeiro, 1965. p. 137.

¹⁸ *Idem.*

¹⁹ MORAIS SILVA, A. 1957. p. 55.

²⁰ Ver argumentação convincente em MOURA, Lillian. *Redução Fonológica do item Senhor e seu estatuto sintático*. Monografia de Bacharelado, Universidade Federal de Minas Gerais, Belo Horizonte, 2004. Seguem-se aqui algumas ocorrências colhidas em entrevistas livres.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABAURRE, M. B.; GALVES, C. As diferenças rítmicas entre o português europeu e o português brasileiro: uma abordagem otimalista e minimalista. *D.E.L.T.A.*, 14, 2, p. 377-402, 1998.
- ALKMIN, M. *Negativas sentenciais no dialeto de Ouro Preto: uma abordagem sociolinguística*. Tese (Doutorado), UFMG, Belo Horizonte, 2001.
- ALVES, N. *As formas você e cê e a indeterminação do sujeito no português brasileiro*. Dissertação (Mestrado), UFMG, Belo Horizonte, 1998.
- AMARAL, A. *O dialeto caipira*. São Paulo: Secretaria da Cultura, Ciência e Tecnologia, [1923], 1976.
- ANDERSON, S. *A-Morphous Morphology*. Cambridge: Cambridge University Press, 1992.
- AOUN, J.; HORNSTEIN, N.; SPORTICHE, D. Some Aspects of Wide Scope Quantification. *Linguistic Inquiry*, 18, p. 537-577, 1981.
- BASSANI, D. *Arceburgo: a presença italiana em Minas*. Monografia (Iniciação Científica) - inédito, UFMG, Belo Horizonte, 2002.
- BASTO, Cláudio. *Formas de tratamento em Postuguês*. Revista Lusitana XXIX: 182-202.

- BENMAMOUN, A. *Inflectional Morphology: Problems of Derivation And Projection*. Unpublished MS. University of South California, 1990.
- BIDERMAN. Formas de Tratamento. *Alfa*, 18-19, p. 339-381, 1972.
- BITTENCOURT, V. Gramaticalização e discursivização no português oral do Brasil: o caso “tipo (assim)”. *Scripta*, 4, n. 2, p. 23-38, 1999.
- BONET, E. Feature Structure of Romance Clitics. *Natural Language & Linguistic Theory*, 13, 4, p. 607- 647, 1995.
- BRANDÃO DE CARVALHO, J. Phonological Conditions on Portuguese Clitic Placement: on Syntactic Evidence for Stress and Rhythmical Patterns. *Linguistics*, 27, p. 405-436, 1989.
- BUENO, F. da S. *Grande dicionário etimológico-prosódico da língua Portuguesa*. São Paulo: Saraiva, 1967. p. 3700.
- BYBEE, J.; HOPPER, P. *Frequency and the Emergence of Linguistic Structure*. Amsterdam/Philadelphia: John Benjamins, 2001.
- _____; PERKINS, R.; PAGLIUCA, W. *Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press, 1994.
- CAMARGOS, M. *Negativas sentenciais no dialeto mineiro*. Monografia (Iniciação Científica), UFMG, Belo Horizonte, 1998.
- _____. *Negativa [V não]: uma abordagem minimalista*. Dissertação (Mestrado), UFMG, Belo Horizonte, 2002.
- CAMPBELL, Lyle; JANDA, Richard. Introduction. *Language Sciences* 22.4 [Special Issue on *Grammaticalization: A Critical Assessment*; L. Campbell (Ed.)], p. 1-31, 2001.
- CARDINALETTI, A.; STARKE, M. The Typology of Structural Deficiency on the Three Grammatical Classes. In: RIEMSDIJK, H. (Ed.). *Clitics in the Languages of Europe*. Berlin/New York: Mouton de Gruyter, 1999.

