

ESTUDOS

Lingüísticos e Literários

Número 19 / março 1997



Programa de Pós-Graduação
em Letras e Lingüística
Universidade Federal da Bahia

ESTUDOS

Lingüísticos e Literários

Estudos Lingüísticos e Literários, nº 19, Salvador, Programa de Pós-Graduação em Letras e Lingüística, Universidade Federal da Bahia, março 1997, 286p.

15,5 x 22,5 cm.

1. Letras - Periódicos I. Mestrado em Letras,
Universidade Federal da Bahia.

CDU 8 (05)

ISSN 0102-5465

ESTUDOS

Lingüísticos e Literários

Número 19 / março 1997



Programa de Pós-Graduação
em Letras e Lingüística
Universidade Federal da Bahia

ESTUDOS

Linguísticos e Literários

Número especial

PROGRAMA DE PÓS-GRADUAÇÃO
EM LETRAS E LINGÜÍSTICA

Universidade Federal da Bahia
PUBLICAÇÃO SEMESTRAL

Cordenador do PPGLL

Celia Marques Telles

Editor

Celina Scheinowitz

Co-Editor

Evelina Hoisel

Conselho Editorial

Lígia Guimarães Telles (UFBa)

Luiz Antônio Marcuschi (UFPe)

Maria da Conceição Paranhos (UFBa)

Regina Zilberman (PUC/RS)

Rosa Virgínia Mattos e Silva (UFBa)

Serafina Pondé (UFBa)

Assessoramento Editorial

Conceição Torres (UFBa)

Jacques Salah (UFBa)

Robélia Cabral (UFBa)

Projeto Gráfico / Editoração

Manoel Boullosa e Bete Capinan

UNIVERSIDADE FEDERAL DA BAHIA

Reitor
Luis Felipe Serpa

INSTITUTO DE LETRAS

Diretor
Evelina Hoisel

Vice-Diretor
Iracema Luiza de Souza

Programa de Apoio a Publicações Científicas

SCT/PR



CNPq



FINEP

INSTITUTO DE LETRAS
DA UNIVERSIDADE FEDERAL DA BAHIA
Campus de Ondina, CEP 40.170-290,
Salvador, Bahia, Brasil
Telefone: (071) 366-0790
Fax: (071) 336-8355
E-MAIL: PGLET@UFBA.BR

SUMÁRIO

Apresentação	7
Rosa Virgínia Mattos e Silva	
Creoles, markedness and the Language Bioprogram Hypthesis	11
Ian Roberts	
A gramaticalização	25
Ataliba Castilho	
A relevância dos processos de pidginização e crioulização na formação da língua portuguesa no Brasil	65
Alan Baxter e Dante Lucchesi	
From Latin weak pronouns to Romance clitics	85
Giampaolo Salvi	
Do português clássico ao português europeu moderno: uma análise minimalista	105
Charlotte Galves	
Mudança sintática: clíticos, negação e um pouquinho de <i>scrambling</i>	129
Ana Maria Martins	
Importância das fontes judiciais no conhecimento do português seiscentista	163
Rita Marquilhas	
O termo sintaxe	179
Maria Carlota Rosa	
A acentuação gráfica em João de Barros e os elementos nulos no SN	189
Sônia Cyrino	
Aspectos da sintaxe da ordem nas sentenças finitas na 'Obra Pedagógica de João de Barros'	197
Maria Aparecida Moraes	
A 'Obra Pedagógica de João de Barros': a sintaxe da ordem nas sentenças encaixadas	217
Ilza Ribeiro	
Adverbiais na 'Crônica de D. Pedro'	239
Sônia Bastos Borba Costa	
Observações sobre a variação no uso dos verbos <i>ser, estar, haver, ter</i> no galego-português ducentista	253
Rosa Virgínia Mattos e Silva	

Apresentação

Rosa Virgínia Mattos e Silva

UFBA

A **Revista Estudos Lingüísticos e Literários** do Programa de Pós-graduação em Letras e Lingüística e do Instituto de Letras da Universidade Federal da Bahia, na sua longa história — longa história, sim, em função dos parâmetros de longevidade de revistas acadêmicas da nossa Área: a nossa **Revista** existe sem descontinuidade desde de 1984 — define alguns de seus números como temáticos. Neste que agora se edita o foco é “Lingüística Histórica e história da língua portuguesa”. Tal tema foi proposto à **Revista** pelo “Programa para a história da língua portuguesa - PROHPOR”, grupo de pesquisa vinculado ao Departamento de Letras Vernáculas e ao PPGLL. Esse grupo de pesquisa tomou a si a sua organização.

Para tanto, o grupo PROHPOR convidou colegas estrangeiros e brasileiros, especialistas nesse campo da Lingüística, além de publicar aqui trabalhos de alguns dos membros do grupo. Os colegas convidados têm tido contacto de vária natureza com as atividades do PROHPOR, quer como professores visitantes, como Ian Roberts, Alan Baxter e Ataliba de Castilho; quer como responsáveis por seminários e conferências, como Charlotte Galves, Giampaolo Salvi, Ana Maria Martins; quer como companheiros, que lidam na mesma seara, como Rita Marquilhas, Maria Carlota Rosa, Maria Aparecida Morais e Sônia Cyrino, as duas últimas já associadas ao PROHPOR por projeto coletivo comum, ainda em curso, intitulado “Do português arcaico para o moderno”; mas todos como estimuladores e interlocutores de nosso trabalho.

É, portanto, com muito grata satisfação que apresento aos Leitores esses colegas de longe e de próximas terras e também os membros do PROHPOR que aqui publicam: Dante Lucchesi, Ilza Ribeiro, Sônia Borba Costa e eu própria.

Ian Roberts, da Universidade de Wales, Banghor, é dos mais conhecidos romanistas gerativistas, especialista em sintaxe comparada. Traz-nos em seu artigo a questão das línguas crioulas no quadro teórico da teoria chomskyana recente, em que os crioulos são interpretados, devido à sua gênese a partir de *pidgins*, como línguas tipologicamente caracterizadas por notável número de propriedades sintáticas decorrentes da seleção de parâmetros não-marcados, o que não ocorre com exclusividade nos crioulos, distinguindo-se esses, entretanto, das outras línguas tipologicamente semelhantes, por sua gênese histórica.

Ataliba Castilho da Universidade de São Paulo e da Universidade de Campinas, muito conhecido especialista no português brasileiro atual, coordenador do grande projeto interinstitucional “Gramática do português falado”, traz à discussão no seu artigo, com base em bibliografia diversificada, questões teóricas, ilustradas com fatos de várias línguas, referentes ao complexo problema da “gramaticalização”. Apresenta-se esse estudo como uma atualização de questões que há muito vêm perseguindo os lingüistas que se orientam no sentido de explicitar o constituir-se

das línguas no tempo, no contexto do processo comunicativo, orientação que marca as vertentes funcionalistas da Lingüística.

Alan Baxter, da Universidade da La Trobe, Austrália, é reconhecido especialista em línguas crioulas de base portuguesa e juntamente com Dante Lucchesi, membro do PROHPOR, estruturaram e vêm desenvolvendo o projeto "Vestígios de criouliização em comunidades afro-brasileiras isoladas". O artigo que apresentam, tal como o de Ian Roberts, tem como foco a questão das línguas crioulas. Diferentemente deste, que enquadra a instigante questão na teoria chomskyana, visando a tipologia sintática geral das línguas, A. Baxter e D. Lucchesi, também discutindo teorizações sobre línguas crioulas, centram-se no problema, ainda não resolvido pela sociolingüística, que é o da designada "Criouliização prévia do português popular brasileiro" e se concentram na análise de dados de uma comunidade afro-brasileira isolada, Helvécia, no sul do estado da Bahia, apresentando resultados parciais de pesquisa ainda em curso.

Do conjunto dos treze trabalhos que compõem este número, esses três iniciais são os mais definidamente teóricos, no âmbito da Lingüística histórico-diacrônica, e percorrem vertentes teóricas conviventes hoje nesse campo, quais sejam a da sintaxe diacrônica gerativa, a dos funcionalismos e a da sociolingüística histórica.

Os três artigos seguintes, todos no quadro teórico gerativista, discutem questões referentes a um problema crucial das sintaxes românicas, que é a dos chamados clíticos.

O artigo de Giampaolo Salvi, romanista ítalo-suíço da Universidade de Budapeste, Hungria, aborda o problema no latim e nos desenvolvimentos diacrônicos subseqüentes em direção aos clíticos românicos. Centra-se numa interpretação diacrônica dos clíticos nas línguas românicas antigas, descendentes dos pronomes pessoais fracos do latim. No quadro teórico gerativista atual, enfoca esse aspecto das gramáticas das línguas que tem sido intensamente debatido pelos sintaticistas gerativistas nesses últimos anos.

O estudo de Charlotte Galves, da Universidade de Campinas, lingüista sintaticista teórica e que vem ultimamente se dedicando a questões diacrônicas da história do português para além da sua vasta obra sobre a sintaxe brasileira, focaliza a análise da mudança ocorrida na colocação dos clíticos no português europeu do século XIX, no quadro teórico do minimalismo chomskyano.

Também nessa abordagem é o artigo de Ana Maria Martins, filóloga e sintaticista gerativista da Universidade Clássica de Lisboa. Centrando-se nos clíticos, discute a Autora outros problemas sintáticos a eles correlacionados. Parte da colocação dos clíticos no português europeu atual, em comparação com outras línguas românicas, e percorre a questão no português dos séculos XIII ao XVII, em confronto com o português europeu contemporâneo.

Rita Marquilhas, filóloga e lingüista, também da Universidade Clássica de Lisboa, associa, no seu artigo, o trabalho filológico de edição de textos à análise lingüística, articulando grafia, realizações fônicas e interpretação fonológica. Traz novos dados significativos para a cronologia de mudanças fônicas já em curso,

pelo que indica a documentação analisada, no português europeu do século XVII, quando a tradição filológico-lingüística estabelece para o século subsequente fatos que a Autora faz recuar no tempo histórico do português.

O artigo de Maria Carlota Rosa, filóloga e lingüista, da Universidade Federal do Rio de Janeiro, autora de uma significativa tese de doutoramento sobre o primeiro impressor em Portugal, Valentin Fernandes, aborda uma análise conceitual do termo "sintaxe" no percurso histórico do seu uso desde o século XV até à atualidade, articulando assim historiografia lingüística, um dos seus interesses de pesquisa atualmente, com um aspecto semântico-lexical na metalinguagem histórica dos estudos sintáticos.

O conjunto de três artigos com base no *corpus* da 'Obra pedagógica de João de Barros', datada de 1540, é parte do projeto coletivo, em andamento, do PROHPOR, intitulado 'Do português arcaico para o moderno'.

Sônia Cyrino, da Universidade de Londrina, Paraná, Doutora em Lingüística pela UNICAMP, avalia recurso ortográfico proposto e aplicado por João de Barros, primeiro gramático prescritivista da língua portuguesa, para analisar e interpretar aspectos sintáticos do português quinhentista, numa abordagem no quadro teórico da gramática gerativa.

Ainda nessa linha teórica, analisam aspectos da sintaxe da ordem em João de Barros Maria Aparecida Morais, da Universidade de São Paulo e Doutora em Lingüística pela Universidade de Campinas e Ilza Ribeiro, da Universidade Estadual de Feira de Santana, também Doutora em Lingüística pela UNICAMP, membro do grupo de pesquisa PROHPOR. A primeira focaliza aspectos referentes à ordem sintática em sentenças finitas, não-encaixadas, enquanto Ilza Ribeiro aborda questões relativas à ordem, mas em sentenças encaixadas, ambas discutindo o problema da sintaxe V2 no português quinhentista.

Essas três últimas Autoras prosseguem nesses artigos a sua linha de pesquisa em sintaxe diacrônica que já resultou em vários artigos, comunicações, além de suas teses de doutoramento defendidas recentemente na UNICAMP.

Os dois artigos finais se fundam em dados exaustivos sobre *corpora* do período arcaico do português, ou seja, a partir do século XIII e adentrando-se pelo século XVI, e se orientam numa linha de análise descritiva interpretativa.

O de Sônia Bastos Borba Costa analisa exaustivamente os elementos "adverbiais", sobretudo as locuções, na **Crônica de d. Pedro** de Fernão Lopes, primeiro prosador da língua portuguesa, cuja saborosa e rica escrita se situa na primeira metade do século XV. Dá, com esse estudo, a Autora continuidade a seu projeto individual de pesquisa no PROHPOR, que já resultou em publicações, com base em outros *corpora*, sobre esse aspecto da estrutura do português.

O último artigo, de minha autoria, também é parte de projeto individual do PROHPOR. Nele se enfoca a variação nos usos de *ser*, *estar*, *haver* e *ter*. Tem como base um *corpus* ducentista galego-português diversificado, que engloba produção poética, **Cantigas de Santa Maria** e o **Cancioneiro da Ajuda**, e produção em prosa não-literária — notarial, oficial e privada, e foral, com o **Foro Real de Afonso**

X. Busca-se nesse trabalho precisar os usos variáveis, que resultarão em mudanças significativas posteriores, dos referidos verbos no galego-português do século XIII, nas estruturas existenciais, atributivas e possessivas.

Parece-me caber, para concluir esta **Apresentação**, a pergunta: qual o sentido dos estudos aqui reunidos no contexto da Lingüística Histórica da atualidade e das pesquisas para a construção de uma história renovada e rigorosa da língua portuguesa?

Seguindo-se à hegemonia dos estudos sincrônicos que perdurou por quase todo o nosso século, os estudos histórico-diacrônicos voltaram à cena — não como vertente hegemônica, é claro, como no século XIX — pela força de orientações teórico-metodológicas que vêm avançando depois da década de sessenta, tais como: a teoria da variação e mudança, que tem a obra de W. Labov e seus seguidores como carro-chefe; as reformulações que apontam no gerativismo de fins de setenta e vieram a desembocar na teoria de princípios e parâmetros, que, a par da aquisição lingüística, passou a considerar a mudança lingüística como necessária à sua argumentação e estudos de sintaxe comparada como forma de fundamentar o modelo paramétrico proposto; também pela via funcionalista, que, pesquisando processos comunicativos, não pode restringir-se aos mecanismos apenas sincrônicos, já que nesses processos a diacronia se revela.

No Brasil, é nos inícios dos anos oitenta que os estudos histórico-diacrônicos começam a ressurgir, tanto pela via da sociolingüística como pela gerativista e, mais recentemente, no âmbito dos estudos funcionalistas.

O conjunto de trabalhos aqui publicados espelha essas abordagens atuais e mostram como estão pesquisando no complexo campo da Lingüística Histórica e, nele, na história da língua portuguesa, investigadores estrangeiros e brasileiros. A predominância de estudos de sintaxe histórica reflete, sem dúvida, um novo momento da Lingüística Histórica que, até meados deste século, deu prioridade à mudança fônica. E, no caso da língua portuguesa, era, como se sabe, um terreno quase virgem, o do conhecimento de mudanças sintáticas que ocorreram do século XIII, momento em que o português inicia o seu percurso de língua não-ágrafa, até os nossos dias.

As contribuições aqui reunidas, que perpassam pelos sete séculos de história da língua portuguesa, trazem aos Leitores trabalhos calçados em teorias e métodos atuais, sobretudo na análise de dados empíricos, fundamentais para a escrita da história de qualquer língua no seu percurso diacrônico.

Estamos cientes, pesquisadores brasileiros, portugueses e estrangeiros, que apenas começamos a reunir novos dados e novas interpretações, para a reconstrução da história da língua portuguesa, fundados nos avanços das teorias e dos métodos da Lingüística Histórica contemporânea.

Este número de **Revista Estudos Lingüísticos e Literários**, julgo, reveste-se como uma contribuição nessa direção.

Salvador, 2 de março de 1997

Rosa Virgínia Mattos e Silva

Creoles, Markedness and the Language Bioprogram Hypothesis¹

Ian Roberts

University of Wales, Bangor

Resumo

Este artigo trata de traços sintáticos partilhados por crioulos na perspectiva de uma versão da teoria de princípios e parâmetros que contém a noção de marcação. O argumento central é que os traços comuns aos crioulos que Bickerton (1981, 1984) apresentou como evidência para a Hipótese do Bioprograma da Linguagem podem ser vistos como reflexos do estabelecimento de parâmetros não-marcados (essa idéia está antecipada em Bickerton (1984)). A noção de marcação se liga à noção de simplicidade ou elegância de representações. As regras de movimento criam estruturas de adjunção (exclusivamente, cf. Kayne (1994)). Estruturas de adjunção são mais complexas do que estruturas de não-adjunção, em um sentido óbvio. Seguindo Chomsky (1995), parto do princípio de que todo parâmetro depende dos valores dos traços abstratos das cabeças funcionais: a variação paramétrica decorre do fato de uma dada cabeça funcional H poder desencadear movimento de vários tipos. Se H desencadeia movimento, cria-se uma representação relativamente complexa. O valor desencadeador de movimento de um parâmetro (o valor "forte" na terminologia de Chomsky) é então o valor marcado. Assim, dada a noção de expressão de parâmetro do tipo desenvolvido por Clark & Roberts (1993), combinada com a teoria da marcação, vê-se que só valores de parâmetros marcados precisam ser expressos; no caso da não-expressão de um parâmetro, o *default* será o valor não-marcado. Há evidências para o fato de que valores marcados de parâmetros possam ser ou sintaticamente expressos (pela ordem das palavras) ou morfológicamente expressos (por paradigmas flexionais associados a cabeças funcionais). A propriedade especial dos crioulos é que, dada à circunstância da gênese do crioulo (aquisição da língua tendo por base um *pidgin* como *in put*), os valores marcados dos parâmetros não eram expressos. Daí terem os crioulos uma preponderância de parâmetros com valores não-marcados. Neste artigo, argumenta-se que um número de notáveis propriedades sintáticas que são comuns a muitos crioulos é atribuível a parâmetros com valores não-marcados. Dessa conclusão decorre que não há nada de especial nos crioulos, além de sua história.

¹ This paper is part of a longer paper ("Verb-Movement and Markedness" in M. deGraff and A. Pierce, eds, *Creolization, Language Change and Language Acquisition*, MIT Press, 1996), which relates the claims made here about creoles to the loss of verb-movement in 16th-century English and to the "root-infinitive" stage of the acquisition of English, French and various other languages discussed in Wexler (1994), Rizzi (1994) and elsewhere. Parts of this material have been presented in lectures at the Universities of Paris, Florence, Venice, Oxford and Wales. I'd like to

Introduction

In a series of very interesting papers, Bickerton (1981, 1984) has argued that creoles give a direct insight into Universal Grammar. His central idea is that creoles are acquired on the basis of a radically impoverished trigger, so impoverished that UG (or the Language Bioprogram) has a more direct relationship with the final-state system than in the case of non-creole languages. Non-creole languages are, of course, underdetermined by experience, but the standard assumption is that aspects of experience profoundly influence the final state by setting parameters, etc. On the other hand, even such aspects of experience as these are argued to be absent in the trigger experience for acquisition of creoles and so creoles provide a more direct reflection of Universal Grammar than do non-creoles. Creoles are unique window on the Language Bioprogram.

Bickerton's evidence for his point of view comes from the striking morphosyntactic similarities that hold among creoles that are based on different lexifier languages and are widely dispersed both geographically and historically. I will review the most salient of these below. Bickerton argues that it is extremely unlikely that these similarities are the result of historical borrowing or contact, and most unlikely that they are due to chance.

The hypothesis that I would like to entertain is that creoles tend to have weak values of parameters. As such, they will have a number of properties in common (but notice that these common properties could perfectly well be shared by non-creole languages, something that the strongest version of the Language Bioprogram Hypothesis would not admit). We can relate the tendency towards weak parameter values to the nature of the trigger. Assume that a significant part of the trigger consists of pidgin. Such vernaculars lack inflectional morphology, and so the morphological triggers for strong feature values are missing. Similarly, we can speculate that pidgins do not robustly encode syntactic triggers, owing to their impoverished and artificial status. Notice that we do not need to make any very strong claim about the trigger for the acquisition of first-generation creoles, only that the trigger is morphologically and syntactically defective on crucial points, perhaps just in

thank the audiences at those presentations for their comments. I'd also like to thank Robin Clark, from whom the ideas about learning theory are derived, and the students in my Seminar on Language Change at the University of Maryland — especially Jairo Nunes, who suggested that I look at creoles. Thanks also to David Lightfoot and Juan Uriagereka for interesting discussions and much more. Finally, my grateful thanks to Derek Bickerton and Michel deGraff for very helpful comments on an earlier draft. All the errors are mine.

the sense that certain properties required for the triggering of strong features were not expressed with sufficient frequency or clarity. This strikes as a plausible and conservative speculation as to the circumstances of creole genesis. Granted this much, if the triggering data for strong values is not available to the learner the learner's preference for maximally elegant representations will always favour weak feature values since these give rise to representations that are simpler than those arising from strong feature values. Hence weak feature values will tend to predominate. So we can relate the unusual circumstances of creole acquisition to a propensity for unmarked, i.e. weak, feature values.