- CARDOSO, E. *O crioulo da Ilha de S. Nicolau e Cabo Verde*. Lisboa: Instituto Cultural de Língua Portuguesa, 1989.
- CASTILHO, A. A gramaticalização. *Estudos lingüísticos e literários*. Salvador: Universidade Federal da Bahia, n.19, p. 25-64, 1997.
- CASTRO, Ivo. Para uma história do português clássico. In: DUARTE, I.; LEIRIA, II (Org.). *Actas do Congresso Internacional sobre o português*. Lisboa: Colibri, 1996.
- CHAMBERS, J. K. *Sociolinguistic Theory*. Oxford/Cambridge: Blackwell, 1995.
- CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris Publications, 1981.
- _____. *Knowledge of Language*. New York: Praeger, 1986.
- _____. Some Notes on Economy of Derivation and Representation. In: FREIDEN, R. (Ed.). *Principles and Parameters in Comparative Grammar*. Cambridge: The MIT Press, 1991.
- _____. *The Minimalist Program*. Cambridge: The MIT Press, 1995.
- CINQUE, G. On si Constructions and the Theory of Arb. *Linguistic Inquiry*, 19, p. 521-581, 1988.
- CINTRA, L. F. L. *Sobre forma de tratamento na língua portuguesa*. Lisboa: Livros Horizonte, 1961.
- COELHO, M. do S. Uso de você no Norte de Minas. Dissertação (Mestrado), UFMG, Belo Horizonte, 1999.
- COHEN, M. A. Gramaticalização e reanálise na língua portuguesa: um estudo diacrônico. *Caligrama*, 1, p. 543-552, 1988.
- COHEN, M.; RAMOS, J. (Org.). *Dialeto mineiro e outras falas*. Belo Horizonte: FALE/UFMG, 2002. *Corpus Anotado de Português Clássico Tycho Brahe*, coordenado por Charlotte Galvez, USP e UNICAMP.

- CORRADELLO, M. *Quem é você?* Análise de um pronome pessoal. Dissertação (Mestrado), UNICAMP, Campinas, 1997.
- CORRÊA, L. T. *A forma clítica de pronome pessoal no dialeto mineiro: uma variante sociolinguística.* Dissertação (Mestrado), UFMG, Belo Horizonte, 1998.
- COSTA, J. Adverb Positioning and V-Movement in English: Some More Evidence. *Studia Linguistica*, 50, 1, p. 22-34, 1996.
- COUTINHO, I. da S. *Gramática histórica.* Rio de Janeiro: Ao Livro Técnico, 1954.
- CYRINO, S.; DUARTE, M. E.; KATO, M. Visible Subjects and Invisible Clitics in Brazilian Portuguese. In: KATO, M.; NEGRÃO, E. (Org.). *Brazilian Portuguese and the Null Subject Parameter.* Vurvuert: Iberoamericana, 2000.
- CYRINO, S. M. L. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: ROBERTS, I.; KATO, M. (Org.). *O português brasileiro: uma viagem diacrônica.* Campinas: Editora UNICAMP.
- DAVIDSON, D. *Inquiries into Truth and Interpretation.* Oxford: Oxford University Press, 1984.
- DI CHRISTO, A. *De la microprosodie à L'intonosyntaxe.* Aix-en-Provence: Publications Université Aix-en-Provence, 1985.
- DICKEY, E. Forms of address and terms of reference. *Lingua*, 33, p. 255-274, 1997.
- DOBROVIE-SORIN, C. Clitics Across Categories: The Case of Romanian. In: RIEMSDIJK, H. (Ed.) *Clitics in the Languages of Europe.* Berlin/New York: Mouton de Gruyter, 1999.
- DUARTE, M. E. L. *Variação e sintaxe: clítico acusativo, pronome lexical e categoria vazia no português do Brasil.* Dissertação (Mestrado), PUC/SP, São Paulo, 1986.

- _____. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; KATO, M. (Org.). *O português brasileiro: uma viagem diacrônica*. Campinas: Editora UNICAMP, 1993.
- _____. *A perda do princípio "Evite Pronome" no PB*. Tese de Doutorado, UNICAMP, Campinas, 1995.
- _____. *A sociolinguística paramétrica: perspectivas*. Comunicação apresentada no 1º Seminário Nacional de Estudos Linguísticos, João Pessoa, 1997.
- DUTRA, R. *A Gramática do português oral: folha de exemplos*. Artigo inédito. UFMG, Belo Horizonte, 1991.
- ERNST, T. Negation in Mandarin Chinese. *Natural Language and Linguistic Theory*, 13, p. 665-707, 1995.
- FARACO, C. A. *The Imperative Sentences in Portuguese: a Semantic and Historical Discussion*. Phd Dissertation, University of Stanford, UK, 1982.
- FERREIRA, M. B. *Argumentos nulos em português brasileiro*. Dissertação (Mestrado), UNICAMP, Campinas, 2000.
- FILETTI, E. *O fenômeno de objetos implícitos nos verbos do português do Brasil*. Dissertação (Mestrado), UFMG, Belo Horizonte, 1999.
- FONTANA, J. *Phrase Structure and the Syntax of Clitics of Spanish*. Phd Dissertation, University of Pennsylvania, Pennsylvania, 1993.
- _____. On the Integration of Second Position Phenomena. In: KEMERADE, A. van; VINCENT, N. (Ed.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 1997.
- FROTA, S.; VIGÁRIO, M. *Aspectos de prosódia comparada: ritmo e entoação no PE e no PB*. Artigo inédito. Universidade de Lisboa, Lisboa, 1999.

- GALVES, C. C. *V-Movement Levels of Representation and the Structure of S*. UNICAMP. Comunicação apresentada no XIII GLOW, Cambridge, 1990.
- _____. O enfraquecimento da concordância no português brasileiro. In: ROBERTS, I.; KATO, M. (Org.). *O português brasileiro: uma viagem diacrônica*. Campinas: Editora UNICAMP, 1993. p. 387-408.
- _____. Relatório nº 1 do projeto temático “Padrões Rítmicos, Fixação de Parâmetros e Mudança Lingüística”. UNICAMP, Campinas, 1999.
- _____; BRITTO, H. *A construção do corpus anotado do português Histórico Tycho Brahe e o sistema de anotação morfológica*. Comunicação apresentada no IV PROPOR, 1999.
- _____. *Ensaio sobre as gramáticas do português*. Campinas: Editora UNICAMP, 2000.
- GIVÓN, T. *Toward a Discourse Definition Syntax*. Unpublished MS. University of California at Los Angeles, 1974. *Apud* Schwegler, 1991.
- _____. *On Understanding Grammar*. New York: Academic Press, 1979.
- GONÇALVES, F. M. R. *Negação frásica em português: caracterização sintáctica com referência ao processo de aquisição*. Dissertação (Mestrado), Universidade de Lisboa, Lisboa, 1994.
- GONÇALVES, V. *Aspectos da gramaticalização no português*. Dissertação (Mestrado), UFMG, Belo Horizonte, 1987.
- GREEN, G. Main Clause Phenomena in Subordinate Clauses. *Language*, 52, p. 382-397, 1976.
- GRIMSHAW, A. D. Unawareness without Ignorance: Unexploited Contributions in Sociology. *Contemporary Sociology*, 20, p. 843-848, 1991.

- HAEGEMAN, L.; ZANUTTINI. *Negative Heads and Negative Concord*. Unpublished MS. University de Genève, 1991.
- HALLE, M.; MARANTZ, A. Distributed Morphology and the Pieces of Inflection. In: HALE, K.; KEYSER, S. J. (Ed.). *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge/Mass.: The MIT Press, 1993.
- HALPERN, A.; FONTANA, J. X^o and Xmax Clitics. *Proceedings of WCCFL*, 12, p. 251-266, 1993.
- _____ ; ZWICKY, A. (Ed.) *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications, 1996.
- HARRIS, A.; CAMPBELL, L. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press, 1995.
- HEINE, B.; HÜNNEMEYER, B.; CLAUDI, U. *Grammaticalization: a Conceptual Framework*. Chicago: The University of Chicago Press, 1991.
- _____ ; REH, M. Grammaticalization and Reanalysis. In: *African Languages*. Hamburg: Helmut Booke Verlag, 1984.
- HIGGINBOTHAM, J.; MAY, R. Questions, Quantifiers and Crossing. *The Linguistic Review*, 1, p. 41-79, 1981.
- HOCK, H. H. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter, 1991.
- HOLMBERG, A.; PLATZACK, C. *On the Role of Inflection in Scandinavian Syntax*. Unpublished MS. University of Stockholm, 1988.
- HOPPER, P. J. On Some Principles of Grammaticalization. In: TRAUGOTT, E.; HEINE, B. (Ed.). *Approaches to Grammaticalization I*. Amsterdam: John Benjamins, 1991. p. 17-35.
- HOPPER, P.; TRAUGOTT, E. *Grammaticalization*. Cambridge: Cambridge University Press, 1993.