Some Properties of Creoles

We must now show that the syntactic characteristics of creoles can be accounted for as the reflexes of unmarked feature values. Muysken (1981, 1988) discusses six properties that he takes as characteristic of creoles. We will now discuss these one by one, and show how, with one partial exception, they can all be viewed as deriving from unmarked values of parameters.

1. Lack of verb movement.

This property is particularly interesting in creoles whose lexifier languages have verb movement. A clear case of this type is Haitian, whose lexifier language is French. Haitian very clearly lacks V-to-I movement, as has been shown in recent work by deGraff (1994) and deGraff & Dejean (1994). DeGraff (1994) explicitly applies the Pollock-Emonds tests for V-to-I movement (cf. Emonds (1978), Pollock (1989)) to Haitian, and the results very consistently show that V-to-I movement is lacking in Haitian. Contrasts with French of the following type are illustrative (although deGraff discusses a much wider range of data, including many different classes of adverbs):

Adverb-placement:

- (1) a. *Bouki pase deja rad yo. (Haitian)
 Bouki iron already cloth the
 b. Bouki deja pase rad yo.
 "Bouki has already ironed their clothes."
- (2) a. Bouqui repasse déjà le linge. (French)
 Bouqui irons already the cloth.
 "Bouqui is already ironing the clothes."
 b. *Bouqui déjà repasse le linge.

Negation:

- (3) a. Boukinèt pa renmen Bouki. (Haitian)
Boukinèt NEG love Bouki.
"Boukinèt does not love Bouki."
b. *Boukinèt renmen pa Bouki.
- (4) a. *Jean ne pas aime Marie. (French)
b. Jean n'aime pas Marie.
"John does not love Mary."

These examples clearly show that elements that are usually held to intervene between I and VP, and whose position relative to a verb is therefore a test for verb-movement, always precede V in Haitian. (Floated quantifiers are not found in Haitian (de Graff, p.c.); inversion, i.e. I-to-C movement, is not found in Haitian (DeGraff (1993:75))). The conclusion seems clear: Haitian lacks V-to-I movement. As both deGraff and deGraff & Dejean point out, Haitian is also a typical creole in that there is no verbal inflectional morphology. There is no subject-verb agreement at all, and information regarding tense, mood, etc. is carried by preverbal particles (to which we return below). Hence, the correlation between the lack of V-to-I movement and the lack of verbal inflectional morphology that has been observed in the history of various Germanic languages (see Roberts (1985, 1992), Rohrbacher (1994), Vikner (1995)) holds up in the relationship between French and Haitian.

2. SVO order.

It appears that creoles are without exception SVO (Bickerton (1981, 1988), Mühlhäusler (1986), Muysken (1981, 1988)). This is particularly striking in creoles that derive from OV lexifier languages. This is the case of Berbice Dutch, derived from Dutch and Ijo, both surface OV languages. SVO is of course a very common order among non-creoles, but SOV is just as common, and there is a significant minority of VSO languages (the other word-order types are very rare, cf. Greenberg (1963)). So creoles as a group can be distinguished from non-creoles in that they do not show non-SVO typologies; on the other hand, SVO itself is not confined to creoles. This is a good example of how creoles occupy just part of the space of variation that is attested in language in general.

It is plausible to regard VO as the unmarked order. Following Kayne (1994), Chomsky (1995) and Zwart (1994), I take VO to be the only available underlying order and surface OV to be derived by DP-movement to Spec,AgROP for Case-checking. So OV results from the presence of a strong

N-feature in AgrO. The strong feature of AgrO entails more complex representations for the learner, since it induces more overt movement. Hence it is marked. So we conclude that OV order is marked. In this light, the fact that creoles are always VO is a reflection of their tendency to show unmarked values for parameters, given the extremely degenerate trigger for creole genesis.

If we apply the above reasoning to subject movement, the straightforward prediction, assuming some version of the VP-internal subject hypothesis, would be that subject-raising out of VP is a marked option. This predicts that the "least marked" word order is SVO (since verb-movement, also a marked option, is required for VSO) where the subject is obligatorily adjacent to the verb (and the verb to the object). In such a system VP-adverbs and negation would always precede the subject, as would tense, mood and aspect particles and any I-type material. As examples like (1) and (3) show, this is not the situation in creoles; these elements typically intervene between the subject and the verb. We are led to the conclusion that creoles typically show the marked value for subject movement and the unmarked value for object movement. Why should this be?

We can show that abandoning the VP-internal subject hypothesis is not a solution. The proposal would be that subjects are base-generated in SpecAgrS, and that therefore there is no raising to this position, hence SXVO is not a marked order. However, there is clear evidence of raising constructions elsewhere in Haitian, hence evidence that AgrS has a strong N-feature independently of what we assume about the base position of subjects. DeGraff (1993) gives pairs like the following:

- (5) a. genlè Jak damou
 seem Jak in-love
 "It seems that Jak is in love."
 b. Jak genlè damou
 Jak seems in-love
 "Jak seems in love."

This clearly shows that raising to subject of a familiar kind exists in Haitian, and hence that AgrS has strong N-features.

We propose that the strength of AgrS's N-feature in Haitian reflects a property of UG, not a parametrically variant property or a property of the learning device. We can phrase the generalisation as follows:

- (6) AgrS has a strong N-feature.

- (6) corresponds to the Extended Projection Principle (EPP) of earlier

work (cf. Chomsky (1982)). In the minimalist framework, the EPP reduces to a feature-checking requirement, as (6) states. I assume that the strength of AgrS's N-feature does not vary parametrically. If so, then no markedness issue arises, and we of course expect creoles to have a strong N-feature in AgrS.² It remains an open question why (6) should hold. I will not speculate on this here.

3. No referential null subjects.

This property seems to be generally true of creoles, according to Muysken (1981), Lightfoot (1991). It is particularly striking in creoles that are based on null-subject lexifier languages like Spanish and Portuguese. A case in point is Papiamentu, based on Spanish. Muysken (1988:291) gives the following examples to show that this is not a null-subject language:

- (7) a. e ta kome
he ASP eat
(él está comiendo)
- b. *ta kome
(está comiendo)
- c. *ta kome maria
ASP eat Maria
(está comiendo Maria)

We can attribute this to the fact that creole agreement systems are too impoverished to permit recovery of the content of referential *pro*.³ Assuming, following Rizzi (1986a), that this ability to recover *pro*'s content is the fundamental property of null-subject languages, we can ask how this should be expressed in terms of the system of parametric variation adopted here. Given what we said above, it cannot be that the ability to license null subjects is connected to AgrS's N-feature, since (6) would then imply that all languages are null-subject languages. It must then be a property of AgrS's V-feature (recall that we assume that only these two features can vary). Accordingly, we adopt the following approach to licensing *pro*:

² This proposal implies that in VSO languages the subject raises to SpecAgrS and the verb higher. Such an analysis may be tenable, but to go into this matter would take us too far afield here.

³ We have nothing to say about the difficult question of null arguments in languages which entirely lack overt agreement, such as Chinese, Thai, etc.

- (8) a. Pro is formally licensed by a strong N-feature.
- b. Pro's content is identified by spec-head agreement with the relevant inflection.

(8b) requires that the morphology that permits the recovery of pro's features be present in AgrS. This is necessary at LF for the correct interpretation of pro, and it is necessary at PF for the correct identification (and perhaps elimination — cf. Chomsky (1995)) of pro's ϕ -features. Assuming that there is no PF-movement, this means that pro must be in the relevant configuration at Spell Out. Moreover, assuming Greed, the head bearing the independent inflection must be required to move into a position satisfying (8b) for pro for reasons independent of pro. The relevant morphology may be a grammaticalised pronoun, as in many Northern Italian dialects, or it may be the verbal morphology. In the latter case, (8b) requires that V be in AgrS at Spell Out, i.e. that AgrS have a strong V-feature. In this way, we derive the result that referential null-subject languages always have V-to-AgrS movement (although the converse does not necessarily hold: languages with V-to-AgrS movement do not necessarily allow referential null subjects — cf. French, Middle English, Icelandic, etc.). Thus (8b) gives us the result that referential null subjects are a marked property, contingent on V-movement. It then follows from what we said above about verb-movement that creoles will not have referential null subjects. Independent evidence for (8b) comes from the fact that referential null subjects cannot occupy the "freely-inverted" position in Italian (Rizzi (1987)).

(8) is a part of UG: it is a well-formedness condition on particular element, pro. As such it is not a parameter. The only syntactic variation with respect to null subjects concerns V-movement; the verb must be in a position where (8b) is satisfied in order to license subject pro at Spell Out. Hence, AgrS's V-feature plays a crucial role in the null-subject parameter. There is also morphological variation of an ill-understood kind: the verbal inflection must be "rich" enough to permit recovery of pro's features. We take it that inflection has whatever properties it has at PF, and that this may suffice or not to satisfy the licensing condition in (8b) at PF; there is no syntactic variation on this point beyond verb-movement. Our approach can solve at least one puzzle: (8a) prevents pro from appearing in Spec,CP (since this is not a Case position), and so (8b) predicts that, if V always moves to C in verb-second languages, referential null subjects cannot be licensed whatever the PF-properties of inflection. This explains why verb-second languages never allow referential null subjects, and in particular explains why Icelandic, a language which appears to have "enough" inflection to recover the content of pro, does not allow referential null subjects.

We conclude that Haitian does not allow referential null subjects.⁴ However, it is clear from examples like (9) that Haitian does allow expletive null subjects:

- (9) te fè frèt
ANT make cold
"It was cold." (deGraff's (2), p. 72)

Also, Haitian allows apparent violations of the COMP-trace filter, suggesting that it has expletive *pro*:

- (10) kimoun ou kwè (ki) pa vini
who 2sg believe (COMP) will come
"Who do you think will come?" (deGraff's (43), p. 80)

(8a), in combination with (6), allows any language to have expletive *pro* in SpecAgrS. Hence we expect to find it in Haitian. The question now becomes: why does English apparently not have expletive *pro* here? This may reduce to a simple matter of lexical variation: the inventory of pronouns in English contains expletive elements, that of Haitian does not (deGraff (1993, Note 3) points out that an overt expletive must appear with certain adjectives which have clausal complements, e.g. "difficult," "impossible," etc.; following Bennis (1986), these look like correlative argument pronouns). The inventory of pronouns in referential null-subject languages seems to consistently lack weak subject pronouns. This has usually been accounted for in terms of the Avoid Pronoun Principle, which is presumably some kind of PF Economy principle; I have nothing to add to such an account here.

The licensing principles for *pro* in (8), combined with the unexplained observation that expletive objects are not found, implies that expletive null subjects are only found in SpecAgrSP. They are thus banned from initial position in verb-second languages, an observation that has often been made

⁴ In a recent paper, deGraff (1993) argues that Haitian has null subjects. While he clearly shows that Haitian (and other creoles) can have null expletive subjects, deGraff's evidence that Haitian has referential null subjects is not convincing. His argument is that "'subject' pronouns in HA [Haitian — RC & IGR] do not appear in subject position, but that they are clitics phonologically spelling person and number features of INFL" (p. 73). However, there is nothing in deGraff's discussion which distinguishes the Haitian subject pronouns from those of Standard French. Standard French subject clitics have usually been regarded as phonological clitics (cf. Kayne (1983), Rizzi (1986b), Rizzi & Roberts (1989)) or, more recently, as weak pronouns (Cardinaletti & Starke (1993)). The data in (15 - 19) provide no reason to consider the Haitian pronouns as distinct from their Standard French counterparts. Hence they do not provide evidence that Haitian allows referential null subjects. The Haitian situation parallels that of Standard French and differs from what is found in typical Northern Italian dialects, favouring the conclusion that Haitian is like French and not like a Northern Italian dialect.

regarding the Germanic verb-second languages (cf., among others, Bennis (1986), Cardinaletti (1990), Vikner (1994)). This conclusion (as well as the one drawn above regarding Icelandic) assumes that the position of an initial subject in verb-second languages is not SpecAgrSP, pace Zwart (1993, 1994).

Expletive null subjects are also found in other creoles. DeGraff (1993:84) gives the following examples, among others, all of which have expletive null subjects:

- (11) a (a) (bi-) kendi/koto
 it TNS hot/cold
 "It was hot/cold." (Saramaccan; Byrne (1987:76))
- b. pro tawata jobe
 PAST rain
 "It was raining." (Papiamentu; Kouwenberg (1990:46))

This kind of data is consistent with the Haitian data presented by deGraff, and with our treatment of it.

We see then that referential null subjects, to the extent that they depend on V-movement to AgrS, represent a marked parametric option. On the other hand, expletive null subjects are a lexical, rather than a parametric, option (in fact, it is arguable that *pro* is universally available but its distribution varies depending on the other expletive pronouns available in the language; in English, for example, it is arguably available in locative-inversion constructions — cf. Hoekstra & Mulder (1990)). Hence, if creoles tend to favour unmarked parameter values, we expect that referential null subjects will not be found in these languages while expletive null subjects may be.

4. No complement clitics.

This property is particularly striking in those creoles which derive from Romance languages with rich systems of complement clitics. Haitian is again a good example, since it derives from French. DeGraff (1994: 11f.) gives the following contrasts with French:

- (12) a. Bouqui l'aime.
 Bouqui 3sg-like
 "Bouqui likes him/her/it."
 b. *Bouqui aime le/la.
 Bouqui like 3sg-m/3sg-f.
- (13) a. Bouki renmen li.
 Bouki like 3sg.
 b. *Bouki li renmen.

Haitian is quite typical in this respect.

There are various views on the nature of Romance clitics and cliticisation

processes. One view, due to Sportiche (1988) and elaborated by Rizzi (1993), is that Romance clitics undergo a combination of D- and DP-movement. In that case, some strong feature must trigger the movements, and hence clitics are marked. Another view, recently elaborated by Sportiche (1992), is that Romance clitics are themselves heads which trigger movement of pro. Although Sportiche claims that pro can move either before or after Spell Out, the system of pro-licensing in (8) implies that pro would have to move prior to Spell Out (although overtly-doubled elements could procrastinate, as Sportiche proposes). Hence in this system, too, clitics imply more movement than non-clitics. We do not need to choose between these two different views of Romance clitics here; the important thing is that both imply the existence of an overt movement operation for clitics which is not required where there are no clitics. So, on either view, complement clitics represent a marked parameter. We thus expect clitics of this kind to be lacking in creoles.

5. Preverbal tense/mood/aspect (TMA) particles.

This is another property that has often been noticed (cf. Bickerton (1981, 1984), Muysken (1981, 1988)). These particles carry information that is typically indicated inflectionally in languages with richer morphology, including many lexifier languages. The particles themselves are usually related to auxiliaries of the lexifier language, e.g. the Haitian anteriority marker *te* derives from either *été* or *était*, both past forms of French *être* ("to be"). Muysken (1981) shows that these elements have the following properties:

- they occur adjacent to the first verb;
- they cannot be fronted by predicate-clefting;
- they only occur with the first verb in serial-verb constructions.

Muysken concludes that these elements are auxiliaries. Although seemingly correct, Muysken's conclusion leaves important questions open. English auxiliaries, for example, show quite different syntactic and morphological properties internal to the class; it is generally assumed that modals and *do* are generated in I, while *have* and *be* are generated in syntactically lower V-positions. The main reason for this is that modals and *do* are always finite and always precede negation, while *have* and *be* are both able to be non-finite and follow negation. What they have in common is that they are athematic; these elements appear to lack argument structure. *Have* and *be* are also able to appear as "main verbs," where they may have an argument structure, although this is rather unclear (see Pollock (1989) and Williams (1994) for discussion of this point from different perspectives).

It is interesting to consider the Haitian TMA markers from the perspective

of English. In his discussion of these elements, deGraff (1993: 74-75) makes two important observations. First, negation always precedes all these markers:

- (14) a. Jan pa t ava ale nan mache
Jan NEG ANT MOOD go in market
"John would not have gone to the market."
b. Jan te (*pa) ava (*pa) ale (*pa) nan mache
Jan ANT NEG MOOD NEG go NEG in market

Second, most of the TMA markers are able to act as main verbs in isolation (deGraff, p. 75):

Pral, marking future, also means 'to go'; *dwe*, marking obligation or possibility, also means 'to owe'; *fini*, marking completion, also means 'to finish'; *konnen*, marking habituality, also means 'to know'; *sòti*, marking recent past, also means 'to leave'; etc.

It seems clear from the above that the TMA markers of Haitian at least are auxiliaries comparable to English *have* and *be*, rather than to English modals and *do* (deGraff also points out (Note 14) that Saramaccan and Antigua Creole are comparable to Haitian in this respect). However, an important difference as compared to *have* and *be* is that the TMA markers never raise, consistent with the general absence of V-movement in Haitian that we discussed above.

It has often been argued that auxiliaries like *have* and *be* should be treated as verbs heading their own VPs (this proposal goes back to Ross (1969), and has been adopted in recent work by Chomsky (1986), Pollock (1989), Roberts (1992) and many others). In that case, they are verbs which are at least able to be thematically defective, even if they may also have thematic structure when they appear as main verbs. However, if we maintain the proposal made in Chapter Two that functional heads are underspecified lexical heads, then these elements must strictly speaking be functional heads. If we adopt standard assumptions about clause structure, then the fact that these elements appear lower than negation indicates that they represent a second layer of functional structure below AgrS, T and AgrO. A proposal of this type has recently been made for complex auxiliary structures by Cinque (1996). Notice that, if these elements are functional heads, then they are functional heads with weak features, since they do not trigger overt incorporation of lower verbs. This is of course consistent with our general proposal about creoles. These elements are licensed as LF affixes (when they have no argument structure; when they have an argument structure they are licensed by assigning thematic roles, like all other verbs). Furthermore, at the very least tense features of some part of the verbal complex must be

checked with T. This clearly shows that T has a weak V-feature, another result consistent with our general claim about creoles.

We see, then, that the TMA system is a complex of functional heads with weak features. Above negation, there are at least three further functional heads, all of which have weak features (except for AgrS's invariant strong N-feature). All the features which are capable of variation thus have weak values in Haitian, and as far as we are aware Haitian is quite typical of creoles generally in these respects. This is consistent with our general view of creoles.

6. Conclusion

The above is a somewhat sketchy and second-hand overview of a number of salient syntactic properties of creoles. We saw that it is possible to analyse all of these properties as reflexes of weak feature values of functional heads, i.e. as unmarked parameter-settings. Hence creoles tend to have unmarked values for parameters. It is reasonable to speculate that this is due to the peculiar circumstances of creole genesis: if the trigger experience is made up largely of pidgin, then morphological triggers will be wholly lacking and syntactic triggers defective. The learner will then employ the elegance property in such a way that the simplest possible representations, i.e. those with the minimum amount of movement, will be selected. The result is that overt movement will be lacking, and creoles will have weak values for all or most parameters. In this precise and rather limited sense, Bickerton's claim that creoles can tell us something special was right. However, it is not the case, contra Bickerton, that creoles give us a unique insight into the workings of UG. Creoles, even if they truly have all parameters set to the weak value, are still just instantiations of what UG in combination with the learning device creates as possible variation. Moreover, languages other than creoles may have similar constellations of weak features: English for example has no verb movement to the higher functional domain, surface SVO order, no referential null subjects, no complement clitics and a system of preverbal TMA markers quite similar to those found in creoles. Moreover, all of these properties, except for the null-subject one, have been innovated in the recorded history of the language (and the null-subject parameter must have changed since Indo-European, if not since Proto-Germanic). What may give us a privileged view of UG, and the nature of the learner, is not so much creoles as language change.