- HUBER, J. *Altportugiesisches Elermentrbuch*. Heidelberg: Carl Winters Universitätsbuchhandlung, 1933. Trad. Portuguesa: DELILLE, M. M. G. *Gramática do português antigo*. Lisboa: Fundação Calouste Gulbenkian, 1986.
- ILARI, R. Locuções negativas polares: reflexões sobre um tema de todo mundo. *Linguística: questões e controvérsias*, Série Estudos 10, Uberaba, 1984.
- ILARI, R. *et al.* Considerações sobre a posição dos advérbios. In: CASTILHO, A. (Org.). *Gramática do português falado*. 2 ed. Campinas: Editora UNICAMP, 1991. v.1.
- JACKENDOFF, R. *Semantic Interpretation in Generative Grammar*. Cambridge: The MIT Press, 1972.
- JANDA, Richard. Beyond "Pathways" and "Unidirectionality": On the Discontinuity of Language Transmission and the Counterability of Grammaticalization. *Language Sciences*, 22, 4, 2001.
- JERPERSEN, O. The philosophie de la grammaire. Tradução francesa de The Philosophy of Grammar. Paris: Editions de Minuit, 1971.
- _____. *Negation in English and other Languages*. Kobenhavn: Bianco Lunos Bogtrykkeri, 1917.
- KATO, M. *Português brasileiro falado*: aquisição em contexto de mudança lingüística. Artigo inédito. Congresso Internacional sobre o Português, Lisboa, 1994.
- _____. Strong and Weak Pronominals in the Null Subject Parameter. *Probus*, 11, p. 1-37, 1999.
- KAYNE, R. *French Syntax: the Transformational Cycle*. Cambridge: The MIT Press, 1975.
- _____. Chains, Categories External to S and French Complex Inversion. *Natural Language and Linguistic Theory*, 1, p. 107-139, 1983.

- _____. *Facets of Romance Past Participle Agreement*. In: BENINCÁ, P. (Ed.). *Dialect Variation and the Theory of Grammar*. Dordrecht: Foris, 1997.
- KEMENADE, A. van; VINCENT, N. (Ed.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 1987.
- _____. *The Antisymmetry of Syntax*. Cambridge/ Mass.: The MIT Press, 1994.
- KLAVANS, J. *Some Problems in a Theory of Clitics*. Bloomington: Indiana University Linguistics Club, 1982.
- KLIMA, E. Negation in English. In: FODOR, J. A.; KATZ, J. J. (Ed.). *The Structure of Language*. Prentice-Hall: Englewood Cliffs, N.J., 1964. p. 246-323
- KROCH, A. Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change*, 1, p. 199-244, 1989.
- KROCH, A.; TAYLOR, A. *Remarks on the XV/VX Alternation to Early Middle English*. Unpublished MS, University of Pennsylvania, 1994.
- _____. Verb Movement in Old and Middle English: Dialect Variation and Language Contact. In: KEMENADE, A. van; VINCENT, N. (Ed.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 1997. p. 297-325.
- KURYLOWICZ, J. *The Inflectional Categories of Indo-European*. Heidelberg: Carl Winter, 1964.
- LABERGE; SANKOFF, G. Anything YOU can do (1976). In: SANKOFF, G. *The Social Life of Language*. Philadelphia: University of Pennsylvania, 1980.
- LABOV, W. Contraction, Deletion, and Inherent Variability of the English. *Language*, 45, p. 715-762, 1969.
- _____. Negative Attraction and Negative Concord. *Language*, 48, p. 773-818, 1972.

- _____. Where Does the Linguistic Variable Stop? A Response to Lavandera. Working Papers. In: *Sociolinguistics*, 44, Austin: Southwest Educational Development Laboratory, 1978.
- _____. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell, 1994.
- LADUSAW, W. Polarity Sensitivity as Inherent Scope Relations. Phd Dissertation, The University of Texas at Austin, 1979.
- LAKA, I. *Constraints on Sentence Negation: the Case of Basque*. Unpublished MS, MIT, 1988.
- LEHMANN, C. Thoughts on Grammaticalization. A Programmatic Sketch. Unpublished MS. Köln: Arbeiten des Kölner Universalien-Projekts, 1982. v.1.
- LIGHTFOOT, D. *Principles of Diachronic Syntax*. Cambridge: The MIT Press, 1979.
- _____. *The Development of Language*. Oxford: Blackwell, 1999.
- LEHISTE, I. *Suprasegmentals*. Cambridge: The MIT Press, 1970.
- LUFT, C. Tratamento depreciativo. *Revista Brasileira de Filologia*, v. 3, p. 193-207, 1957.
- LUZ, Marilina Santos. *Formas de Tratamento no Português Arcaico: subsídios para o seu estudo*. Revista Portuguesa de Filologia VII, VIII e IX, 1956-8.
- MACHADO FILHO, A. *Português fora das gramáticas*. 2 ed. Rio de Janeiro, 1965.
- MAHAJAN, A. K. LF Conditions on Negative Polarity Licensing. *Lingua*, 80, 1990. p. 333-348.
- MARTELOTTA, M.; VOTRE, S.; CEZARIO, M. M. (Org.). *Gramaticalização no português do Brasil – uma abordagem funcional*. Rio de Janeiro: Tempo Brasileiro, 1996.
- MARTIN, J. GÊNERO? *Revista Brasileira de Lingüística*, 2. Rio de Janeiro, p. 3-8, 1968.

- MARTIN, R. *La négation de virtualité du moyen français*.
România, 1992. p. 20-49.
- MARTINS, A. M. *Clíticos na história do português*. 1994. Dissertação de Doutorado, Universidade de Lisboa.
- MARTINS, E. E. *Sentential Negations in Spoken Brazilian Portuguese*. Master's Research Paper. Washington: Georgetown University, 1997.
- MATA, A. *A variável de negação no dialeto pessoense*. Dissertação (Mestrado em andamento), s/d.
- MATOS; SILVA, R.V. *Estruturas trecentistas: elementos para uma gramática do português arcaico*. Lisboa: Imprensa Nacional/Casa Da Moeda, 1990.
- MATTOSO CAMARA JR., J. *História e estrutura da língua portuguesa*. Rio de Janeiro: Editora Padrão, 1979.
- MEILLET, A. L'evolution des formes grammaticales. In: *Esquisses linguistiques*. Munich: Fink, 1958. p. 38-54.
- _____; VENDRYES, J. *Traité de grammaire comparée des langues classiques*. 2e éd. Paris: Lib. Honoré Champion, 1948.
- MENON, O. P. S. *Analyse sociolinguistique l'indetermination du sujet dans le portugais parlé au Brésil, à partir des données du NURC/SP*. Université de Paris VII. Tèse (Doctorat), 1994.
- MILNER, J-C. *Ordres et raisons de langue*. Paris: Le Seuil, 1982.
- MILROY, J. *Linguistic Variation and Change*. Blackwell: Oxford, 1992.
- MIOTO, C. *Negação sentencial no português brasileiro e o critério WH*. Tese (Doutorado), UNICAMP, Campinas, 1992.
- MOIGNET, G. L'opposition Non/Ne en ancien français. In: *Travaux Linguistiques Littéraires*. Strasbourg, p. 41-65, 1965.
- MONACHESI, P. *A Lexical Approach to Italian Cliticization*. Stanford: CSLI Publications, 1999.

- MORAIS E SILVA, A. *Grande dicionário da língua portuguesa - 10ª edição revista, corrigida, e muito aumentada e actualizada*. Lisboa: Editorial Confluência, 1957.
- MORITZ, L.; VALLOIS, D. Pied-Piping and Specifier-Head Agreement. *Linguistic Inquiry*, 25, 4, p. 667-707, 1994.
- MOURA, L. *Redução Fonológica do Item Senhor e seu Estatuto Sintático*. Monografia. UFMG, Belo Horizonte, 2004.
- MULLER, C. L'association négative. *Langue Française*, p. 59-94, 1984.
- MUSSAFIA, A. *O enclisi o proclisi del pronome personale atono quale oggetto*. Romania, 1898. p. 145-146.
- 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.
- NARO, A.; BRAGA, M. L. A interface sociolingüística/gramaticalização. *Gragoatá*, 9, p. 125-134, 2000.
- NASCIMENTO, M. do. *Sur la position du sujet dans le portugais du Brésil*. Tèse (Doctorat), Université de Paris VIII, 1984.
- NESPOR, M.; VOGEL, I. *Prosodic Phonology*. USA: Foris Publications, 1986.
- NEWMAYER, Frederick J. Deconstructing grammaticalization. *Language Sciences*, 23, p. 187-220, 2001.
- NUNES, J. M. *O famigerado SE: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. Dissertação (Mestrado), UNICAMP, Campinas, 1990.
- NUNES, J. J. *Compêndio de gramática histórica portuguesa*. Lisboa: Livraria Clássica Editora, 1945.
- OGANDO, V. A Colocación do pronome átono en relación co verbo no galego-portugués medieval. *Verba*, 7, p. 251-282, 1980.