Abstract

This paper looks at a number of syntactic features shared by creoles from the perspective of a version of principles-and-parameters theory which contains a notion of markedness. The central claim is that the common features of creoles which Bickerton (1981, 1984) adduced as evidence for the Language Bioprogram Hypothesis can be

seen as the reflections of unmarked parameter settings (this idea was anticipated in Bickerton (1984)). The notion of markedness relies on a notion of simplicity or elegance of representations. Movement rules create adjunction structures (exclusively, see Kayne (1994)). Adjunction structures are more complex than non-adjunction structures in an obvious sense. Following Chomsky (1995), I assume that all parameters are concerned with the abstract feature-values of functional heads: in particular, parametric variation concerns whether a given functional head H can trigger movement of various kinds. If H triggers movement, a relatively complex representation is created. The movement-triggering value of a parameter (the "strong" value in Chomsky's terminology) is thus the marked value. Now, given a notion of parameter-expression of the kind developed by Clark & Roberts (1993) combined with a theory of markedness, we see that only marked values of parameters need to be expressed; in the case of non-expression of a parameter, the default will be the unmarked value. There is evidence that marked values of parameters can be either syntactically expressed (by word order) or morphologically expressed (by inflectional paradigms associated with functional heads). The special property of creoles is that, given the circumstances of creole genesis (language acquisition on the basis of pidgin input), marked parameter values were not expressed. Hence creoles tend to have a preponderance of unmarked parameter values. In this paper I argue that a number of notable syntactic properties that are common to many creoles are attributable to unmarked parameter values. This conclusion entails that there is nothing special about creoles, other than their history.

References

- Bennis, H. (1986) **Gaps and Dummies**. Dordrecht: Foris.
- Bickerton, D. (1981) **Roots of Language**. Ann Arbor: Karoma.
- (1984) "The Language Bioprogram Hypothesis," **Behavioral and Brain Sciences** 7.2:173-222.
- Byrne, F. (1987) **Grammatical Relations in a Radical Creole**. Amsterdam: John Benjamins.
- Cardinaletti, A. (1990) **Pronomi nulli e pleonastici nelle lingue germaniche e romanze: Saggio di sintassi comparata**, Dottorato di ricerca in linguistica, Università di Padova.
- & M. Starke (1994) "The Typology of Structural Deficiency," ms. Universities of Geneva and Venice.
- Chomsky, N. (1982) **Some Concepts and Consequences of the Theory of Government and Binding**. Cambridge, MA: MIT Press.
- (1986) **Barriers**. Cambridge, MA: MIT Press.
- (1993) "A Minimalist Program for Linguistic Theory." In K. Hale & S.J. Keyser (eds) **The View from Building 20**, Cambridge, MA: MIT Press, pp. 1-52.
- (1994) "Bare Phrase Structure," **MIT Occasional Papers in Linguistics** 5.
- (1995) **The Minimalist Program**. Cambridge, MA: MIT Press.
- Cinque, G. (1996) "Adverbs and the Universal Hierarchy of Functional Projections," ms. University of Venice.
- Clark, R. & I. Roberts (1993) "A Computational Approach to Language :Learnability and Language Change," **Linguistic Inquiry**, 299-345.
- Emonds, J. (1978) "The Complex V - V' in French," **Linguistic Inquiry** 9: 151-175.
- deGraff, M. (1993) "Is Haitian Creole a Pro-Drop Language?" In F. Byrne & J. Holm (eds) **Atlantic Meets Pacific**. Amsterdam: John Benjamins, pp. 71-90.
- (1994) "To Move or Not to Move? Placement of Verbs and Object Pronouns in Haitian Creole and in French." Ms. University of Michigan.

- deGraff, M. & Y. Dejean (1994) "On Haitian Creole's 'very strict' Adjacency Principle." Paper given at the meeting of the society for Pignin and Creole Linguistics, Boston.
- Green, J. (1988) "Romance Creoles." In M. Harris & N. Vincent (eds) **The Romance Languages**, London: Routledge, pp. 420-474.
- Greenberg, J. (1963) "Some universals of grammar with particular reference to the order of meaningful elements." In J. Greenberg (ed) **Universals of Language**. Cambridge, MA: MIT Press.
- Hoekstra, T. & R. Mulder (1990) "Unergatives as Copular Verbs: Locational and Existential Predication," **The Linguistic Review** 7: 1-79.
- Kayne, R. (1983) "Chains, Categories External to S, and French Complex Inversion," **Natural Language and Linguistic Theory** 1:109-37.
- (1994) **The Antisymmetry of Syntax**, Cambridge, MA: MIT Press.
- Kouwenberg, S. (1990) "Complementizer *pa*, the finiteness of its complements and remarks on empty categories in Papiamentu." **Journal of Pidgin and Creole Languages** 5.1:39-51.
- Lightfoot, D. (1991) **How to Set Parameters: Degree-0 Learnability**. Cambridge, MA: MIT Press.
- Muysken, P. (1981) in P. Muysken & C. Lefebvre (eds) **Generative Studies in Creole Languages**. Dordrecht, Foris, pp.
- Muysken, P. & D. Bickerton (1988) "The Linguistic Status of Creole Languages: Two Perspectives." In F. Newmeyer (ed) **Linguistics: The Cambridge Survey, Vol. 2**. Cambridge, Cambridge University Press, pp. 267-306.
- Pollock, J.-Y. (1989) "Verb Movement, UG and the Structure of IP," **Linguistic Inquiry** 20:365-424.
- Rizzi, L. (1986a) "Null Objects in Italian and the Theory of *pro*," **Linguistic Inquiry** 17:501-557.
- (1986b) "On the Status of Subject Clitics in Romance." In O. Jaeggli & C. Silva-Corvalán (eds) **Studies in Romance Syntax**. Dordrecht: Foris, pp. 391-419.
- (1987) "Three Issues in Romance Dialectology," talk given at the GLOW Workshop on Dialectology. 10th GLOW Colloquium, University of Venice.
- (1993) "Some Notes on Romance Cliticisation," ms. University of Geneva.
- (1994) "Some Notes on Linguistic Theory and Language Development: The Case of Root Infinitives." Ms. University of Geneva.
- & I. Roberts (1989) "Complex Inversion in French," **Probus** 1:1-39.
- Roberts, I. (1985) "Agreement Parameters and the Development of English Modal Auxiliaries," **Natural Language and Linguistic Theory** 3: 21-58.
- Roberts, I. (1992) **Verbs and Diachronic Syntax**. Dordrecht: Kluwer.
- Rohrbacher, B. (1994) **The Germanic VO Languages and the Full Paradigm: A Theory of V to I Raising**, PhD Dissertation, University of Massachusetts, Amherst.
- Ross, J.R. (1969) "Auxiliaries as Main Verbs." In W. Todd (ed) **Studies in Philosophical Linguistics**, Series 1. Evanston: Great Expectations Press.
- Rottet, K. (1993) "Functional Categories and Verb Raising in Louisiana Creole," paper given at the Society of Pidgin and Creole Linguistics, Los Angeles.
- Sportiche, D. (1988) "Conditions on Silent Categories," Ms. UCLA.
- Sportiche, D. (1992) "Clitic Constructions," Ms. UCLA.
- Vikner, S. (1994) **Verb Movement and the Licensing of NP Positions in the Germanic Languages**. Oxford: Oxford University Press.
- Wexler, K. (1994) "Finiteness and Head Movement in Early Child Grammars." In D. Lightfoot & N. Hornstein (eds) **Verb Movement**. Cambridge: Cambridge University Press, pp. 305-350.
- Williams, E. (1994) "A Reinterpretation of Evidence for Verb Movement in French." In D. Lightfoot & N. Hornstein (eds) **Verb Movement**. Cambridge: Cambridge University Press, pp. 189-206.
- Zwart, J.-W. (1993) "Dutch is Head Initial," Ms. University of Groningen.
- Zwart, J.-W. (1994) **Dutch Syntax**. PhD Dissertation, University of Groningen.

A gramaticalização

Ataliba T. de Castilho

(USP - CNPq)

Resumo

Argumento neste texto que os estudos sobre gramaticalização não se situam claramente numa perspectiva teórica. Proponho distinguir *estágios* (sintaticização, morfologização, redução fonológica e alterações semânticas) de *princípios* de gramaticalização (analogia, reanálise, continuidade e unidirecionalidade), examinando fatos da língua portuguesa sob essa ótica. Concluo que a gramaticalização, juntamente com a discursivização e a semanticização, compõem um plano sistemático de estudo dos processos constitutivos da língua, no quadro de uma teoria modular.

Apresentação

Este trabalho tem por objetivo examinar criticamente a literatura atual sobre a gramaticalização, propondo no final um plano sistemático de investigações sobre esse e outros processos de criação das expressões lingüísticas. Ele foi preparado após participação em cursos ministrados por Joan Bybee e Paul Hopper no Instituto de Lingüística da "Linguistic Society of America", realizado na Universidade do Novo México, no verão de 1995, graças a financiamentos concedidos pela FAPESP e pelo Fundo Robert H. Robins. Agradeço a Miguel Salles pela identificação de fenômenos de gramaticalização em parte da literatura dialetológica sobre o Português do Brasil, e a Lygia C.D. de Moraes por observações críticas à versão anterior do texto. Os enganos continuam de minha responsabilidade.

A bibliografia consultada revela de parte dos autores certa despreocupação em situar-se claramente numa perspectiva teórica. Isto provavelmente se explique pela dificuldade em escolher uma única orientação, se o objetivo é dar conta da impressionante multiplicidade dos "casos de gramaticalização".

De fato, encontra-se aqui de tudo no que respeita a uma teoria sobre a gramática: desde uma visão estrutural (com seus sistemas hierarquicamente ordenados), passando por uma visão gerativista (a gramaticalização obedece a princípios e parâmetros), chegando a uma visão funcionalista, ainda pouco elaborada (a gramaticalização tem seu ponto de partida nas necessidades comunicativas do discurso). Paralelamente, encontra-se pouco no que respeita a uma teoria sobre a língua, sobre que se fundariam as teorias sobre a gramaticalização.

No corpo deste trabalho apresento um re-arranjo dos argumentos apresentados na bibliografia a que tive acesso, separando os *estágios* dos *princípios* mais gerais que os regem. Nas Conclusões, parto de uma concepção lingüística modular, e sugiro que a gramaticalização é apenas um dos processos constitutivos da língua. Dois outros processos, o de semanticização e o de discursivização, devem ser levados em conta, caso se deseje pôr ordem nas idéias e dispor de um plano sistemático de investigações sobre a criatividade lingüística. O objetivo maior deste texto é submeter ao debate essa proposta, que será aplicável tanto ao plano sincrônico quanto ao diacrônico, se é que nesta altura dos acontecimentos essa dicotomia ainda se mantém.

O texto se compõe de três partes: (1) A gramaticalização na Lingüística Contemporânea, (2) Estágios da gramaticalização, (3) Princípios da gramaticalização.

1. A Gramaticalização na Lingüística Contemporânea

Procedo aqui a uma resenha muito ligeira da vasta bibliografia sobre a gramaticalização.

Heine e Reh (1984), apud Hopper e Traugott (1993: 87), sintetizam com as seguintes palavras o entendimento que se vem tendo desse fenômeno: *"evolução em que unidades lingüísticas perdem em complexidade semântica, liberdade sintática e substância fonética"*. Por outras palavras, a gramaticalização é o estudo de mudanças lingüísticas situadas no continuum que se estabelece entre unidades independentes, localizadas em construções menos ligadas, e unidades dependentes tais como clíticos, partículas, auxiliares, construções aglutinativas e flexões: Traugott (1988).

A formulação de Heine e Reh sintetiza os quesitos que se mostraram recorrentes nos debates sobre gramaticalização, e que poderiam ser assim enumerados: (i) de que "unidades lingüísticas" se está falando: lexicais, gramaticais, semânticas, discursivas? (ii) de que "evolução" se está falando: evolução das categorias gramaticais, no plano diacrônico? emergência de categorias gramaticais e discursivas, no plano sincrônico? (iii) em que consiste a "complexidade semântica"?

Algumas respostas satisfatórias têm sido dadas, relativamente à perda de liberdade sintática e à redução fonológica por que passam as expressões que se gramaticalizam. No mais, tudo continua ainda muito obscuro.

Particularmente com respeito ao quesito (ii), surgiu desde logo uma distinção de natureza terminológica: trata-se dos termos *gramaticização* e *gramaticalização*, este de veiculação maior na literatura, desde que Meillet

(1912) o cunhou. Ambos demonstram que fenômenos diacrônicos e sincrônicos sempre cruzaram a argumentação dos lingüistas.

Por *gramaticização* entende-se o fenômeno da transição entre categorias, apanhadas em seu processo no plano sincrônico. Esse processo é particularmente perceptível quando examinamos a emergência da gramática na conversação.

Por *gramaticalização* entende-se o produto final que resulta da gramaticização, no plano diacrônico. Entretanto, por uma questão de comodidade e também por que "gramaticalização" se tornou um termo mais difundido, adotarei sistematicamente este último no presente texto.

Pelo menos três vertentes deram origem ao debate sobre a gramaticalização na Lingüística Contemporânea: a tipologia lingüística, o problema da mudança histórica e, mais recentemente, a observação do dinamismo da sintaxe da língua falada.

1.1 - Gramaticalização e tipologia lingüística

Humboldt (1822), apud Lehmann (1982b: 1), falando sobre a origem das formas gramaticais, hipotetizou que a significação das formas gramaticais sofre uma mudança em quatro estágios: (i) as categorias gramaticais estão completamente escondidas nos lexemas e nas configurações semântico-sintáticas de expressões idiomáticas; (ii) as palavras passam a ter uma ordem fixa, e algumas formas vacilam entre palavras plenas de conteúdo e palavras esvaziadas, com funções gramaticais; (iii) estágio da aglutinação: as "palavras vacilantes" são aglutinadas (como no Turco e no Finlandês), e sufixos "grudam-se" a palavras plenas; (iv) estágio da flexão: palavras gramaticais ligam-se a raízes, com alteração do esquema acentual, apagamento dos limites entre as palavras e criação de regras de ligação. Esses passos refletem os tipos morfológicos da tipologia lingüística de seu tempo: os estágios (i) e (ii) remetem às línguas isolantes, o estágio (iii) às línguas aglutinantes, e o estágio (iv) às línguas flexionais.

As idéias de Humboldt ficaram conhecidas como a Teoria da Aglutinação, embora este termo retrate apenas o estágio (iii). Elas tiveram uma enorme repercussão entre indo-europeístas, tais como Franz Bopp, Karl Brugmann e Georg von Gabelentz.

Não poderei aqui proceder a um rastreamento das noções sobre gramaticalização na vasta literatura sobre tipologia lingüística. Tornaram-se bem conhecidos os estudos tipológicos sobre línguas que situam o verbo na posição final da sentença e que, ao mesmo tempo, praticam posposições

de variada ordem. Lembrem-se também os estudos sobre as transformações dos Demonstrativos em Artigos e em Marcadores de gênero, as pesquisas em línguas africanas compendiadas em Heine-Claudi-Hünemeyer (1991), e as indagações sobre as relações entre gramaticalização do significado e gramaticalização da morfologia, com base num corpus de 66 línguas de diferentes famílias: Bybee-Perkins-Pagliuca (1994).

1.2 - Gramaticalização e Mudança Lingüística

Meillet (1912) deu continuidade à tradição indoeuropeísta, sobretudo em seu artigo *L'évolution des formes grammaticales*, a que se seguiu, em 1915, *Le renouvellement des conjonctions*. Juntamente com a analogia, ele considera a gramaticalização como os dois principais processos de mudança gramatical. Em seu estudo de 1915 ele distingue *criação*, que é a introdução de categorias gramaticais e/ou formais inexistentes previamente, de *renovação*, em que um termo como Lat. *nam* é substituído por *quare*, donde o Fr. *car*.

Meillet propõe inicialmente a existência de três classes de palavras, as *palavras principais*, as *palavras acessórias* e as *palavras gramaticais*, indicando que entre elas há uma transição gradual. A esse processo de transição ele chamou *gramaticalização*, entendida como a "*atribuição de um caráter gramatical a um termo outrora autônomo*": Meillet (1912: 131).

Essa transição diz respeito ao esvaimento tanto do sentido quanto da forma, de tal sorte que em dado momento uma palavra acessória se transforma numa palavra gramatical, podendo mesmo juntar-se a uma palavra principal para atribuir a esta um papel gramatical de que ela não dispunha previamente. Para exemplificar o processo, ele chama a atenção para os usos do verbo francês *être* como palavra principal (= verbo locativo), como palavra acessória (= verbo de ligação) e como palavra gramatical (= verbo auxiliar, que pode combinar-se com o próprio *être*, enquanto verbo principal). A exemplificação de Meillet leva a crer que por gramaticalização se pode entender tanto um processo diacrônico, isto é, a derivação de usos "acessórios" e "gramaticais" de um uso "principal", quanto um processo sincrônico, isto é, a convivência dos usos assim constituídos num mesmo recorte do tempo.

Meillet se vale da metáfora da espiral para captar a transição de um tipo de palavra para outro, o que insinua a idéia de que a gramaticalização é um processo inacabado, sem fim.

Alguns exemplos de Meillet levantam o problema das relações entre *gramaticalização* e *lexicalização*, entendida esta como sinônimo de Etimolo-

gia. Ele explica que Al. *heute* deriva do Alto Alemão antigo *hiu + tagu*, literalmente “neste dia”, semelhante ao Port. *hoje*, Arc. *ogano*, do Lat. *hoc + die*, *hoc + anno*, respectivamente. Esse processo de agregação de um Demonstrativo a um Nome é extremamente produtivo em muitas línguas, mas conduz, na verdade, à criação de um item lexical. Ora, casos como esses não têm interesse no quadro dos estudos de gramaticalização, tal como o próprio Meillet a definiu, por se tratar de lexicalizações.

É. Benveniste, discípulo de Meillet, voltaria ao assunto em 1968, repisando os argumentos e os exemplos de seu mestre, e lançando o termo *auxiliação* para indicar a passagem de um verbo pleno a verbo auxiliar: Benveniste (1968).

Kurylowicz definiu a gramaticalização como processo em que se verifica 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*: Kurylowicz (1965), apud Lehmann (1982). Ao gramaticalizar-se, o item lexical perde substância semântica (“bleaching”, “fading”) e fonológica. Quando um item lexical “migra” para a gramática, ele passa a comportar-se como morfema livre (no caso dos Auxiliares), ou como morfema preso (no caso dos afixos derivados de um morfema livre). Não deixa de ser curioso que Kurylowicz admita uma “graduação na gramaticalidade”, perspectiva em que as flexões seriam mais gramaticais que os morfemas-vocábulo, o que parece “tipologicamente incorreto”.

Roberts (1993) vê a gramaticalização como “*a mudança de uma categoria léxica para uma funcional, associada à perda do conteúdo lexical*”. No caso dos verbos, ele postula que esse processo passa pelos seguintes estágios: verbo pleno > construção predicativa > forma perifrástica > aglutinação. A diferença entre o verbo pleno e o perifrástico está em que apenas o primeiro atribui papéis temáticos.

1.3 - Gramaticalização, Sincronia e Sintaxe Conversacional

Os autores examinados até aqui focalizam o aspecto diacrônico da gramaticalização, uma posição que predominou até aos 70. Entretanto, se nossa leitura de Meillet não estiver errada, vê-se que sincronia e diacronia se entrecruzaram desde logo nas pesquisas sobre gramaticalização. Mattos e Silva (1995: 22) recua para os anos 50 o momento de superação da famosa dicotomia saussuriana.

Sapir (1921: 104-105) definiu quatro conceitos, ao estabelecer um continuum entre os diferentes tipos de conceitos lingüísticos como uma base

para a sua tipologia sincrônica: I) Conceitos fundamentais, concretos, *"tais como objetos, ações, qualidades"*. II) Conceitos de derivação, *"menos concretos que I, e mais que III, normalmente expressos pela afixação de elementos não-radicais aos elementos radicais, ou por uma modificação interna destes últimos"*. III) Conceitos Concretos de Relação, *"expressos por afixação (...), diferem fundamentalmente do tipo II porque indicam ou implicam relações que transcendem do vocábulo particular a que estão imediatamente ligados"*. IV) Conceitos Puros de Relação, *"destinam-se a relacionar entre si os elementos concretos da proposição, dando-lhe forma sintática definida"*.

É visível a gradiência que assinala essas classes, tanto do ponto de vista semântico (do concreto para o abstrato), quanto do ponto de vista morfológico (de palavras independentes para elementos afixais), uma constatação recorrente entre os que focalizaram o processo da alteração semântica, que acompanha a gramaticalização: v. item 2.2.

Durante o Estruturalismo, o debate sobre a gramaticalização conheceu certo refluxo, rompido por Hodge (1970) e Givón (1971, 1979 e 1983).