- OLIVEIRA, M. A. Variável lingüística: conceituação, problemas de descrição gramatical e implicações para a construção de uma teoria gramatical. *D.E.L.T.A.*, n. 3, v. 1, p. 19-35, 1987.
- OLIVEIRA E SILVA, G. M. Escolarização. In: CARDOSO, S. (Org.). *Diversidade lingüística*. Salvador: Editora UFBA, 1996.
- OUHALLA, J. *Functional Categories and Parametric Variation*. London: Routledge, 1991.
- PAGOTTO, E. *A posição dos clíticos em português: um estudo diacrônico*. Dissertação (Mestrado), Universidade Estadual de Campinas, Campinas, 1992.
- _____. Clíticos, mudança e seleção natural. In: ROBERTS, I.; KATO, M. (Org.). *O português brasileiro: uma viagem diacrônica*. Campinas: Editora UNICAMP, 1993.
- PARCERO, L. M. de J. *Fronteamentos de constituintes no português dos séculos XV, XVI e XVII*. Dissertação (Mestrado), UFBA, Bahia, 1999.
- PAREDES SILVA, V. L. Variação e funcionalidade no uso de pronomes de 2ª pessoa do singular no português carioca. *Revista de Estudos da Linguagem*, v. 7, n. 2, p.121-138, 1998.
- PERINI, M. *Gramática do infinitivo português*. Petrópolis: Vozes, 1977.
- PINTO, I.; FIORETT, M. T. *Tutorial para o pacote Varbrul*. 1992.
- POLLOCK, J-Y. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry*, 20, 3, p. 365-474, 1989.
- PONTES, E. *Verbos auxiliares em português*. Petrópolis: Vozes, 1973.
- _____. *O tópico no português do Brasil*. Campinas: Pontes Editora, 1987.
- PONZO, E. *Aquisição de clítico em posição de sujeito no português brasileiro*. Tese (Doutorado em andamento), UFMG, Belo Horizonte, s/d.

- PRADO, J. A. A. *Sistema dos pronomes pessoais na prosa portuguesa do século XIV ao XVI*. Tese (Doutorado), USP, São Paulo, 1975.
- PROGOVAC, L. Clitics in Serbian/Croatian: Comp as the Second Position. In: HALPERN, A.; ZWICKY, A. *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications, 1996.
- QUEIROZ, S. M. M. *A língua do negro da Costa*. Um remanescente africano em Bom Despacho. Dissertação (Mestrado), UFMG, Belo Horizonte, 1984.
- RADANOVIC-KOCIC, V. The Placement of Serbo-Croatian Clitics: A Prosodic Approach. In: HALPERN, A.; ZWICKY, A. *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications, 1996.
- RAMOS, J. Mudança sintática e teoria gramatical. *Cadernos de Estudos Lingüísticos*, 20, p. 33-40, 1991.
- _____. *Marcação de caso e mudança sintática no português brasileiro*. Tese (Doutorado), UNICAMP, Campinas, 1992.
- _____. *A alternância entre "não" e "num" no dialeto mineiro: um caso de mudança lingüística*. Comunicação apresentada no Encontro da Abralín. Maceió, 1997.
- _____. O uso das formas você, ocê e cê no dialeto mineiro. In: D. DA HORA (Ed.). *Diversidade Lingüística no Brasil*. João Pessoa: Idéia Editora, 1997a. p.43-60.
- _____. *Formas de tratamento: a díade pai e filho*. 17 p. Inédito, 2000.
- RAMSDEN, H. *Weak-Pronoun Position in the Early Romance Languages*. Manchester: Manchester University Press, 1963.
- REIS, M. C. F. *Sintagmas adjetivais e sintagmas nominais em função adverbial: a questão dos traços de concordância*. Dissertação (Mestrado), UFMG, Belo Horizonte, 1997.

- RIBEIRO, I. A formação dos tempos compostos: a evolução histórica das formas ter, haver e ser. In: ROBERTS, I.; KATO, M. *O português brasileiro: uma viagem diacrônica*. Campinas: Editora UNICAMP, 1993.
- RICHARDS *et al.* *Dictionary of Language Teaching & Applied Linguistics*. England: Longman, 1995.
- RIEMSDJIK, H. (Ed.). *Clitics in the Languages of Europe*. Berlin/ New York: Mouton de Gruyter, 1999.
- RITTER, E. On the Syntactic Category of Pronouns and Agreement. *Natural Language and Linguistic Theory*, 13, p. 405-443, 1995.
- RIVERO, M. L. On Two Locations for Complement Clitic Pronouns: Serbo Croatian, Bulgarian and Old Spanish. In: KEMENADE, A. van; VINCENT, N. *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 1997.
- RIZZI, L. *Issues in Italian Syntax*. Dordrecht: Foris Publications, 1982.
- _____. *Relativised Minimality*. Cambridge: The MIT Press, 1990.
- _____. *Residual Verb Second and the Wh-Criterion*. Unpublished MS, Université de Genève, 1991.
- _____. On the Status of Subject Clitics in Romance. In: JAEGGLI; CORVALAN, C. S. (Org.). *Studies in Romance Linguistics*. Dordrecht: Foris Publications, 1986.
- ROBERTS, I.; KATO, M. (Org.). *O português brasileiro: uma viagem diacrônica*. Campinas: Editora UNICAMP, 1993.
- ROMANELLI, R. C. *O supletivismo indo-europeu na morfologia latina*. Belo Horizonte: Imprensa Universitária, 1975.
- RONCARATTI DE SOUZA, C. N. A negação do português falado. In: MACEDO, A. T. *et al.* *Variação e discurso*. Rio de Janeiro: Tempo Brasileiro, 1996.