Hodge admite dois estágios: (i) sintaxe forte e morfologia fraca, que de um modo geral corresponde aos estágios I e II de Humboldt, e (ii) sintaxe fraca e morfologia forte, que corresponde aos estágios III e IV de Humboldt. Ele criou o slogan *"a sintaxe de ontem é a morfologia de hoje"*.

Essa afirmação encontrou eco em Givón (1971: 413), já nos quadros da Gramática Funcional, quando este autor declarou que *"a morfologia de hoje é a sintaxe de ontem"*. Essa formulação recebeu, posteriormente, o seguinte aditamento: *"a sintaxe de hoje é o discurso pragmático de ontem"*: Givón (1979: 208-209). Surgiu assim a escala *Discurso > Sintaxe > Morfologia > Morfofonêmica > Zero*. Nesse mesmo texto, Givón sustenta que no processo de gramaticalização o modo pragmático de comunicação dá lugar ao modo sintático; com isso, expressões lingüísticas com vinculação sintática fraca se transformam em expressões sintáticas fortemente ligadas.

Note-se que até aqui os autores vinham operando com dois módulos lingüísticos: o Léxico e a Gramática. O Funcionalismo givoniano agregou um novo módulo, o do Discurso, considerando-o prioritário em relação à Gramática. Essa idéia conheceu uma série de desdobramentos, sendo um deles o da "iconicidade sintática". A sintaxe "icônica" revela um grau de não-arbitrariedade das estruturas com relação ao que elas significam. Assim, abrimos uma sentença com uma construção de tópico porque o referente sobre que se declara algo deve preceder a declaração propriamente dita, isto é, o comentário.

Outro desdobramento é o da chamada *Gramática Emergente*, enten-

dimento da gramática como um conjunto de parcelas *"cujo estatuto vai sendo constantemente negociado na fala, não podendo em princípio ser separado das estratégias de construção do discurso"*: Hopper (1988: 118). Do ponto de vista do tempo, este modelo vê a língua como uma atividade no tempo real, cujas regularidades são provisórias e continuamente sujeitas à negociação, à renovação e ao abandono. Com isto, a rigor não existe gramática, existe gramaticalização, *"que é a cristalização das formas discursivas mais produtivas"*: Castilho (1994b).

Estas postulações tiveram importância para impulsionar uma Teoria do Discurso. Entretanto, como pretendo demonstrar nas Conclusões, há grandes riscos em estabelecer relações de precedência entre módulos linguísticos tais como o Discurso e a Gramática. Essa afirmação vai implicar em conseqüências indesejáveis no entendimento dos processos constitutivos da língua.

Finalmente, lembre-se o incipiente debate sobre a Sintaxe Conversacional, vale dizer, sobre a sintaxe do oral. Diferentes grupos de pesquisadores destacam que o fundamental na língua falada são os processos de constituição das estruturas, a que contrapõem como inadequado um estudo estático-descritivo do oral, mais centrado nos produtos do que no processamento propriamente dito. Num trabalho que apresentei ao X Congresso Internacional da ALFAL, comprovei a dimensão sincrônica da gramaticalização ao examinar casos de repetição na língua falada, em que uma Construção de Tópico se sintaticiza como Argumento Interno de um nome ou de um verbo, passando em seguida a Adjunto, migrando para a direita da sentença, como Sujeito, sobrevivendo sua elipse como estágio final do processo: Castilho (1993). Sobre as relações entre Conversação e Gramática, v. Castilho (1995a).

2. Estágios da Gramaticalização

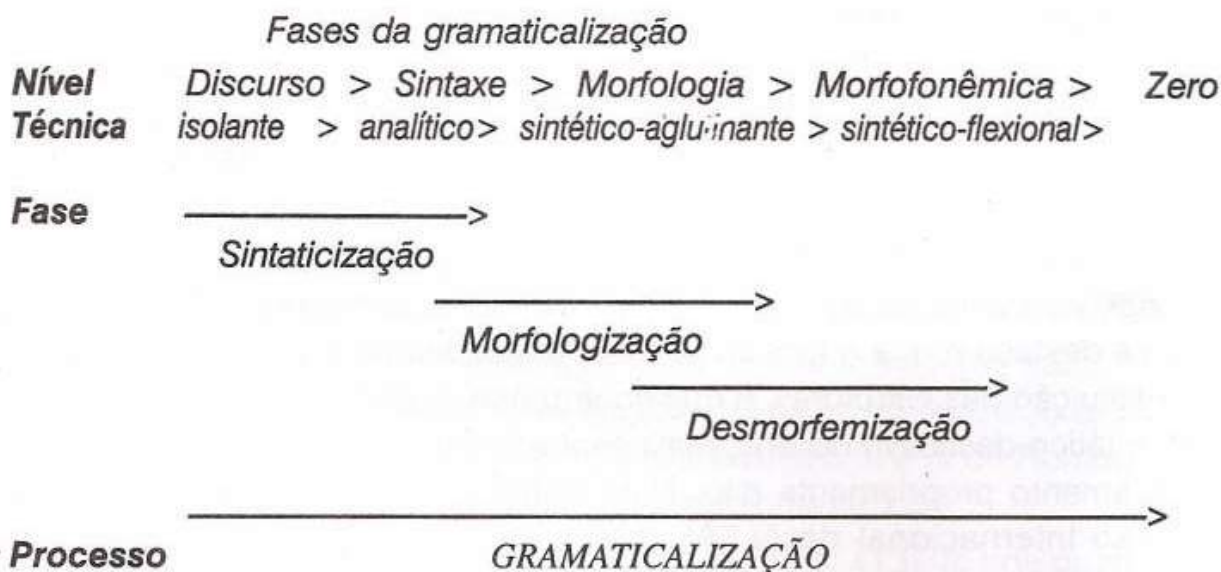
Após esse breve histórico das idéias sobre gramaticalização, vejamos resumidamente os resultados das pesquisas nesta área, separando-os por estágios.

Entendo por gramaticalização o trajeto empreendido por um item lexical, ao longo do qual ele muda de categoria sintática (= recategorização), recebe propriedades funcionais na sentença, sofre alterações morfológicas, fonológicas e semânticas, deixa de ser uma forma livre, estágio em que pode até mesmo desaparecer, como conseqüência de uma cristalização extrema. Esse trajeto se dá tanto no tempo real quanto no tempo aparente. Num sentido mais amplo, a gramaticalização é a codificação de categorias

cognitivas em formas lingüísticas, aí incluídas a percepção do mundo pelas *diferentes culturas, o processamento da informação, etc.*

Os seguintes estágios, ou fases, podem ser identificados no processo de gramaticalização (i) sintaticização, (ii) morfologização, (iii) redução fonológica, (iv) estágio zero, fase que reinstaura todo o processo. Simultaneamente, os itens lexicais sofrem (v) alterações semânticas.

A esse respeito, Lehmann (1982b: 13) organizou um quadro bastante ilustrativo. Vou aqui transcrevê-lo:



Desconsiderando o Discurso como ponto inicial no processo, passo a detalhar os estágios acima listados. Advirto que essa forma de disposição dos argumentos não é clara na literatura específica, correspondendo a uma leitura minha.

2.1 - Sintaticização

A sintaticização de um item lexical é (i) sua recategorização: uma "classe de palavra X" passa a "classe de palavra Y"; (ii) sua categorização funcional: categorias sintagmáticas recebem propriedades argumentais e de adjunção na sentença. Passo a considerar essas faces da sintaticização.

2.1.1- Recategorização sintática

Cada item lexical dispõe de propriedades gramaticais que os integram em diferentes classes léxico-sintáticas. Sua gramaticalização será, portanto, mais propriamente um processo de recategorização.

Na recategorização de categorias lexicais, tem-se observado o seguinte continuum:

Categoria maior [Nome, Verbo, Pronome] > Categoria mediana [Adjetivo, Advérbio] > Categoria menor [Preposição, Conjunção]: Hopper-Traugott (1993: 104).

A criação das categorias léxico-gramaticais tem sido tradicionalmente apresentada como processos de composição e de derivação. A recategorização de um item pré-existente, por fugir a esses processos, é descrita como um caso de "derivação imprópria". Os casos talvez mais dramáticos são o da transformação de um Verbo pleno em Verbo auxiliar ou em Complementizador, e o de Nomes em Preposições.

Este item está assim organizado: (1) Grupo Verbal, (2) Grupo Nominal, (3) Grupo Pronominal, (4) Grupo Adverbial e (5) Grupo dos Nexos.

2.1.1.1 - O Grupo Verbal

Como categoria primária, aparentemente o verbo não deriva de outra classe lexical, exceto nos casos em que um sufixo verbal é acrescentado a uma base nominal.

Assim, do ponto de vista da gramaticalização o fenômeno mais interessante é o da transformação de um verbo pleno num verbo funcional, e deste num verbo auxiliar.

Verbos plenos são os que funcionam como núcleos do predicado. Os verbos funcionais transferem esse papel para os constituintes à sua direita, tornando-se portadores de Pessoa, Número, Tempo e Modo. Os verbos auxiliares acompanham verbos nucleares na forma nominal, aos quais atribuem as categorias de pessoa e número, especializando-se como auxiliares de tempo, modo e aspecto.

No português, essa escala é exemplificada, entre outros, por *sedere* > *ser* "estar assentado", e *stare* > *estar* "estar de pé". Nos primeiros momentos da língua, esses verbos atuavam como núcleo predicador da sentença, selecionando sujeitos não agentivos, especificando o lugar ocupado por eles. Eram, portanto, Locativos.

Os Locativos dão surgimento aos Existenciais, a aos Possessivos, transitando de verbos plenos para verbos funcionais.

Tratando da gramaticalização de um Locativo como Existencial, Lyons (1968: 409-410) mostra que é natural raciocinar que se uma pessoa ou algo ocupa determinado lugar é porque existe. O lugar é cognitivamente falando mais concreto que a existência, que é mais abstrata. Um sinal dessa derivação pode ser encontrado naquelas línguas cujas estruturas Existenciais ainda mantêm um Adv. de lugar, seja como um "sujeito dummy" (Ingl. *there is*),

seja acompanhando o verbo (Fr. *il y a*, Ital. *ci sono*, Esp. *hay* (< *habet ibi*)) e Port. Arcaico e regional do Brasil *hai*: cf. *num hai quem num sabia, não hai home mais home do que outro não*: Amaral (1922: 77), Marroquim (1945: 213).

Os Locativos dão surgimento também aos Possessivos. A construção latina de esse com Dativo recipiente mostra essa gramaticalização: *mihi est liber*, “eu tenho um livro”, lit. “um livro é para mim”. V. também Ingl. *The children are with Mary* —> *Mary has the children with her*.

Verbos tipicamente Possessivos, como Lat. *habere*, *tenere* (e formas românicas derivadas), Germânico **kap* “agarrar, tomar” > **haf* > Ingl. *have* tinham o sentido original de “agarrar com as mãos”, “ter nas mãos”. De *tenere* derivam Port. *tenente*, Esp. *tierrateniente*, “proprietário”. Este fato teve conseqüência na organização do Aspecto Inceptivo, em que verbos como *pegar*, *agarrar*, associados a um infinitivo, passam a indicar o começo do estado de coisas codificado por este último: cf. Port. *pegou a falar* “começou a falar”, *garrou a criar uma coisa assim*, *garrou a atacar*, identificados no falar piracicabano por Rodrigues (1974), à semelhança do Esp. *se agarró hablando*.

Tratando de *ser*, *estar*, *ter*, *haver* no plano diacrônico, Mattos e Silva (1994) assim resume a questão: “1. a variação de *ser* e *estar* e o avanço deste em estruturas atributivas, 2. a variação de *haver* e *ter* e o avanço deste em estruturas possessivas, 3. a variação singular/plural e a perda de concordância no particípio passado em estruturas com *haver/ter* analisáveis como tempo composto (...)”. Há, também, confluências entre *ter* e *estar* no Port. coloquial, como em *eu tive lá*, por *eu estive lá*, *eu tive estudando na sua casa*. Como nos casos anteriores, essas ocorrências mostram que os verbos mais gramaticalizados integram uma classe especial, em que muitas permutas de formas são feitas.

O Copulativo é um item que transforma uma expressão nominal em predicado sentencial. Nas línguas Indoeuropéias, a cópula foi gramaticalizada como um verbo, como o Port. *ser*. No Chinês, é um Demonstrativo ou um Pronome Anafórico que assume essa função. Li and Thompson (1977), apud Lehmann (1982: 28), mostram que nessa língua a sentença nominal não tinha cópula, e seu sujeito era deslocado para a esquerda, sendo retomado no interior da estrutura por um Demonstrativo, gerando a estrutura [SN [Dem SN]]. O Demonstrativo no Chinês Arcaico é *shi*. Essa estrutura se gramaticalizou, sendo reanalisada como [SN Dem SN], e assim o Dem. *shi* passou a funcionar como cópula, deixando de ser um Demonstrativo.

Outro canal de gramaticalização é aquele em que uma construção nominal já preenchida por uma cópula é reforçada por um Dem., como acontece no Fr. *Vivre est apprendre à mourir*, progressivamente substituído por *Vivre c'est apprendre à mourir*.

Um verbo funcional pode transformar-se em verbo auxiliar, e este em afixos, morfologizando-se. Tais processos seqüenciais de gramaticalização podem ser assim representados:

Verbo pleno > Verbo funcional > Verbo Auxiliar > Clítico > Afixo.

Os verbos auxiliares codificam o Aspecto, o Tempo, o Modo e a Voz.

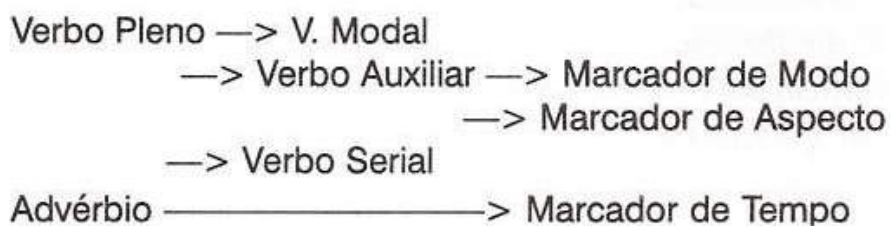
Na história das línguas, muitos *Auxiliares de Tempo* se morfologizam e são substituídos por outros. O processo parece ser desencadeado por uma sorte de competição entre essas formas. Assim, o Futuro Latino *amabo*, ele mesmo fruto da afixação de um antigo verbo IDE. **bho* "vou", passou a sofrer a concorrência de *amare habeo*, uma perífrase que de início codificava o Modo, com a idéia de obrigatoriedade. É de supor-se que durante algum tempo a forma simples, que poderíamos representar por A, sofreu a competição da perífrase, que poderíamos representar por B, até que esta última acabou por suplantar a primeira, num ritmo que poderia ser assim notado: A > A/B > B.

A continuidade da inovação, que funciona como um dos princípios da gramaticalização, leva à criação de uma nova forma flexional, como Port. *amarei*, Fr. *aimerai*, It. *amerò*, etc. (de *amare habeo*) a qual, a seu turno, passa a sofrer a concorrência de uma nova perífrase, desta vez constituída pelos verbos *ir* (Port. *vou / irei amar*, Fr. *je vais aimer*, It. *vado amare*) e *querer* (Fr. *je veux aimer*), retomando-se o processo anterior. No caso de *ir*, note-se que ao gramaticalizar-se esse verbo perdeu o sentido original de movimento de afastamento referido ao falante, por contraste com *vir*, que é movimentar-se em direção ao falante. Com isto, tornaram-se possíveis construções como Port. *vou ir, vai vir*, Ingl. *he is going to come*, que seriam contraditórias, caso o sentido original se tivesse mantido. No caso de *querer*, línguas como o francês acabam promovendo confusões entre *aller* "ir" e *vouloir* "querer" (cf. *je veux* "eu quero" e *je vais* "eu vou"), como em *le train veut partir*, no lugar de *le train va partir*. Sobre a associação entre a noção de movimento e a gramaticalização do futuro, v. Bybee-Pagliuca-Perkins (1991). Para mais evidências, v. Bybee-Perkins-Pagliuca (1994: 243-280).

Também os *Auxiliares de Modo* derivam de verbos plenos. Lat. existencial esse combinando-se com um gerúndio passa a expressar a obrigatoriedade: *mihi est eundum*, lit. "ir existe para mim", donde "eu tenho de ir". Verbos de volição dão lugar aos modais derivativos, como Ingl. *will*. Para uma elaboração sobre a gramaticalização dos modais, v. Abraham (1991) e Bybee-Perkins-Pagliuca (1994: 176-242).

No Latim Vulgar, e portanto nas Línguas Românicas, o *Auxiliar de Voz Passiva* esse substituiu a Voz Passiva afixal. Verbos dêiticos como *ir* e *vir*, por sua vez, substituem esse, como no It. *viene detto, va detto*.

Como se sabe, é larga a discussão sobre o estatuto categorial dos verbos auxiliares. Lehmann (1982: 38) considera infrutífera essa discussão: *two grammatical categories connected on a grammaticalization scale are neither the same nor distinct. The difference between them is gradual, and there is no clear-cut dividing line*. Ele assim sintetiza a criação das categorias de Aspecto, Tempo e Modo:



2.1.1.2 - O Grupo Nominal

Os seguintes processos de recategorização ocorrem no Grupo Nominal:

1) Verbos dão origem a Nomes substantivos, como os infinitivos (*comer é bom*), os participios (em expressões nas quais se apagou o nome, como "*collecta pecunia > coleta*", *promessa*, etc.), gerundivos no plural neutro (*fazenda, legenda > lenda*).

2) Adjetivos se recategorizam como Substantivos (*galgo < cannis galicus, o pobre*), Substantivos se recategorizam como Adjetivos (*vermelho < vermiculus, louro < laurus*): Maurer (1959: 248-250).

Os exemplos acima mostram que a recategorização dos Nomes pode ser assim formulada: *Verbo > Nome [Substantivo > Adjetivo, e Adjetivo > Substantivo]*.

2.1.1.3 - O Grupo Pronominal

Aparentemente, as línguas naturais organizam seu quadro pronominal privilegiando a codificação das pessoas do discurso e dos lugares ocupados por elas no espaço físico. Os pronomes são, portanto, fundamentalmente dêiticos.

Com a gramaticalização dessas formas, elas vão perdendo seu cará-

ter dêitico, isto é mostrativo, abstratizando-se esse valor, que se desloca para o da identificação do lugar no discurso, ou anáfora.

1) Os Demonstrativos são basicamente pronomes definidos dêiticos, servindo para especificar o lugar ocupado por pessoas ou objetos. Eles fazem as vezes dos gestos, como em *este aqui, aquele lá*, etc., e ocupam no SN o lugar dos Determinantes.

No caminho de sua gramaticalização, a primeira coisa que acontece é a perda do caráter dêitico e a assunção do valor anafórico, como aconteceu com Lat. *ille*, Port. *aquele*, e Pron. Pessoal *ele*. Depois, os Demonstrativos. podem transformar-se em Artigos, cliticizando-se ao N, como no Port. *o homem*, < Lat. Vulgar *illu homine*, Gr. *hó ánthropos* "o homem", e reduzindo seu caráter anafórico. O Artigo, com efeito, fornece uma instrução muito genérica para a busca do referente do Nome que ele determina.

Em suma, se a hipótese acima for verdadeira, as línguas começam a organizar seu quadro pronominal a partir dos Demonstrativos.

2) Os Pronomes Pessoais da primeira e da segunda pessoa são evidentemente dêiticos, e os da terceira são anafóricos. Segundo Apolônio Díscolo, a terceira pessoa só é dêitica na ostensão, como em *eu vou, mas ele aí não vai*. Isto cinde o quadro em dois tipos distintos de pronomes, com reflexos na sintaxe e em sua interpretação semântica.

Novas formas são obtidas, especialmente para a primeira e a segunda pessoa, pela incorporação de expressões de tratamento, como Port. *Vossa Mercê* > *você*, que deslocou *tu*, ou de sintagmas nominais indefinidos, como Port. *a gente*, que desloca presentemente *eu* e *nós*, Fr. *on*, do Lat. *homo*.

Com isto, duas são as fontes de gramaticalização dos Pronomes Pessoais: (i) Demonstrativos para a terceira pessoa determinada, e (ii) expressões nominais para a terceira pessoa indeterminada (cf. Lat. *homo* > Port. Arc. *ome*, Fr. *on*), para a primeira pessoa (cf. Port. *a gente* em lugar de *eu / nós*) ou para a segunda pessoa (cf. *Vossa Mercê* > *Você*): Lehmann (1982: 41).