- ROUVERET, A. Cliticisation et temps en portugais européen. *Revue des Langues Romanes*, XCIII, 2, 1989.
- _____. *La syntaxe des dépendances lexicales: identité et identification dans la théorie syntaxique*. Tèse (d'État), Université de Paris VII, 1987.
- SAID ALI, M. *Dificuldades da língua portuguesa*. Livraria Acadêmica: Rio de Janeiro, 1950.
- _____. *Gramática histórica da língua portuguesa*. Rio de Janeiro: Edições Melhoramentos, 1966.
- SALES FILHO, A. *A Negação em Vila dos Confins*. Rio de Janeiro: Tempo Brasileiro, 1980. (*Apud* RONCARATI, 1996).
- SANKOFF, G.; BROWN, P. The Origins of Syntax in Discourse: A Case Study of Tok Pisin Relatives. *Language*, v. 52, n. 3, p. 631-667, 1976.
- SANTOS, M. T. M. dos. *Uma análise espectrográfica dos sons fricativos surdos e sonoros do PB*. Dissertação (Especialização), Escola Paulista de Medicina, São Paulo, 1987.
- SCHERRE, M. *Introdução ao pacote Varbrul para microcomputadores*. Artigo inédito. Rio de Janeiro, 1993.
- SCHWEGLER, A. Predicate Negation in Contemporary Brazilian Portuguese – a Change in Progress. *Orbis*, 1991.
- STEVENS, K. N.; KLATT, D. H. *Studies of Acoustic Properties of Speech Sound*. Air Force Cambridge Research Laboratories (Scientific Report 8), Massachusetts, 1968.
- STAUF, I. *Recherches sur le "ne" redondant*. Paris: Rousseau, 1928.
- STOWELL, T. *Origins of Phrase Structure*. PhD Dissertation. Cambridge: MIT, 1981.
- SUNER, M. Negative Elements, Island Effects and Resumptive No. *The Linguistic Review* 12, p. 233-273, 1995.
- TARALLO, F. L. *Relativization Strategies in Brazilian Portuguese*. PhD Dissertation, University of Pennsylvania, 1983.

- _____. *A pesquisa sociolingüística*. São Paulo: Ática, 1985.
- _____; NUNES, J.; RAMOS, J. *A sociolingüística paramétrica*. Artigo inédito. Texto apresentado na Anpoll de Recife, 1990.
- _____; KATO, M. Harmonia trans-sistêmica: variação inter e intralingüística. In: *Preedição*, n. 5, Campinas, 1989.
- TAYLOR, A. *Clitics and Configurationality in Ancient Greek*. Doctoral Dissertation, University of Pennsylvania, 1990.
- TAYLOR, A. A Prosodic Account of Clitic Position in Ancient Greek. In: HALPERN; ZWICKY, A. (Ed.). *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications, 1996.
- THOMAS, E. *The Syntax of Spoken Brazilian Portuguese*. Nashville: Vanderbilt University Press, 1969. (Apud SCHWEGLER, 1991)
- TOBLER, A. Review of J. Le coultre, de l'ordre des mots dans chrétien de troyes. *Vermischte Beiträge zur Französischen Grammatik 5*. Leipzig: Hirzel, p. 395-414, 1912.
- TORRES DE MORAIS, M. A. *Do português clássico ao português moderno*. Tese (Doutorado), UNICAMP, Campinas, 1995.
- TRABALHO DE INICIAÇÃO CIENTÍFICA NAMIUTI, IEL/ UNICAMP. *Interpolação no português clássico*. Campinas, Unicamp, 2001.
- TRAUGOTT, E. Meaning-Change in the Development of Grammatical Markers. *Language Sciences*, 2, p. 44-61, 1980.
- _____; HEINE, B. (Org.). *Approaches to Grammaticalization*. Amsterdam: John Benjamins, 1991.
- TSIMPLI, I-M.; ROUSSOU, A. Negation and Polarity Items in Modern Greek. *The Linguistic Review*, 13, p. 49-81, 1996.
- URIBE-ETXEBARRIA, M. *Towards an LF Theory of Negative Polarity Licensing*. Unpublished MS, University of Connecticut, MIT, s/d.