Pronomes Pessoais átonos se cliticizam, como Port. *me, te, se, o, lhe*. O reflexivo *se* tende a generalizar-se para todas as pessoas, sobretudo nas modalidades não padrão das línguas, como Port. pop. *eu se alembro, você se alembra, ele se alembra*. Depois deste estágio em que se perdeu a categoria de pessoa, o reflexivo se cliticiza ao verbo, como no Fr. não-padrão *je sarrête, nous se reverrons, vous se privez*. Exemplos como esses fazem pensar na perda da representação morfológica da categoria de Pessoa.

2.1.1.4 - O Grupo Adverbial

As seguintes classes podem recategorizar-se, gramaticalizando-se como Advérbios:

1) Nome > Advérbio, como no Fr. *chez*, que deriva do Lat. *casa*; Port. *passo*, Fr. *pas* < Lat. *passu*, originalmente após um Adv. de negação com verbos de movimento (como em Port. “*não andou passo*”, literalmente, “*não deu um passo*”, Fr. “*Ils n’ont parti pas*”), e depois com outros verbos (Port. “*nom vejo passo*”, Fr. “*Je ne veux pas*”), e finalmente sem a primeira forma negativa, assumindo *passo / pas*, integralmente, a noção de negação (Port. “*olhou passo*”, Fr. “*connais pas*”). Nesta fase, completa-se a gramaticalização do N *passo*, que não pode mais ser articulado, como no Port. * “*olhou um passo*”, Fr. * “*ils ne fument un pas*”: Hopper (1991: 31). Outro fenômeno que pode aqui ser capitulado é a progressiva abstratização do N *tipo* no português falado. De “*isto aqui é um tipo de camisa*”, em que *tipo* é o núcleo do SN [*um tipo de camisa*], temos inicialmente a reanálise dessa unidade, em que *tipo* cede o núcleo a *camisa*, recategorizando-se como Especificador desse SN, constituindo a expressão adverbial Delimitadora. Num segundo momento, *tipo* se localiza fora do SN, ampliando seu uso, como em “*bom... tipo... que tal comer uma pizza ?*”.

2) Verbo > Advérbio, como no Fr. *voire*, que aparece nos Adv. de designação *voici, voilà*.

2.1.1.5 - O Grupo dos Nexos

1) Muitas Preposições são recategorizações de outras classes, tais como:

1.1) Nome > Preposição. Partes do corpo humano são freqüentemente reanalisadas como Preposições secundárias (ou “Nomes funcionais”, como também são denominados). Ingl. *back* “costas” pode funcionar para indicar o espaço (como em *three miles back* “três milhas atrás”) ou para indicar o tempo (como em *three years back* “três anos atrás”). A transformação desse Nome em Preposição implica também na reanálise do sintagma em que ele aparece. Assim, em *the back is strong*, *back* é o núcleo do SN. Já em *back of the body*, “atrás do corpo”, *back* é o Adjunto do SP, nos termos da gramática dessa língua. O mesmo se aplica a *face* (cf. *in the face of*), e ao Port. *amor*, em *por amor de* “por causa de”, e suas variantes *prumode, pramode*, mencionadas por Amaral (1922) e Nascentes (1953). Note-se que esses Nomes não podem ser articulados, nem receber do Verbo sentencial as propriedades sintáticas habituais, mas conservam de certa forma seu sentido original. O seguinte esquema capta a gramaticalização de um Nome:

N > N relacional > Prep. secundária > Prep. primária > Clítico > Afixo

Obviamente a fase clítica e de transformação em Afixo ocorrerá com as Preps. primárias. Nestas Preps., aliás, as regramaticalizações ocorrem com frequência, por meio de reforços pela mesma ou por outras Preps., aparentemente porque elas vão se tornando opacas quanto à representação do espaço. Em Port., conhece-se o caso de *desde* < *de + ex + de*. No Francês há uma incrível série de sobreposições, criadas a partir do Lat. *in*, logo reforçado por *tus*, donde *intus*, e em seguida por *de*, donde *de intus* (> *dans*), *de de intus* (> *dedans*), *de de de intus* (> *de dedans*). Fenômeno semelhante ocorreu com a Prep. *com*, cf. Port. Arc. *migo* (< Lat. *mecum*) > Port. Mod. *comigo*: Lehmann (1982: 18).

1.2) Verbo > Preposição. Verbos em formas nominais podem recategorizar-se como Preposições, como Ing. *regarding*, *considering*, habitualmente utilizadas para a introdução de novos tópicos no discurso. Port. *exceto*, *salvo*, *durante*, *mediante* são casos semelhantes. Mithun (1984) denomina *incorporação* esses casos.

1.3) Numeral Ordinal > Preposição: cf. Port. *segundo*.

(2) As Conjunções podem também proceder da recategorização de várias classes.

Pesquisas sobre línguas não indo-européias e sobre a modalidade falada têm evidenciado a precariedade da linha que separa a coordenação da subordinação. No caso do Português, basta lembrar as tênues relações entre coordenadas explicativas e subordinadas causais, entre coordenadas adversativas e subordinadas concessivas, etc. Agregue-se a isso o estatuto das correlatas como um terceiro tipo de relações intersentenciais.

A gramaticalização das relações intersentenciais aparentemente dispõe as coordenadas e as subordinadas nos extremos de um continuum, mediado pelas correlatas: Oiticica (1955). Se duas sentenças independentes, de verbo finito, compartilham o mesmo sujeito, elas podem ligar-se, ou apresentando uma delas seu verbo em forma não finita, ou antecedendo-se de conjunção. A primeira assumirá então uma hierarquia de superordenação. Assim, de "*X aprende*" e "*X estuda*", obtém-se "*X não só estuda como também aprende*", de que decorrem "*X aprende estudando / porque estuda / quando estuda / se estuda*". Qualquer uma dessas estruturas está à disposição do falante, desde a menos ligada até a mais ligada.

A conhecida assimetria entre dependência semântica e dependência estrutural, de um lado, e as motivações discursivas das relações de causa, condição, tempo, modo e lugar "[*que*] permeiam todo o texto, independentemente do nível das unidades estruturais (micro ou macro-estruturais) envol-

vidas: *sintagmas, orações, frases, parágrafos, capítulos, etc.*”, de outro, têm sido alvo de intensas investigações: Thompson-Mulac (1991), Hopper-Traugott (1993:171-203), Neves (1996).

Outra área de pesquisa importante aqui é o da gramaticalização das conjunções. Sweetser (1990: 78) mostra que as conjunções são polifuncionais, atuando no domínio do conteúdo, no domínio epistêmico e no domínio conversacional. Vejamos mais de perto como se constitui a classe das conjunções.

1) Nome > Conjunção. A palavra *modo* constitui a locução conjuncional causal *a modos que*, cujas variantes *amoque, mode, pramor de* foram registradas por Amaral (1922) no falar caipira: “*hei d’i na vila dumingo pramor de vê se compro os perciso*”. Outras variantes assumem o papel de Complementizador, como em “*Mas home, ocê mode que não imagina*”, com o sentido de “parece”: Marroquim (1945).

2) Verbo > Conjunção. Em várias línguas da África, *uerba dicendi et sentiendi* transformam-se em Conjunções. Hopper-Traugott (1993: 14) mostram que mesmo no Inglês isso é possível, quando uma conjunção condicional pode ser substituída pelo verbo *to say*, como em “*Say / If the deals falls through, what alternative do you have*”? O fenômeno ocorre também em algumas línguas românicas, como Fr. “*Mettons qu’on puisse y aller*”, Esp. “*Pongamos que yo lo sepa*”, Port. “*Digamos que eu possa*”. Em todos esses casos, segundo Lygia Corrêa Dias de Moraes (com. pessoal), o V em processo de transformação parece funcionar como um índice de modo irreal.

3) Advérbio > Conjunção. Um caso notável é o da transformação do Advérbio latino *magis* “mais” na conjunção coordenativa *mas*. Esse fenômeno, examinado por Ducrot-Vogt (1978), pode ser claramente documentado em segmentos da língua falada, como demonstraram Andrade (1995: 124) e Castilho (1995b). Encontramos aí as seguintes fases no continuum Adv. de Inclusão > Conjunção Adversativa: (1) *Mas* preserva o valor original: “*a gente vive de motorista o dia inteiro, mas o dia inteiro*” [D2 POA 360: 94]; “*nós temos tantos amigos desintegrados...mas nós só temos amigos assim de família desestruturada*” [D2 RJ 147: 167]. Nesses dois exemplos, *mas* soma sintagmas nominais e sentenças, notando-se que o segundo termo não tem relação de contrajunção com o primeiro, tanto assim que não é parafraseável por uma concessiva: “** temos amigos desintegrados... embora nós só tenhamos amigos assim...*”. A literatura dialetal registra com abundância o *mais / mas* inclusivo: “*stava a Rosita, mas seu Roberto, sem cerimônia*”: Laytano (1981); “*passarim avôe mais baixo, quando ocê cantá mais eu*”, “*minha filha está pra casar mais o filho do Manuel cargueiro*”, “*e fosse bolir mais ele, fosse!*”: Marroquim (1945). (2) Aparentemente, contextos de negação explí-

cita dão início à recategorização do Adv., como em *"ela está lá mas não funciona"* [D2 SP 343: 101], isto é, *"ela está lá embora não funcione"*. (3) Por metonímia, *mas* incorporou o sentido de negação, recategorizando-se plenamente como Adversativa, dispensando contextos com a negação, como em *"as mais velhas estão entrando na adolescência mas são muito acomodadas"* [D2SP 360: 41]. Em (2) e (3), a conjunção altera as expectativas geradas na sentença anterior.

4) Pronome > Conjunção. Dialetoólogos como Amaral (1922), Nascen-tes (1953) e lingüistas como Câmara Jr. (1972) e Tarallo (1983) vêm mostrando que os pronomes relativos estão perdendo suas propriedades pronomi-nais, com uma severa redução em seu quadro, restringindo sua atuação gramatical à de uma conjunção sem papel funcional. As relativas com pro-nomes-lembrete e as relativas cortadoras decorrem dessa mudança: *"este é o homem que eu falei com ele ontem"*, *"eu tenho um amigo que ele com-prou um terreno na praia"* / *"este é o homem que eu falei ontem"*, *"a casa que eu morei"*.

5) No item anterior mencionei indiretamente a redução dos pronomes relativos a apenas um item, *que*. Esse mesmo item está passando por outras formas de generalização de uso, operando como conjunção aditiva / tempo-ral / condicional / comparativa, como nestes exemplos de Marroquim 1945: *"eu tomei pula vareda, qui (= e) quando caí no engano tinha andado meia légua"* / *"eu não sei qui ele veio" (= quando)*, *"se o olhar fosse alfinete e que (= se) desse alfinetada, tu ficava furadinha, que (= que nem) só renda de almofada"*.

2.1.2 - Categorização Funcional

O estudo da atribuição de propriedades funcionais acentua o caráter dinâmico da gramaticalização. Esse programa de pesquisas tem-se revela-do menos desenvolvido que o anterior. Nos itens a seguir menciono alguns estudos sobre a emergência das categorias funcionais da sentença. Tratei desse fenômeno na língua falada em Castilho (1994a), de um ângulo sincrônico.

Situando-se em outra perspectiva, Tarallo (1991: 70) oferece uma ele-gante ligação entre o problema da recategorização léxico-sintática e o da categorização funcional. Ao estudar a emergência de uma "gramática brasi-leira", ele hipotetiza que *"a re-organização do sistema pronominal (...) teve como conseqüências importantes a implementação de objetos nulos no sis-tema brasileiro de um lado, e sujeitos lexicais mais freqüentes de outro"*, fato que repercutiu, também, nas estratégias de relativização.

Neste item, considerarei superficialmente a gramaticalização do Tópico, o preenchimento das funções de Sujeito e de Objeto Direto no Português do Brasil, e a ordem dos constituintes funcionais.

2.1.2.1 - Tópico > Sujeito

Línguas como o Japonês marcam estruturalmente o tópico (mediante a partícula *wa*) e o sujeito (mediante a partícula *ga*). *Wa* separa uma entidade dentre outras possíveis a predicar (donde o efeito de ênfase), separando-a, também, “*daquilo que constitui a proposição inteira*”: Shibatani (1991: 97).

Em outras línguas, Tópico e Sujeito não são estruturalmente diferenciados, e “*os sujeitos são essencialmente tópicos gramaticalizados*”: Li e Thompson (1976: 484), apud Hopper-Traugott (1993: 28). Em tais línguas, o Tópico se gramaticalizou ou como um Nominativo (nas línguas nominativo-acusativas), ou como um Absolutivo (nas línguas ergativas). O Sujeito que daí resulta pode ter tanto uma expressão nominal agentiva, nas sentenças transitivas, quando codificar-se como o único nominal, nas sentenças intransitivas. Pode-se dizer, então, que Tópico e Sujeito se fundiram estruturalmente.

Tem-se reconhecido o Português como uma língua que integra as duas tipologias. É uma língua de tópico, porque permite a ocorrência de um sintagma nominal separado da sentença, retomado ou não em seu interior por um pronome-cópia. E é uma língua de sujeito, definido como um constituinte que suscita a concordância no verbo, além de outras propriedades sintáticas: Pontes (1987).

Decat (1989: 114) investigou as pressões estruturais que levaram as construções de tópico, em alguns casos, a serem reanalisadas como “*sujeito da sentença em que aparecem*”. Seus dados procedem dos séculos XVIII a XX. Ela demonstrou que em estruturas como [Tóp. 0 V X], em que 0 representa a ausência do sujeito, o tópico ocupa seu lugar.

2.1.2.2 - Preenchimento do Sujeito e do Objeto Direto

A simplificação da morfologia do Português do Brasil implicou no aumento de sujeitos preenchidos; inversamente, a alteração no quadro dos pronomes pessoais implicou no aumento dos objetos elididos. Tarallo (1991: 82-83) mostra que em 1981 podiam-se identificar 79.4% de sujeitos retidos, contra apenas 18.2% de objetos também retidos. Voltando atrás no tempo, esses valores eram os seguintes:

1725	1775	1825	1880
Sujeito: 23.3%	26.6%	16.4%	32.7%
OD: 89.2%	96.2%	83.7%	60.2%

Duarte (1993) examina peças teatrais brasileiras dos sécs. XIX e XX, encontrando valores muito próximos aos de Tarallo.

2.1.2.3 - Ordem dos Constituintes Funcionais

A gramaticalização da ordem dos constituintes funcionais tem demonstrado, no Português do Brasil, um enrijecimento progressivo. Berlinck (1989) estuda a construção [V-SN1] dos séculos XVIII a XX. Consultando autores tão separados no tempo como Jerônimo Soares Barbosa (1803) e Eunice Pontes (1987), ela notou que para o primeiro a ordem SV ou VS "são ambas naturais", ao passo que, para Pontes, a ordem SV é a predominante, e VS "se mantém em casos especiais, sobretudo em orações marcadas em relação à oração declarativa, afirmativa, neutra". Esta diferente percepção do mesmo fenômeno certamente aponta para uma mudança. E justamente, seus dados mostraram um decréscimo da ordem VS, que obteve 42% de ocorrências no séc. XVIII, 31% no séc. XIX, e apenas 21% no séc. XX. Em sua análise, as regras de colocação do sujeito se alteraram fortemente. De discursivas no primeiro momento examinado (o status informacional especificava a ordem), essas regras passaram a estruturais no segundo (o tipo sintático do verbo especifica a ordem).

2.2 - Morfologização

A morfologização é a criação de formas presas, sejam afixos flexionais, sejam afixos derivacionais.

2.2.1 - Criação de Afixos Flexionais

2.2.1.1 - Flexão Verbal

Os indoeuropeístas sustentam que os afixos de Pessoa derivam da cliticização de Pronomes Pessoais, os quais, por sua vez, provêm de Demonstrativos. Particularmente com respeito à P2, o morfema IDE. *{-t} é uma forma reduzida do Dem. neutro *to-: Lehmann (1982a: 15).

No Português Europeu ocorre fenômeno semelhante: nos dialetos em

que ocorre a assimilação de {-mos} a nós, o Pronome Pessoal nós é dito nes, o sufixo {-mos > -nes}, como em *andábanes*, documentado no Inquérito Lingüístico Boléo. No Português do Brasil, a confluência de fenômenos tais como a simplificação da morfologia verbal e as alterações do quadro dos pronomes pessoais anuncia a cliticização destes ao verbo, e o possível surgimento de prefixos número-pessoais. Tais proclíticos desempenharão a função de sujeito, cuja progressiva retenção foi mencionada no item 2.1.2.

Um fato análogo é a migração de morfemas de Pessoa para as formas nominais de Infinitivo e de Gerúndio. Num primeiro momento, essas formas admitem um sujeito pessoal, e logo depois aparecem os morfemas de Pessoa. Para o Infinitivo Pessoal, v. Maurer (1952). Para o Gerúndio Pessoal, cf. Português Europeu dialetal *andândomos*, documentado igualmente no Inquérito Lingüístico Boléo.

Quanto à gramaticalização dos morfemas de Tempo, nota-se com enorme freqüência que eles procedem de um Verbo Auxiliar. Foi o que ocorreu com o futuro românico, em que *habere* se transforma no morfema {re}, na verdade {e}, se considerarmos que {r} é o morfema do infinitivo. O novo morfema se integra no sub-sistema morfológico da língua, passando a ter um uso obrigatório.

O afixo Ingl. {-ed}, que assinala o tempo passado, deriva de um verbo Auxiliar, o Pré-Germânico *dhe* (Ingl. *to do*), cuja forma de passado era *dhedhoht* > *did*. Posposta a radicais verbais, como em *solpa* + *dheedhoht*, obtiveram-se *solpadhedhet* (morfologização), depois *solboded* (haplologia), Gót. *salboda*, donde o Al. Mod. *salbte*.

Os parágrafos anteriores podem dar a impressão de que os morfemas de Tempo derivam sempre de Auxiliares, o que não é correto. Também Adv. podem originá-los: no Tok Pisin, a expressão adverbial do Ingl. *by and by* originou o marcador temporal *baimbai*; o Adv. *já* indica o passado em várias línguas; no Papiamento, a forma espanhola *luego* se reduz a *lo* e se transforma num marcador de futuro, como em *lo mi kanta*, "eu cantarei".

2.2.1.2 - Flexão Nominal

1) Os morfemas de Caso representam a gramaticalização de itens independentes. Dado o caráter cíclico da gramaticalização, a seguinte hierarquia pode ser observada nos Nomes que já dispõem de Caso: N + Caso > Prep - N + Caso > Prep - N > Caso + N, em que o último estágio é por assim dizer a imagem especular do primeiro.

2) Os três Gêneros do Indo-Europeu põem a questão da origem do feminino. O masculino remetia aos seres animados, o neutro aos inanima-

dos. O morfema do feminino pode ter-se calcado no neutro plural {-a}, que veiculava um sentido de coletivo. Assim, Lat. *iuga*, Gr. *gyné* (em que {-é} está por {-a}) podem ser o coletivo de gemas (de ovo) e de mulheres. Uma evidência paralela para esta afirmação está em que no Gr. o Sujeito neutro plural deixa o verbo no singular.

3) Nem todas as línguas do mundo dispõem da categoria de Número. Nesse caso, podem-se usar Nomes coletivos, sufixados a Nomes /+humano/ ou a Nomes relevantes para a cultura local. É esse o caso do Hixkaryana *harye komo* "batatas doces", em que *komo* é uma partícula pós-nominal de sentido coletivo.

Segundo Lehmann (1982: 58), o próximo passo é a aglutinação dos afixos de número, sufixais em sua maior parte, ocorrendo a mudança *coletivo* > *plural*. Além dos coletivos, também numerais e quantificadores podem dar origem à categoria de Número, como no Tok Pisin *dok*, "cão", *wanpela dok* "um cão", *tupela dok* "dois cães", *tripela dok* "três cães", *ol dok* "cães".

As categorias de número nas línguas do mundo compreendem o Singular, o Plural, o Dual, o Trial e, em alguns casos, o Paucal, que corresponde a uma sorte de quantificação indefinida, como em *poucas maçãs*, sem codificação gramatical no Port.

Línguas como o Port. e o Fr. podem reanalisar o morfema de plural {-s} do Art. Def., que passa de sufixo a prefixo, inicialmente nos casos em que segue um N começado por vogal, como Port. pop. *os olhos* [ozóio] > *zóio*, Fr. *le zofficiers*, e depois com Ns quaisquer, mesmo iniciados por consoante, como Fr. pop. *cinqq z hommes*, *sett z hommes*, etc.

Os chamados plurais cumulativos dão origem a novas regras de pluralização, como se observa nas seqüências mencionadas por Santos Jota, apud Gonçalves (1994: 96): *ilhó* > *ilhós* > *ilhoses*; *ananá* > *ananás* > *anana-ses*; *eiró* > *eirós* > *eiroses*.

O grau mais avançado da gramaticalização do Número ocorre nos casos de alternância vocálica, como Port. *avô* / *avós*, Ingl. *mouse* / *mice*.

Um fenômeno curioso é o modo pelo qual os afixos de plural penetram no SN. No Ingl., essa marca se concentra no núcleo nominal: *the dogs*. Em línguas românicas como o Port. pop. e o Fr., essa marca só ocorre no Det.: Fr. *les hommes* [lezòme], Port. *as casa pequena*. Nos dois casos, evita-se a redundância das marcas de plural.

2.2.2 - Criação de Afixos Derivacionais

Um caso interessante foi o a morfologização do Nome latino *mente*,

feminino, que entre outros sentidos tinha o de "modo, maneira". Esse N pospôs-se a Adjetivos, no sintagma nominal, dando origem a expressões como *agir cristã mente*, isto é, "agir de um modo cristão". Por uma redução fonológica, *mente* se tornou átono, cliticizando-se ao Adjetivo, num processo ainda não consumado, se levarmos em conta expressões como *falou franca e asperamente*.

2.3 - Redução Fonológica

A redução fonológica é um dos processos de gramaticalização mais visíveis. Formas livres fundem-se com outras formas livres, transformando-se em formas presas, gramaticalizando-se como afixos. Braga (1994: 17) menciona "*a relação inversa entre a dimensão dos vocábulos e sua frequência de uso*", formulada por Zipf, a que agrega "*a atuação de um princípio icônico, qual seja, o princípio de que ao conceito familiar se atribui uma expressão reduzida*".

O exemplo sempre lembrado a respeito de redução fonológica é o da formação do futuro nas línguas românicas, em que a forma trissilábica *habeo* reduziu-se ao ditongo *ei*.

2.4 - Estágio Zero

Um morfema tão trabalhosamente construído pode desaparecer, retomando-se o processo *da capo*, com o que uma nova expressão perifrástica fará sua aparição, como se observou acima.

Por que as línguas são assim? Um longo processo de gramaticalização produz categorias fixas, rígidas, inalteráveis. Elas se tornam, por isso mesmo, antifuncionais. A cristalização de estruturas contraria a natureza mesma da língua, no que ela tem de criativo, de transitório, de re-elaborador. Os argumentos de Franchi (1977: 32) em defesa da linguagem como uma atividade de permanente reformulação caem aqui como uma luva: "*A linguagem na medida em que 'dá forma' é bem já uma atividade quase-estruturante, mas não necessariamente 'estruturada', no sentido estrito do termo, ou se concebermos 'estrutura' como uma organização estável de categorias. (...) A linguagem natural permanece sempre o instrumento de uma prática primitiva de estruturação dos fatos da experiência, de revisão e reformulação: uma espécie de 'lógica' primitiva e fraca que não se cinge às restrições das propriedades formais*".

O estágio zero é o momento máximo de exaustão da estrutura, e anuncia a retomada do processo contínuo que é a gramaticalização.

2.5 - Alteração semântica: metáfora e metonímia

O surgimento do futuro e do pretérito perfeito composto nas línguas românicas implicou em que *habere* perdesse seu sentido original de posse, assumindo o valor modal de "dever", no caso do futuro, e o valor aspectual resultativo e temporal de passado, no caso do pretérito perfeito composto. O mesmo pode dizer-se de *ter*, em *tenho de ir*, e de *ir*, que na indicação do futuro deixou de significar movimento no espaço físico, para indicar um movimento mais abstratizado, no espaço-tempo.

Alteração semântica sofreu também *passu*, termo que se construía com verbos de movimento, como em Port. *non vou passo*, Fr. *je ne vais pas*, isto é, "não percorrerei o espaço de um passo". Com o termo, perde-se completamente o sentido original, e *passo / pas* tomam o sentido de negação, importado de *non/ne*, a ponto de se poder omitir esses advérbios, sem perda do sentido de negação: *vou passo, vais pas*.

A gramaticalização implica, portanto, na perda do sentido original da forma sob transformação, desenvolvendo-se novos sentidos.

Vários autores notaram a diferente carga semântica das palavras. Meillet, como vimos, distingue "palavras principais", isto é, plenas de sentido, de "palavras acessórias" e de "palavras gramaticais". Vendryes (1921) fala em "palavras lexicais" e "instrumentos gramaticais". Ullman (1964: 94), reportando-se a diversos autores, separa "palavras plenas ou auto-semânticas", "*significativas por si próprias*" (Nomes, Adjetivos, Verbos, Advérbios [Predicativos]), das "palavras-forma, ou sinsemânticas", "*significativas apenas quando aparecem com outras palavras*" (Preposições, Conjunções, Artigos e Advérbios [Não-Predicativos]).

Em seu conjunto, esses autores captaram um dos mecanismos do que estou chamando *semanticização*: em dados contextos lingüísticos, as palavras sofrem ligeiras alterações de suas propriedades intensionais de base; em outros contextos, essas propriedades são afetadas mais severamente, desenvolvendo-se sentidos novos. Tais alterações têm sido denominadas *dessemantização*, *bleaching*, *fading*. Trata-se de termos impróprios, pois, como se disse, o que se observa é uma mudança mais ou menos severa, mas nunca uma perda semântica, um esvaziamento do significado. Em qualquer um desses casos, dizemos que o item tramitou do Léxico para a Gramática, isto é, gramaticalizou-se.

Uma questão não resolvida é se a mudança sintática acarreta a mudança semântica, ou se a mudança semântica é um estágio precoce da mudança sintática. Só mesmo postulações teóricas sobre a natureza desses módulos da linguagem permitem resolver o impasse.

Pesquisas sobre a gramaticalização têm mostrado que no continuum criado por seus estágios é possível identificar pontos focais, não discretos, como no surgimento de verbos auxiliares. Os itens em processo de gramaticalização são polissêmicos, apresentando diversas acepções relacionadas umas com as outras. Relações de *metáfora* e de *metonímia* podem ser estabelecidas entre esses pontos focais.

2.5.1 - A Metáfora

A metáfora é a transferência de um sentido A para um sentido B, por haver alguma similaridade entre eles; trata-se, basicamente, de um processo cognitivo. A metáfora não opera mudanças radicais nas propriedades intensionais do item. Assim, um verbo como *ver* "olhar, perceber com os olhos", passa a significar "saber, perceber com a razão", e depois, "compreender". De A para B, e de B para C, percebe-se um movimento que vai de um sentido básico, mais concreto, para sentidos derivados, mais abstratos, preservado o sentido de base de "captar algo". A Teoria dos Casos capta muito bem esse movimento do concreto, discursivamente motivado, para o abstrato, estruturalmente motivado. Segundo Parret (1994: 28), a teoria "semilocalista" dos casos sustenta que há três casos "gramaticais", de motivação estrutural (o nominativo, caso do sujeito; o acusativo, caso da determinação do verbo; o genitivo, caso da determinação do nome), e quatro casos "locais", de motivação discursiva (o ablativo, lugar 'de onde'; o locativo, lugar 'onde'; o dativo, lugar 'para onde'; o instrumental, lugar 'conjunto').

Os itens que significam partes do corpo humano e espaço físico oferecem outros exemplos de continuum metafóricos: a gramaticalização dos primeiros dá origem às preposições; a dos segundos, aos auxiliares temporais, e assim por diante, conforme anotado em 2.1.1.

Da noção de Espaço físico para a de Espaço-Tempo é outro salto, de natureza igualmente cognitiva: lembre-se que freqüentemente a gramaticalização do tempo futuro e do passado se dá mediante o aproveitamento dos verbos de movimento. Metaforizados, eles passam a significar (i) o trânsito do tempo presente para o tempo futuro, mediante a utilização de *ir* (cf. "*vou buscar comida*"), ou (ii) o trânsito do passado para o presente, mediante a utilização de *vir* (cf. "*venho de encontrá-lo ali na rua*").

Tem-se notado que há um ritmo unidirecional nos processos de metáfora, do mais concreto para o mais abstrato. Mas o importante neste processo está em que "*muito da polissemia se deve ao uso metafórico; de fato, não apenas a linguagem, mas nossa cognição e portanto a língua opera metaforicamente*": Sweetser (1990: 8; v. também pág. 20 e ss).

2.5.2 - A Metonímia

A metonímia é a mudança de sentido desencadeada por itens associados sintaticamente; trata-se, basicamente, de um processo estrutural. A alteração do sentido de *passu* "passo", que assimila de *non* um sentido negativo que não era seu, é um caso de metonímia que afetou a gramática da língua, pois deu surgimento a um novo advérbio de negação. O mesmo pode dizer-se de Lat. *nata*, particípio passado de *nascor*, que assumiu sentido negativo ao associar-se com *nulla*, na expressão *nulla re nata*. Assim, na metonímia a palavra é ressemantizada, e o sentido B não guarda relações com o sentido A, ocorrendo uma perda radical de propriedades intensionais.

A formação de conjunções freqüentemente assenta em alterações metonímicas: Adv. de Inclusão *magis* > Conj. Adversativa *mas*, em que este item passa a codificar a contrajunção, derivada do uso de *mas* em contextos negativos: Castilho (1995a).

3. Princípios da Gramaticalização

Não há acordo entre os lingüistas quanto a uma teoria, se não unificada, pelo menos razoavelmente articulada, para dar conta dos fenômenos ditos de gramaticalização, listados no item anterior. Nem poderia ser diferente, dada a complexidade da matéria e as naturais divergências sobre o que se pode entender por língua e por gramática.

Começarei por enumerar algumas generalizações já elaboradas.

Lehmann (1982a) identificou os cinco princípios seguintes:

- (1) Paradigmatização: de acordo com este princípio, *construções sintáticas integram-se como formas perifrásticas em paradigmas morfológicos [o que] conduz a paradigmas progressivamente pequenos e homogêneos.*
- (2) Obrigatoriedade: *dentro do paradigma, a escolha entre seus membros sujeita-se a regras gramaticais. O conjunto categorial representado pelo paradigma torna-se progressivamente obrigatório nas sentenças da língua.*
- (3) Condensação: *quanto mais um signo se gramaticaliza, tanto menos complexos se tornam os constituintes com os quais ele pode combinar-se.*
- (4) Coalescência: este princípio se manifesta *da justaposição (via cliticização, aglutinação e fusão) para uma alternância simbólica.*
- (5) Fixação: *o signo gramaticalizado tende a ocupar uma posição fixa inicialmente na sintaxe, e depois na morfologia, tornando-se um preenchedor de espaços gramaticais (slots).*

Ele retomaria essas formulações posteriormente, em Lehmann (1982b: 121 e ss). Nesse texto, ele argumenta que a gramaticalização diz respeito

essencialmente à autonomia do signo. Quanto mais autônomo o signo, menos gramaticalizado, e quanto menos autônomo, mais gramaticalizado.

Três aspectos permitem medir o grau de autonomia de um signo: (1) o *peso*: para ser autônomo, um signo deve ter certo peso, propriedade que o distingue dos demais membros de sua classe, proporcionando-lhe proeminência no paradigma; (2) a *coesão*: quanto mais relações com outros signos, tanto menos autônomo se tornou o signo; (3) a *variabilidade*: quanto maior a mobilidade do signo, tanto maior será sua autonomia.

O decréscimo no peso e na variabilidade e o aumento da coesão se constituem, portanto, nos três aspectos da gramaticalização. Para melhor operacionalizá-los, será necessário relacioná-los com dois aspectos lingüísticos fundamentais: a seleção e a combinação dos signos, vale dizer, o aspecto sintagmático e o aspecto paradigmático.

O *peso sintagmático* de um signo é sua *integridade*, seu tamanho substancial, tanto semântico quanto fonológico. O *peso paradigmático* é seu *escopo*, a saber, o número de construções nas quais ele entra e que ajuda a formar.

A *coesão paradigmática* de um signo será chamada *paradigmaticidade*, isto é, o grau de integração de um signo no paradigma. A *coesão sintagmática* de um signo será denominada *conexidade* ("bondedness"), isto é, o grau de sua adesão a outros signos.

A *variabilidade paradigmática* de um signo é a possibilidade de usar outros signos no seu lugar, ou mesmo a possibilidade de omiti-lo. A *variabilidade sintagmática* de um signo é a possibilidade de substituí-lo em sua construção.

O seguinte quadro reúne esses seis critérios (v. p. 123):

	Paradigmático	Sintagmático
peso	integridade	escopo
coesão	paradigmaticidade	conexidade
variabilidade	variabilidade	variabilidade
	paradigmática	sintagmática

Ele deixa claro que esses parâmetros, fatores ou critérios dizem respeito ao signo; os graus com que eles afetam os signos deitam luz sobre o processo da gramaticalização, visto que esta pode ser medida a partir da observação daqueles. Entre págs. 125 e 179 ele examina com mais detalhe a

ação e a interação desses parâmetros, resumizando tudo no seguinte quadro:

Parâmetro	Gramaticalização fraca —>	Processo —>	Gramaticalização forte
Integridade	Conjunto de propriedades semânticas; possivelmente polissilábico —>	Desgaste —>	Poucas propriedades semânticas; monosssegmental
Paradigmaticidade	O item participa escassamente no campo semântico —>	Paradigmaticidade —>	Fortemente integrado no paradigma
Variabilidade paradigmática	Livre escolha dos itens, de acordo com as intenções comunicativas —>	Obrigatoriedade —>	Escolha sistematicamente restrita, uso largamente obrigatório
Escopo	O item se correlaciona com constituintes de complexidade arbitrária —>	Condensação —>	O item modifica palavra ou raiz
Coesão	O item é justaposto independentemente —>	Coalescência —>	O item é afixo ou até mesmo suporte de traço fonológico
Variabilidade sintagmática	O item pode mover-se livremente nas estruturas —>	Fixação —>	O item ocupa lugares gramaticais fixos

Entre outros, o mérito desse quadro é o de mostrar que a gramaticalização está sujeita a uma graduação forte / fraca, aspecto que é focalizado pelo princípio da continuidade e gradualismo (v. item 3.3).

Hopper (1991) admite a existência de cinco princípios, alguns deles muito assemelhados:

- (1) Estratificação ("layering"): *Dentro de um domínio funcional amplo, novas camadas emergem continuamente. Quando isto acontece, as camadas mais antigas não são necessariamente descartadas, mas podem continuar a coexistir e a interagir com as camadas mais novas.*

Este "princípio" é a constatação de que uma das conseqüências da gramaticalização é a convivência de soluções gramaticais distintas, num mesmo corte sincrônico. Isso vem reconhecido em Hopper-Traugott (1993: 124). Assim, a gramaticalização dá surgimento à variação lingüística, considerada por alguns autores como o gatilho que dispara a mudança diacrônica.

- (2) Divergência: *Quando a forma lexical gramaticalizou-se num clítico ou num afixo, a forma lexical original permanece como um elemento autônomo e sofre as mesmas mudanças que um item lexical comum.*

Esta generalização já está, de certa forma, contida na anterior.

- (3) *Especialização: Dentro de um domínio funcional, é possível haver em determinado estágio uma variedade de formas com nuances semânticas diferentes; quando a gramaticalização ocorre, estreita-se essa variedade de escolhas formais, e um número menor de formas selecionadas assume significados semânticos mais gerais.*
- (4) *Persistência: Quando uma forma se gramaticaliza, passando de uma função lexical para uma função gramatical, tanto quanto isto seja gramaticalmente viável, alguns traços do seu significado lexical original tendem a aderir à nova forma gramatical, e detalhes de sua história lexical podem refletir-se na sua distribuição gramatical.*

Os itens (3) e (4) captam processos semânticos que acompanham a gramaticalização.

- (5) *Descategorização: Formas em processo de gramaticalização tendem a perder ou a neutralizar as marcas morfológicas e as propriedades sintáticas das categorias plenas Nome e Verbo, e a assumir atributos característicos das categorias secundárias tais como o Adjetivo, o Particípio, a Preposição, etc.*

O termo escolhido por Hopper aqui remete ao ponto de partida do processo de gramaticalização, pois, como é óbvio, Nome e Verbo perdem propriedades ao se recategorizarem como Adjetivo, Preposição, Auxiliar, etc. Encarado o processo como um ponto de chegada, o que se verifica, entretanto, é a aquisição de propriedades novas, isto é, há uma recategorização, como demonstrei em 2.1.1.

Em suma, os princípios acima combinam *processos* e *estágios* aos *princípios* propriamente ditos. Isto é, de um lado existem os mecanismos que levam uma categoria lexical a transformar-se numa categoria gramatical, e de outro lado há os princípios gerais que regem essa mudança de estatuto.

Atendo-me agora às formulações que de fato captam generalizações, vou admitir que quatro princípios podem dar conta dos estágios da gramaticalização: (i) a analogia, (ii) a reanálise, (iii) a continuidade e o gradualismo, e (iv) a unidirecionalidade. Passo a desenvolver estes tópicos.

3.1 - Analogia

A analogia é uma sorte de aproximação psicológica entre categorias em ausência, isto é, entre categorias situadas no eixo paradigmático. Esse princípio leva à criação de formas não explicáveis se tomadas no eixo sintagmático. A analogia não dá surgimento a expressões ou estruturas novas, ela simplesmente estende regras a itens ainda não atingidos, "uniformizando", por assim dizer, as formas da língua. Desnecessário dizer que derivas em sentido contrário conservam as formas irregulares, e por isso a ação da analogia não é uniforme.

Jakobson e Halle (1956), apud Hopper-Traugott (1993: 56), já tinham reconhecido que a analogia opera no eixo paradigmático, por contraste com a reanálise, que opera no eixo sintagmático. A seguinte fórmula capta bem esse processo: $A:B::C:D$, em que D representa a forma surgida por analogia.

O processo da analogia foi identificado pelos neogramáticos, que o aplicaram inicialmente à mudança fonológica. Ficou famosa, em Port., a explicação da nasalidade de *sim* (< Lat. *sic*), dada sua relação por antonímia com *non*.

A analogia também afeta o sistema gramatical. No Port., o pretérito atual de *estar*, *estive*, é analógico ao pretérito arcaico de *ser*, *sive*. Ambos verbos eram usados locativamente, e isto permitiu o trânsito de formas de um para o outro, desaparecendo o pretérito etimológico arcaico *estede* (< Lat. *steti*), em favor da forma analógica *estive*, calcada em *sive*. Assim, *ser* : *sive* :: *estar* : *estive*. No Ingl., temos a seqüência *sing:sung::bring:brung*, forma esta que surgiu no lugar de *brought*.

A analogia e a reanálise podem confluir na gramaticalização de uma mesma expressão. Assim, *ir* + *infinitivo [de destinação] de verbo de movimento*, como em *vou caminhar um pouco*, isto é "vou sair para caminhar um pouco", foi reanalisado como [*vou caminhar*] *um pouco*, em que *ir* se tornou auxiliar do verbo no infinitivo, assinalando a futuridade. Por analogia, verbos não de movimento passaram a construir-se com o auxiliar *ir*, dando surgimento a expressões aparentemente contraditórias, como *vou ficar por aqui mesmo*, consumando-se a noção de futuridade.

3.2 - Reanálise

A reanálise é um processo por meio do qual os falantes mudam sua percepção de como os constituintes de sua língua estão ordenados no eixo sintagmático. Essa mudança de percepção se deve a um tipo de raciocínio conhecido como *abdução*.

A abdução é distinta da indução e da dedução. Segundo Andersen (1973), apud Hopper-Traugott (1993: 39), "*a abdução provém de um resultado observado, invoca uma lei, e infere então o que poderia ser o caso; assim, dado o caso que Sócrates morreu, podemos correlacionar esse fato com a lei geral de que todos os homens são mortais, e portanto Sócrates era um homem*". É bem certo que a abdução (ou inferência) é um tipo fraco de raciocínio, mas, como Peirce já tinha notado, "*ela é a base da percepção humana, e o único tipo de raciocínio através do qual se pode dar origem a novas idéias*" (ibidem: 40).