- VENNEMANN, T. *Explanation in Syntax: Syntax and Semantics*. New York: Academic Press, 1973. v. 2. p. 1-50.
- VIANNA, H. L. G. *A estrutura modal + infinitivo em português: gramaticalização e modalização*. Dissertação (Mestrado), UFMG, Belo Horizonte, 2000.
- VINCENT, N. The Emergence of the D-System in Romance. In: KEMENADE, van; VINCENT, N. (Ed.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 1997.
- VIOTTI, E. *A sintaxe das sentenças existenciais no português do Brasil*. Tese (Doutorado), USP, São Paulo, 1999.
- VITRAL, L. *Sobre a complementação infinitiva em português*. Dissertação (Mestrado), UFMG, Belo Horizonte, 1987.
- _____. Structure de la Proposition et Syntaxe du Mouvement en Portugais Brésilien. Tèse (Doctorat nouveau regime). Université de Paris VIII, 1992.
- _____. A Forma cê e a noção de gramaticalização. *Revista de Estudos da Linguagem*, 5, p.115-124, 1996.
- _____. *A negação: teoria da checagem e mudança lingüística*. *D.E.L.T.A.*, 15, 1, p. 57-84, 1999.
- _____. *Sintaxe formal e gramaticalização: roteiro de uma pesquisa*. Artigo a ser publicado em NICOLAU, E. (Org.). *Estudos sobre a estrutura gramatical da linguagem*. FALE/UFMG, Belo Horizonte, 2000.
- _____. *A interpolação de 'se' e suas conseqüências para a teoria da cliticização*. *Revista da Abralin* 1, número 2, 2002: 161-197.
- _____. e RAMOS, J. *Gramaticalização de 'você': um processo de perda de informação semântica?* *Filologia e Lingüística* 3: 55-64, 1999.

- VOS, R.; VESELOVSKÁ, L. Clitic Questionnaire. In: RIEMSDIJK, H. (Ed.) *Clitics in the Languages of Europe*. Berlin/New York: Mouton de Gruyter, 2000.
- VOTRE, S. Gramaticalização de *achar*. In: *Cadernos de sociolingüística aplicada*. PUC, São Paulo, 1993.
- WILLIAMS, E. *Thematic Structure in Syntax*. Cambridge: The MIT Press, 1994.
- ZANUTTINI, R. The Structure of Negative Clause in Romance. Unpublished MS, University of Pennsylvania, 1989.
- _____. Re-Examining Negative Clauses. In: CINQUE, G.; KOSTER, J; POLLOCK, J-Y; RIZZI, L.; ZANUTTINI, R. (Ed.). *Paths Towards Universal Grammar Studies in Honor of Richard Kayne*. Washington D. C.: Georgetown University Press, 1994. p. 427-451.

Impresso em offset nas oficinas da
FOLHA CARIOCA EDITORA LTDA.
Rua João Cardoso, 23 – Tel.: 2253-2073
Fax.: 2233-5306 – Rio de Janeiro – RJ – CEP 20220-060

O tema *Gramaticalização* tem sido retomado nesses últimos anos, no Brasil e no estrangeiro, permitindo a abordagem de diversos fenômenos de linguagem numa perspectiva renovadora e estimulante. Sabemos, igualmente, que a perspectiva teórica desses novos trabalhos é, comumente, classificada como de orientação funcionalista.

Este livro, que é um registro de pesquisas na área de *Lingüística Teórica*, realizadas, a partir de 1996, na Universidade Federal de Minas Gerais, aposta, no entanto, que os processos de gramaticalização podem também ser abordados numa perspectiva formalista. Para isso, são utilizados pressupostos teóricos da Gramática Gerativa na descrição e análise de diversos fenômenos da sintaxe do português do Brasil. Tais análises se beneficiam também de informações de natureza quantitativa, depreendidas da Teoria da Variação.

ISBN 85-282-0138-4



9 788528 201383