Através da abdução, apagamos os limites entre determinados constituintes, estabelecendo novos "cortes", sem alterar a manifestação superficial da unidade sobre que estamos operando. Entretanto, a reanálise vai ter conseqüências futuras, mesmo no eixo sintagmático, pois uma nova categoria gramatical estará surgindo.

A formação do futuro romance passou por esse processo. Seja a expressão *haec dicere habeo* "disponho de algumas coisas para dizer", em que *habeo* é Possessivo, tendo *haec* por objeto direto. Numa primeira reanálise, ela passou a ser interpretada como [*haec dicere*] *habeo*, "tenho que dizer estas coisas", em que todo o segmento entre colchetes foi analisado como objeto direto de *habeo*, ao mesmo tempo em que este verbo deixava de ser Possessivo, transformando-se em Auxiliar Modal. Numa segunda reanálise, aumenta a conexidade sintática entre *dicere* e *habeo*, e *haec* passa a complemento desse conjunto, reinterpretado como *haec [dicere habeo]*, "tenho de dizer algo". Esta última reanálise abriu as portas à fusão fonológica da expressão entre parênteses, surgindo assim o novo morfema de Futuro {*re*}, desaparecendo os morfemas {-*bo* / -*am*} do latim culto.

A formação do futuro no Ing. teve uma história semelhante. Numa expressão como [*I am going*] [*to sell my car*], o verbo *to go* mantém seu sentido pleno de movimento no espaço físico, e o segundo constituinte assinala a finalidade dessa deslocação. Reanalizada como [*I am going to*] [*sell my car*], uma nova unidade apareceu, o progressivo *I am going to*, cuja gramaticalização prossegue, a ponto de se poder omitir o pronome e o auxiliar, donde *going to* que, reduzido fonologicamente, deu lugar a *gonna*, uma nova forma de futuro. Como no português, o estágio máximo de gramaticalização se consuma quando esta forma passa a combinar-se com verbos estativos, tais como *gonna stay*.

O surgimento do pretérito perfeito composto nas línguas românicas envolveu, igualmente, o processo da reanálise. A expressão *habeo litteras scriptas*, "tenho cartas escritas", em que o constituinte *litteras* era analisado como objeto direto de *habeo*, e *scriptas* como predicativo desse objeto, foi reanalizada como [*habeo scriptum*] *litteras*, em que *habeo* perdeu o sentido de posse, passando a assinalar o resultado presente de uma posse no passado, *scriptum* se deslocou no sintagma em direção a *habeo*, tornando-se invariável e passando a constituir o núcleo da predicação, tomando ambas essas formas o constituinte *litteras* como seu objeto. A história do clítico o em Português é mais um exemplo do efeito da reanálise. Em "*video illum librum*", o recorte "*video [illum librum]*" deu surgimento ao Artigo o < *illum*; a omissão do N deve ter ocasionado a segmentação "[*video illum*] 0", donde o clítico pronominal o.

Os exemplos acima mostram que a reanálise é o *desenvolvimento de novas estruturas a partir de estruturas antigas*: Hopper-Traugott (1993: 56).

3.3 - Continuidade e gradualismo

De acordo com este princípio, que aparece nos escritos de Kurylowicz, apud Lehmann (1982b), a gramaticalização tende continuamente à inovação da estrutura das línguas.

Sapir (1921), e posteriormente Labov (1995) postulam que a variação é o primeiro estágio da mudança sintática. Os trabalhos sobre a gramaticalização invertem essa ordem. O princípio de "estratificação" de Hopper (1991) aponta para essa direção. Lichtenberk (1991: 37) sustenta que *a variação é uma conseqüência necessária do gradualismo da mudança lingüística*. Dado o caráter cíclico da gramaticalização, parece não haver contradição em afirmar que a variação é ao mesmo tempo o ponto de partida e o ponto de chegada da mudança lingüística.

Há um ponto que parece grangear uma aceitação geral entre os lingüistas: o fato de que a gramaticalização não é um processo que possa extinguir-se, sendo muito difícil identificar claramente suas fases. Ele é recorrente e interminável, e alguns exemplos enumerados anteriormente poderão demonstrá-lo. É o que reconhecem Heine e Reh (1984), apud Heine-Claudi-Hünemeyer (1991): *a gramaticalização é um processo contínuo. Qualquer tentativa de segmentá-lo em unidades discretas será de certa forma arbitrária*.

O princípio da inovação permite distinguir gramaticalização de troca lexical. Nos casos de troca lexical, com efeito, um item substitui outro, como ocorreu com *ser* estativo, trocado por *estar* nesse tipo de construção. A gramaticalização é um processo contínuo, pois radica numa propriedade das línguas naturais, que é sua permanente criatividade, muitas vezes motiada por necessidades sociais.

A continuidade da inovação tem por efeito uma continuada re-estruturação das línguas, o que favorece o conceito de "gramática emergente", formulado por Hopper (1987). V. também Hodge (1970). Num dos extremos do continuum, temos o domínio das escolhas, em que diferentes formas lingüísticas co-ocorrem, sendo selecionadas por critérios discursivos. No outro extremo, temos o domínio das determinações, em que estruturas lingüísticas restringem a liberdade do usuário. As primeiras formas, "menos gramaticais", fazem as delícias dos funcionalistas. As segundas, as dos formalistas. DuBois (1985) chamou a isto "motivações em competição".

3.4 - Unidirecionalidade

Ao caráter cíclico, vários autores citados em Lehmann (1982b: 16) agregam que a gramaticalização é unidirecional, isto é, trata-se de um processo irreversível, que só pode desenvolver-se da esquerda para a direita, se tomamos em consideração o quadro "Fases da gramaticalização", reproduzido anteriormente. Não poderia haver, portanto, "desgramaticalização". Assume-se, basicamente, que a um estágio A deve seguir-se um estágio B, e não o contrário: Hopper-Traugott (1963: 95).

O livro de Heine-Claudi-Hünemeyer (1991) traz um extenso relatório sobre o avanço dos estudos de gramaticalização, aplicados às línguas africanas. Quero aqui destacar sua contribuição ao princípio da unidirecionalidade, que eles reconhecem enfaticamente, quando sustentam que as estruturas "menos gramaticais" podem tornar-se "mais gramaticais", porém não o contrário: *"embora tanto a desgramaticalização como a regramaticalização tenham sido observadas, elas são estatisticamente insignificantes e serão ignoradas neste trabalho. Observe-se que muitos casos de uma desgramaticalização alegada discutidos na literatura resultam comprovadamente de uma análise inadequada"* (pp. 4-5).

A dificuldade na admissão da desgramaticalização está em que tais formas deveriam passar por um processo de ressemantização e de reforço fonológico. Exemplos claros desses processos ainda não foram encontrados. Enumero alguns deles:

- (1) A nominalização de alguns sufixos configuraria um caso de desgramaticalização: uma forma presa torna-se livre, como Port. "os *ismos da ciência tal*", em que o morfema derivacional {-ismo} se transforma num Nome, com o sentido de "tendência, direção", Inglês / Espanhol, *bus* "ônibus", que resultou da nominalização do morfema flexional latino {-bus}, destacado da palavra *omnibus*, "para todos", passando a designar o veículo de uso coletivo.
- (2) A perda da "gramaticidade" de constituintes de palavras, obrigando o usuário a duplicá-los, seria outro caso de desgramaticalização. Exemplos disto foram dados anteriormente, a propósito da repetição da mesma Prep. numa perífrase preposicional. No Port., é bem conhecido o caso da forma medieval *migo*, que já continha a Prep. *com* (cf. Lat. *mecum*, de *cum* + *me* > Port. *migo*), e que passou de novo a admiti-la na forma atual *comigo*. Parece bem evidente que a perda do poder relacional das Preps. e conseqüente repetição do item seria melhor descrita como uma *regramaticalização*.
- (3) Outros têm exemplificado como casos de desgramaticalização a formação de delocutivos como *falou ?* e de Marcadores Conversacionais como *tá ? sabe ? entende ? compreende ? viu ?*

né ? Nesses casos, um Verbo deixa de funcionar como núcleo da sentença, "desgramaticalizando-se". O ponto tem sido sustentado por alguns lingüistas brasileiros. Eles argumentam que nos exemplos acima os itens passam a apresentar-se numa forma fixa (não podem ter alterações de tempo e de pessoa), ocupam lugares determinados nos enunciados falados (à direita das unidades discursivas) e não mais organizam sentenças, como núcleos predicadores. Um corolário desse argumento é que seriam postulados como categorias do discurso quaisquer itens privados de propriedades gramaticais - estratégia analítica que, no mínimo, seria muito estranha para a configuração de um campo científico.

Casos como estes, entretanto, podem ser interpretados como fenômenos de *discursivização*. Se postularmos a língua como uma atividade criativa, por que não avaliar os dados acima pelo ângulo do ganho de propriedades discursivas? Além do mais, se reconhecemos que delocutivos e Marcadores Conversacionais são categorias do discurso, por que procurar neles propriedades sintáticas, mais claramente identificáveis no limite da sentença? E por que supor que propriedades discursivas excluem propriedades gramaticais? Não é verdade que os Marcadores Conversacionais têm uma distribuição regular nas unidades suprasentenciais? Isso não é a gramática convivendo com o discurso? Estas considerações nos levam à proposta constante das Conclusões.

Conclusões: discursivização, semantização, gramaticalização

Na Apresentação deste trabalho, observei que os autores consultados não demonstram em seus textos uma preocupação maior em apresentar o quadro teórico que articula suas afirmações. Não será demais lembrar aqui a afirmação de Saussure, para quem a língua é um "objeto científico escondido". Portanto, sem postulações prévias torna-se difícil desenvolver indagações lingüísticas.

Uma teoria modular da língua poderá ser de ajuda para o encaminhamento do debate.

Creio que os primeiros passos na postulação dessa teoria foram dados, na Lingüística Moderna, por Morris (1938). Nesse texto, ele estabelece os domínios da Sintaxe (= estudo da relação de signos entre si), da Semântica (= estudo da relação dos signos com seu referente) e da Pragmática (= estudo da relação dos signos com seus usuários). Entendendo por Gramática a Fonologia, a Morfologia e a Sintaxe, vê-se que por "gramaticalização" apreende-se a criatividade situada apenas num dos domínios das línguas.

Jakobson (1948: 162) retoma o signo como ponto de partida, e mos-

tra que seu "dualismo indissolúvel" *"é o ponto de partida da Lingüística moderna no combate obstinado que ela move em duas frentes (...), o som e os sentidos"*. Fixando-se na Fonologia e na Morfologia, ele as postula como aspectos autônomos das línguas naturais, porém não independentes. Essas formulações são importantes, pois tratam da modularidade no interior da Gramática, destacando a conveniência teórica de tratá-las autonomamente, mas sem perder de vista os pontos de integração entre elas. Este é, aliás, o ponto central da argumentação de Jakobson nesse texto.

Franchi (1977, 1991) reconhece tacitamente os três módulos mencionados por Morris, destaca que eles são articulados pelo Léxico, e insiste em que não há relações de determinação entre eles. Lembro que Franchi rotula-os de "sistemas". Eles devem ser concebidos de tal forma que se preserve a autonomia de cada um, e o lingüista buscará, quando muito, identificar regras de correspondência entre eles.

Considero do maior interesse partir dessas idéias para um debate sobre a gramaticalização. Teríamos a vantagem de evitar "entrar por uma porta e sair pela outra" em matéria de estudo dos processos constitutivos das línguas. Assim, não vejo a menor utilidade em sustentar que uma categoria do discurso se transforma numa categoria sintática - como no caso das Construções de Tópico, em seu trâmite suposto para a categoria funcional de Sujeito. Gostaria de insistir em que qualquer item lexical contextualizado nos usos da língua preserva, ao mesmo tempo, suas propriedades sintáticas, discursivas e semânticas, sem que precisemos estabelecer correlações de precedência genética entre eles.

Nesse debate, a compreensão do que seja o Léxico assume a maior relevância, pois a grande maioria dos autores consultados parte desse componente, afirmando repetidas vezes que a gramaticalização é um conjunto de fenômenos que ocorrem simultaneamente, quando um item lexical passa a ter um papel significativo na estrutura das línguas naturais.

Se considerarmos que a gramaticalização é a migração de um item lexical para uma categoria gramatical, o pressuposto será o de que estamos estabelecendo uma distinção forte entre "itens lexicais", como elementos que compõem o Léxico de uma língua, onde imperam as idiosincrasias, e "categorias gramaticais", como morfemas ou como palavras funcionais (verbos auxiliares, nexos, etc.) que integram a Gramática da língua, onde imperam as regularidades.

Ora, é preciso ter em conta que mesmo fora do enunciado (em "estado de dicionário", no feliz achado de Carlos Drummond de Andrade), as palavras já dispõem de propriedades gramaticais (fonológicas, morfológicas

e sintáticas), semânticas (referencialidade, predicatividade, dêixis, foricidade) e discursivas (topicidade, coesividade, etc). Apesar disso, afirmações colhidas na bibliografia indicam que alguns autores consideram os itens lexicais como uma entidade tábua-rasa, "em branco", sem propriedades lingüísticas de qualquer natureza. Já outros têm uma visão mais ordenada do Léxico, quando dividem as palavras nas seguintes hierarquias, atrás mencionadas:

Léxico

Palavras principais	Palavras acessórias	Palavras gramaticais
Palavras lexicais		Instrumentos gramaticais
Formas livres		Formas presas
Palavras plenas ou auto-semânticas		Palavras-forma, ou sin-semânticas
Categoria maior (N, PRO, V)	Categoria mediana (Adj, Adv)	Categoria menor (Prep, Conj)
Verbo pleno	Verbo funcional	Verbo auxiliar

É portanto muito razoável conceber o Léxico como o módulo central da língua, em que estão depositados itens já marcados por propriedades gramaticais, discursivas e semânticas. Discurso, Gramática e Semântica serão postulados como módulos que dispõem de regras próprias, situando-se em pé de igualdade uns em face dos outros. Um dispositivo pré-verbal, de natureza cognitiva, ativaría a exploração das potencialidades dos itens lexicais, no momento da criação das formas lingüísticas, vale dizer, no momento da enunciação. Com isto, só mesmo por postulação teórica se escolheria o módulo dado como o desencadeador do processo. Os funcionalistas elegem o Discurso, privilegiando a enunciação. Os formalistas, a Gramática, privilegiando o enunciado. A Semântica se dilui nos programas dessas duas grandes tendências teóricas.

Tendo por pano de fundo essa teoria modular da língua, será possível propor três grandes processos de constituição da língua: a *gramaticalização*, a *semantização* e a *discursivização*. Esses processos se desencadeiam simultaneamente num mesmo item, sem excludência.

A seguir, exponho ao debate alguns elementos que poderiam compor um plano sistemático de estudo dos processos constitutivos da língua, procedendo a uma das leituras possíveis de Franchi (1977, 1991).

Discursivização

A *discursivização* é o uso discursivamente relevante de itens lexicais. Algumas regras de correspondência com os sistemas gramatical e semântico poderiam ser estabelecidas para o entendimento de questões como: (1) itens que se prestam à organização da hierarquia tópica; (2) itens que se prestam ao estabelecimento da coesão textual e à manutenção da interação. Neste particular, o caso mais notável é o dos marcadores conversacionais, sobre os quais surgiram muitos estudos ultimamente.

Semanticização

É a alteração semântica por que passam os itens em sua utilização gramatical ou discursiva. Acompanhando sua recategorização gramatical, os itens podem ser ressemantizados, sofrer alterações na direção de seu escopo ou em sua abrangência fórica, e assim por diante. Parece-me inadequado retratar tais casos como fenômenos de "dessemantização". Não há perdas, há modificações de sentidos.

As seguintes fontes de criação conceptual poderiam constituir uma agenda de pesquisas sobre a semanticização: (1) Semântica Lexical e criação dos sentidos: intensão, extensão, metáfora, metonímia, etc. (2) Semântica Sintática ou Composicional e criação dos significados: referência, predicação, dêixis, foricidade, conjunção. (3) Semântica Discursiva e criação das significações: inferência, pressuposição, etc.

Esse plano de estudos precisaria ser melhor elaborado, o que não posso fazer aqui. De todo modo, creio que, juntamente com a gramaticalização, pesquisas sobre a discursivização e a semanticização permitiriam considerar mais adequadamente as línguas naturais em seu dinamismo. Esta é a proposta que submeto ao debate. Quando dispusermos de mais pesquisas nessas áreas, estaremos preparados para dar outro passo: desvendar os processos cognitivos gerais que, em primeira instância, movimentam os processos secundários da gramaticalização, da discursivização e da semanticização.

Abstract

In this paper I argue that recent studies on grammaticalization are not clearly related to an identifiable theory of language and grammar. I propose to distinguish *stages* (such as syntacticization, morphologization, phonological reduction and semantic bleaching) from *principles* of grammaticalization (analogy, reanalysis, continuity and unidirectionality), with examples drawn from Portuguese. In the final remarks I propose that grammaticalization, together with discursivization and semanticization, could be admitted as general creative linguistic processes, in the frame of a modular point of view of language.

Referências Bibliográficas

- Abraham, W. 1991. The grammaticization of the German modal particles, em Traugott-Heine Eds. 1991, vol. 2: 331-380.
- Amaral, A. 1922. **O Dialeto Caipira**, 2a. ed. São Paulo, HUCITEC, 1977.
- Andrade, M. L.C. V.de O. 1995. **Digressão: uma estratégia na condução do jogo textual-interativo**. São Paulo, FFLCH-USP, Tese de Doutorado, mimeo.
- Benveniste, É. 1968. Mutations of linguistic categories, em Lehman and Malkiel, eds., pp. 85-94.
- Berlinck, R.A. 1989. A Construção V SN no Português do Brasil: uma visão diacrônica, em F. Tarallo Org. 1989, pp. 95-112.
- Braga, M. L. 1994. A Dimensão dos Constituintes no Português do Brasil, **Tempo Brasileiro** 117: 17-26.
- Bybee, J. - Pagliuca, R. and Perkins, R.D. 1991. Back to the Future, em Traugott-Heine Eds. 1991, vol. 2, pp. 17-58.
- Bybee, J. - Perkins, R. and Pagliuca, W. 1994. **The Evolution of Grammar**. Tense, Aspect and Modality in the Languages of the World. Chicago and London, The University of Chicago Press.
- Câmara Jr., J. M. 1972. **Dispersos**. Rio de Janeiro, Fundação Getúlio Vargas.
- Castilho, A.T. de 1993. A repetição na língua falada como um processo gramatical, em **Linguística** 5: 1994,
- 1994a. Problemas de Descrição da Língua Falada, **DELTA** 10 (1): 47-71.
 - 1994b. Um ponto de vista funcional sobre a predicação, **Alfa** 38: 75-96.
 - 1995a. Conversação e Gramática, submetido à publicação em **Filologia e Linguística Portuguesa** 1: 1996.
 - 1995b. Língua Falada e Gramaticalização: o caso da conjunção *mas*, em Maria Helena Moura Neves e Ângela C. S. Rodrigues, em organização, **Gramática do Português Falado**, vol. VII.
- Decat, B. 1989. Construções de Tópico em Português: uma abordagem diacrônica à luz do encaixamento no sistema pronominal, em F. Tarallo Org. 1989, pp. 113-139.
- Dias de Moraes, L.C. 1987. **Nexos Coordenativos no Português Falado**. São Paulo, Universidade de São Paulo, Tese de Doutorado, mimeo.
- Duarte, M.E.L. 1993. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil, em I. Roberts e M. Kato Orgs. 1993, pp. 107-128.
- DuBois, J. 1985. Competing Motivations, em J. Haiman Ed. 1985, pp. 343-365.
- Ducrot, O. - Vogt, C. 1978. *De magis a mais*: une hypothèse sémantique. **Revue de linguistique romane** 43: 317-341.
- Franchi, C. 1977. Linguagem - Atividade Constitutiva, **Cadernos de Estudos Linguísticos** 22: 9-39, 1992 [republicação de **Almanaque** 5: 9-26, 1977].
- 1991. Concepção de uma E-Gramática, conferência inédita pronunciada no Dep. de Linguística da UNICAMP, 26 de setembro de 1991.
- Givón, T. 1971. Historical syntax and synchronic morphology: an archaeologist's field trip, em **Chicago Linguistic Society** 7: 394-415.
- 1979. From discourse to syntax: grammar as a processing strategy, em **Syntax and Semantics** 12: 81-112.
- Gonçalves, V.G. 1994. Gramaticalização: conceitos, causas e processos, em **Caderno de Letras [Belo Horizonte]** I (1): 91-104.
- Haiman, J. Ed. 1985. **Iconicity in Syntax**. Amsterdam, John Benjamin.

- Heine, B. Hünnemeyer, B. and Claudi, U. 1991. **Grammaticalization: a conceptual framework**. Chicago, The University of Chicago Press.
- Hodge, C. T. 1970. The linguistic cycle, **Language in Society** 13: 1-7.
- Hook, P. E. 1991. The emergence of perfective aspect in Indo-Aryan languages, in Traugott-Heine Eds. 1991, vol. 2: 59-90.
- Hopper, P.J. 1988. Emergent grammar and the a priori grammar postulate, em Tannen, D. Ed. 1988, pp. 117-134.
- 1991. On some principles of grammaticization, em E.C. Traugott and B. Heine Eds. 1991, vol. 1: 17-36.
- and Traugott, E.C. 1993. **Grammaticalization**. Cambridge, CUP.
- Jakobson, R. 1948. L'Aspect phonologique et l'aspect grammatical du langage dans leurs interrelations, em **Éssais de linguistique générale**. Paris, Les Éditions de Minuit, 1963, pp. 161-175.
- Labov, W. 1995. **Principles of Linguistic Change**. Vol. I: Internal Factors. Oxford & Cambridge, Blackwell.
- Laytano, D. 1981. **O Linguajar do Gaúcho Brasileiro**. Porto Alegre, São Lourenço de Brindes.
- Lehmann, C. 1982a. Grammaticalization: synchronic variation and diachronic change, **Lingua e Stile** 20: 303-318.
- 1982b. **Thoughts on Grammaticalization. A programmatic sketch**. Köln, Arbeiten des Kölner Universalien-Projekts, vol. I, mimeo.
- Lehmann, w. and Malkiel, Y., Eds. 1968. **Directions for Historical Linguistics, a symposium**. Austin, University of Texas Press.
- Lichtenberk, F. 1991. On the gradualness of grammaticalization, em E.C. Traugott and B. Heine Eds. 1991, vol 1: 37-80.
- Lyons, J. 1968. **Semantics**. London, Longman, 2 vols.
- Marroquim, M. 1945. **A Língua da Nordeste**. São Paulo, Companhia Editora Nacional.
- Mattos e Silva, R. V. 1991. Caminhos de mudanças sintático-semânticas no Português Arcaico, **Cadernos de Estudos Lingüísticos** 20: 59-74.
- 1994. **Português Arcaico: Morfologia e sintaxe**. São Paulo, Contexto/UFBA.
- 1995. **Contradições no Ensino de Português**. São Paulo/Salvador, Contexto/EDUFBA.
- Maurer, T.H. 1952. **O Infinito Pessoal em Português**. São Paulo, Companhia Editora Nacional.
- 1959. **Gramática do Latim Vulgar**. Rio de Janeiro, Acadêmica.
- Meillet, A. 1912. L'évolution des formes grammaticales, em Meillet 1958:
- 1915. Le renouvellement des conjonctions, em Meillet 1958:
- 1958. **Linguistique historique et linguistique générale**. Paris,
- Mithun, M. 1984. The evolution of Noun incorporation, **Language**.
- Morris, C.W. 1938. **Foundations of the Theory of Signs**. Chicago, The University of Chicago Press.
- Neves, M.H.M. 1996. Hipotaxe e parataxe. Reflexões iniciais, ms. inédito.
- Oiticica, J. 1955. **Teoria da Correlação**. Rio de Janeiro, Simões.
- Parret, H. 1994. Prehistoire, structure et actualité de la théorie hjelmslevienne des cas, ms. inédito.
- Pontes, E. 1987. **O Tópico no Português do Brasil**. Campinas, Editora Pontes.
- Roberts, I. 1993. **Verb and Diachronic Syntax**. Dordrecht, Foris.
- e Kato, M. Orgs. 1993. **Português Brasileiro. Uma Viagem Diacrônica**. Campinas. Campinas, Editora da UNICAMP.
- Rodrigues, A. N. 1974. **O Dialeto Caipira na Região de Piracicaba**. São Paulo, Ática.

- Sapir, E. 1921. **A Linguagem. Introdução ao Estudo da Fala**, trad: port. de J. Mattoso Câmara Jr. Rio de Janeiro, Instituto Nacional do Livro, 1954.
- Shibatani, M. 1991. Grammaticization of topic into subject, em Traugott-Heine Eds. 1991, vol. 2: 93-134.
- Sweetser, E. **From Etymology to Pragmatics**. Metaphorical and cultural aspects of semantic structure. Cambridge, CUP.
- Tannen, D. Ed. 1988. **Linguistics in Context: connecting observation and understanding**. Norwood NJ, Ablex.
- Tarallo, F. 1983. **Relativization Strategies in Brazilian Portuguese**. University of Pennsylvania, Ph.D. Dissertation, inédito.
- Org. 1989. **Fotografias Sociolingüísticas**. Campinas, Pontes Editores/Editora da UNICAMP.
- 1991. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX, em I. Roberts e M. Kato Orgs. 1993, pp. 105.
- Thompson, S.A. and Mulac, A. 1991. A quantitative perspective on the grammaticization of epistemic parenthetical in English, em Traugott-Heine Eds. 1991, vol. 2: 313-330.
- Traugott, E. C. 1988. Pragmatic strengthening and grammaticalization, em **Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society**, 406-416.
- and Heine, B. Eds. 1991. **Approaches to Grammaticalization**. Amsterdam/Philadelphia, 2 vols.
- Ullmann, S. 1964. **Semântica**, trad. port. de J.A.Osório Mateus. Lisboa, Fundação Calouste Gulbenkian, 2a. ed., 1970.
- Vendryès, J. 1921. **Le Langage**. Introduction linguistique à l'histoire. Paris, Renaissance du Livre.

A relevância dos processos de pidginização e crioulização na formação da língua portuguesa no Brasil

Alan N. Baxter

La Trobe University

Dante Lucchesi

Universidade Federal da Bahia

Resumo

O artigo destaca a relevância do contato entre línguas na formação da realidade lingüística brasileira. Retoma o debate sobre a crioulização prévia do português do Brasil. Apresenta uma visão atualizada dos conceitos básicos da crioulística (crioulo, pidgin, semi-crioulo, transmissão lingüística irregular, etc.). Considerando os parâmetros sociolingüístico e estrutural, propõe a visão do processo de crioulização como um processo variável de transmissão lingüística irregular. Por fim, o artigo aborda aspectos da formação sócio-histórica da língua portuguesa no Brasil, como um processo polarizado, no qual interagem uma norma culta e uma norma vernácula. Os fatos observados no dialeto afro-brasileiro de Helvécia são, então, apresentados como vestígios do processo de transmissão lingüística irregular que caracteriza a constituição das vertentes populares da língua falada no Brasil.

0. Introdução

Em sua *História Universal da Infância*, Borges (1989: 1) nos conta que:

Em 1517, o padre Bartolomeu de Las Casas sentiu muita pena dos índios que se consumiam nos penosos infernos das minas de ouro das Antilhas e propôs ao Imperador Carlos V a importação de negros para que se consumissem nos penosos infernos das minas de ouro das Antilhas. A essa curiosa espécie de benfeitor devemos fatos incontáveis: os Blues de Handy, (...), o impetuoso filme Aleluya, o mestiço que matou Martín Fierro, (...), a cruz e a serpente no Haiti, o sangue das cabras degoladas do papaloi, a habanera de que nasceu o tango, o camdomblé.

Entre os fatos culturais que decorrem do processo histórico da escravidão de populações africanas na América, Borges poderia citar também a gramática original do saramacan, estranhamente constituída por um voca-

1 Este artigo é uma versão ampliada e revisada do texto que serviu de base à comunicação, do mesmo título, apresentada à XIX Reunião Nacional da Associação Brasileira de Antropologia, realizada em Niterói-RJ, no período de 27 a 31 de março de 1994.

bulário em boa parte de origem portuguesa, em terras do Suriname, onde o português não é falado; ou ainda el palenquero, curioso dialeto falado nos arredores de Cartagena, e que até bem pouco tempo era desconhecido, pois seus utentes o dissimulavam diante dos forasteiros.

A relação das línguas pidgins e crioulas na América com a escravidão e as situações de abominável exploração social e econômica que marcaram as suas origens trazem até hoje um certo estigma social aos falantes dessas línguas e a sua própria consideração como fatos culturais do Continente. A ideologia de uma transplantação incólume da cultura européia para o Novo Mundo fez com que no passado muitos estudiosos de renome rejeitassem qualquer influência dessas línguas "depauperadas e degeneradas" na formação do panorama lingüístico da América contemporânea.

Contudo, o recente movimento de resgate das reais condições de colonização da América provocou um crescente interesse pelos processos de contato entre línguas que marcaram a formação da realidade lingüística americana. Geralmente, as línguas e dialetos que mais nitidamente expressam processos prévios de pidginização e/ou criouliização são hoje faladas em regiões restritas e, não raro, em situações de alto isolamento. Esses falares constituem, pois, valiosos testemunhos do processo de constituição histórica das línguas faladas na América e um importante elemento para a compreensão da sua realidade atual.

Este texto pretende chamar a atenção para a importância do estudo desses fatos lingüísticos no território brasileiro. O estudo da fala das comunidades afro-brasileiras isoladas pode oferecer importantes indícios sobre a relevância dos processos de contato entre línguas na formação do português do Brasil, tema que tanta polêmica vem provocando há mais de um século. Mais do que isso, este estudo pode indicar importantes caminhos para uma mais adequada caracterização da realidade lingüística brasileira, não em relação à língua oficial, ou à língua dos círculos culturais de maior prestígio, mas à língua da maioria da população, daqueles cujos antepassados mais proximamente vivenciaram as situações em que o surgimento desse Novo Mundo se configurou: as situações de escravidão.

Tomaremos como exemplo o caso de uma comunidade situada nas proximidades de uma localidade denominada Helvécia, no extremo sul do Estado da Bahia; destacaremos aspectos da gramática desse dialeto que evidenciam processos de aquisição que possibilitam paralelos com a formação das línguas crioulas; e falaremos de como esses fatos podem contribuir para uma melhor caracterização da realidade lingüística brasileira. Mas antes, falaremos brevemente da discussão acerca do que se denominou "origens crioulas do português do Brasil".

1. As origens crioulas do português do Brasil

O estreito contato do português com línguas africanas, durante o período colonial, e possíveis processos de pidginização e criouliização dele decorrentes sempre estiveram na pauta daqueles que se ocuparam da caracterização sócio-histórica do português do Brasil (doravante PB). Já no final do século passado, F. A. Coelho (1880-6) reúne o PB aos crioulos de base portuguesa, com base nas semelhanças identificadas entre eles (em particular, a ausência de concordância no SN), definindo a todos como dialetos do português europeu (doravante PE).

No bojo do movimento de afirmação de uma língua brasileira, no início deste século, R. Mendonça (1933) e J. Raimundo (1933) chegam a atribuir à influência das línguas africanas a motivação de todas as características que distinguem o PB do PE. Na reação que se segue, S. da Silva Neto (1986), G. Chaves de Melo (1981) e S. Elia (1979), embora admitam a formação de crioulos e semicrioulos decorrentes do aprendizado imperfeito do português por falantes africanos, negam qualquer influência destes na constituição do PB. Tal posição se justifica pela imagem de unidade e conservadorismo do PB, por que militavam esses autores, decorrente de uma visão de superioridade cultural do colonizador branco, bem como pela fragilidade de sua concepção acerca dos crioulos e semicrioulos, vistos como uma simplificação e/ou deturpação do português.

Câmara Jr. (1972 e 1985) também negará um papel relevante aos crioulos falados por escravos africanos no Brasil, em função do modelo teórico ao qual se filiava, o estruturalismo, que encerrava na lógica interna do sistema lingüístico os fatores determinantes do desenvolvimento histórico da língua. Desse modo, segundo Câmara Jr., a influência das línguas africanas na constituição do PB se resumiria à aceleração de tendências prefiguradas no sistema lingüístico do português.

O advento da sociolingüística permitirá a retomada do debate em torno da questão, em outros termos. O ponto de partida é o quadro atual de profundas diferenças entre o português popular do Brasil (doravante PPB) e o português padrão. Muitos dos traços que caracterizam o PPB são típicos de línguas crioulas: (i) preferência pela marcação do plural apenas no primeiro elemento do SN; (ii) drástica redução na flexão número-pessoal do verbo; (iii) dupla negação (como no crioulo português de São Tomé e no palanquero) - sendo que esses traços de tipo crioulo do PPB podem ser ainda mais radicais nos dialetos rurais (cf. Jeroslow, 1975).

O debate atual então envolve duas posições: uma que vê essa situação do PPB apenas como o resultado de mudanças lingüísticas internas, já

que "o impulso motor do desenvolvimento do português do Brasil veio já embutido na deriva da língua de Portugal" (Naro e Scherre, 1993); e uma segunda, que destaca a importância do(s) crioulo(s) português(es) falado(s) pelos escravos africanos na compreensão das origens do PPB.

A hipótese da relevância da crioulição prévia na formação do PPB, postulada por G. Guy (1981 e 1989) e J. Holm (1987), encontra um forte amparo nos dados sócio-históricos. O Brasil, por si só, absorveu aproximadamente 40% dos escravos levados para as Américas, ou seja, algo em torno de 3,6 milhões de indivíduos; e a abolição só ocorreu aqui tardiamente, em 1888. Em determinadas regiões, houve concentrações demográficas afro-brasileiras muito elevadas; sendo que, por algum período, a população afro-brasileira superou notavelmente a população européia. Devendo-se acrescentar ainda que, no Brasil colonial, a maioria da população se concentrava no campo, e o campo era o destino da maioria dos escravos.

Desse modo, como destaca G. Guy (1989), parece razoável supor que a língua portuguesa no Brasil pode ter sofrido mudanças do tipo das que afetaram as línguas européias em outros países do Novo Mundo, cuja história também é marcada pelo processo da escravidão de populações africanas; países onde hoje crioulos e post-crioulos são largamente falados, como, por exemplo: a Jamaica (crioulo inglês), o Haiti (crioulo francês) e Curaçao (crioulo espanhol).

Porém, a hipótese da crioulição prévia do PPB é sensivelmente prejudicada pela falta de dados lingüísticos específicos, a ela concernentes. Os defensores da proposta de evolução lingüística interna argumentam que, se variedades de crioulos portugueses existiram no Brasil, elas teriam logo desaparecido, sem que pudessem interferir decisivamente na formação do PB. Contudo, em outros países da América, onde a escravidão existiu e a língua hoje predominante não é o crioulo, línguas crioulas permaneceram em comunidades isoladas, especialmente em áreas em que houve uma considerável concentração de africanos. Assim, na Carolina do Sul, nos EUA, o gullah, uma variedade criouliada do inglês sobreviveu; e, em El Palenque de São Basílio, próximo a Cartagena, na Colômbia, o palenquero, um crioulo espanhol, ~~continua a ser falado em uma comunidade remanescente de um antigo quilombo.~~ O fato de o palenquero só recentemente ter sido reconhecido como um crioulo levanta a possibilidade da sobrevivência de outros crioulos ou post-crioulos em outras comunidades afro-americanas isoladas (cf. Granda *apud* Holm, 1989: 310). Investigar essa possibilidade constitui um dos desafios para a pesquisa sociolingüística no Brasil.

Por outro lado, o debate acerca da relevância de processos anteriores de crioulição na formação da língua falada no Brasil envolve o modo como

o investigador concebe o seu objeto de estudo, num debate conceitual acerca do que se entende por *língua natural*, *língua crioula*, *pidgin*, etc, para o qual concorrem modelos teóricos divergentes e inclinações ideológicas dos mais variadas matizes. Assim, antes de prosseguir, faremos uma breve exposição acerca do conceito de *língua crioula*.

2. O conceito de língua crioula

A definição de língua crioula tem se fundamentado na combinação de dois parâmetros. De um lado, considera-se o contexto sócio-econômico e político, ou seja o conjunto de relações sociais em que a língua emergiu; esse seria o parâmetro que podemos chamar de *sociolingüístico*. De outro lado, consideram-se as semelhanças na estrutura lingüística dessas línguas que se formaram dentro de uma situação excepcional, é o que podemos chamar de parâmetro *estrutural* ou *lingüístico*. Falaremos inicialmente das características sociolingüísticas para em seguida apresentar brevemente algumas das particularidades estruturais atribuídas às línguas crioulas.

2.1. Características Sociolingüísticas

Uma língua crioula é uma língua que nasce em circunstâncias sociolingüísticas especiais que conduzem à aquisição de uma primeira língua, com base em um modelo defectivo de segunda língua. É uma língua falada por uma comunidade cujos antepassados parcialmente perderam os seus traços sociolingüísticos e culturais originais, devido, na maioria dos casos estudados, à colonização européia e à escravatura.

Em muitos casos, a primeira geração de escravos de uma colônia enfrentou uma situação que levou ao uso de uma segunda língua muito rudimentar, fragmentada e variável. Essa segunda língua, não obstante ter como alvo a língua dos dominadores, era fortemente influenciada pelas línguas maternas de seus utentes.

O acesso às línguas européias, sobretudo nas micro-sociedades dos engenhos, costumava ser limitado em função das condições sociais da escravatura. Contudo, os escravos não deixavam de adquirir, através do contato, fragmentos da língua de superstrato. Na maioria dos casos, é provável que essa língua diluída não fosse mais que um jargão, uma parcela de itens lexicais adaptados às estruturas das línguas maternas dos falantes; ou seja, um leque de soluções individuais. Em outros casos, é possível que essa segunda língua rudimentar se tenha cristalizado, constituindo uma língua *pidgin*. É evidente que um fator que teria influenciado o desenvolvimento dessa segunda língua teria sido a homogeneidade, ou não, das línguas

maternas faladas pelos escravos. Nos casos em que os africanos não tinham uma língua em comum, parece mais provável a formação do *pidgin*.

As crianças que nasciam nessa situação eram expostas às línguas nativas de seus pais e também àquela segunda língua baseada na língua européia. Mas, por vários motivos, a segunda língua era mais viável socialmente, e, portanto, acabava por se tornar a sua língua primária.

Contudo, como observamos, essa segunda língua, baseada na língua de superstrato, teria sido defectiva em vários sentidos. E, apesar de contarem apenas com modelos lingüísticos altamente variados, incompletos e possivelmente defeituosos, os novos falantes acabavam por serem capazes de, com essa base, procederem à elaboração formal e funcional de sua língua nativa, o *crioulo* (cf. Holm, 1988: 6-7).

A perspectiva que acabamos de apresentar explicaria, em termos gerais, a formação das línguas crioulas mais conhecidas, como, por exemplo, o crioulo de base francesa do Haiti, os crioulos de base portuguesa de Cabo Verde e São Tomé, e o crioulo de base inglesa da Jamaica.

2.2. Semelhanças estruturais

A partir da década de 60, os lingüistas começaram a insistir no fato de as línguas crioulas apresentarem fortes semelhanças estruturais, a despeito de que línguas estivessem envolvidas na sua formação. Além disso, descobriu-se que essas semelhanças estruturais não são categóricas, mas propensões que dependem de um conjunto de variáveis sociolingüísticas que influenciaram a constituição da língua crioula, tais como:

- (i) as proporções relativas de europeus e não-europeus, e, portanto, a facilidade de acesso à língua do grupo dominante;
- (ii) a homogeneidade/heterogeneidade do grupo dominado, e, portanto, a possibilidade de manutenção das línguas desse grupo.

Assim, pode-se falar de uma crioulização variável que abrange um leque de possibilidades, da mais radical a menos radical.

Ao mesmo tempo, observaram-se fortes paralelos estruturais e de desenvolvimento entre as línguas *pidgins* e crioulas, de um lado, e o processo de aquisição das línguas (primeira e segunda), por outro.

No fim da década de 70, e sobretudo na década de 80, o rumo da discussão da gênese e da forma das línguas crioulas mudou para tratar o processo variável da crioulização como um caso especial de aquisição/criação de primeira língua e para buscar identificar os fatores determinantes desse processo. Mais ainda, os lingüistas começaram a associar as propen-

sões estruturais paralelas das línguas crioulas às tendências lingüísticas universais (observadas nas fases iniciais da aquisição de qualquer língua, crioula ou não, mas não encontradas nas línguas de contato) e, em certos casos, às estruturas específicas provindas das línguas dos grupos dominados. Como exemplos da primeira categoria, a dos universais, teríamos:

- (i) o emprego do mesmo verbo para exprimir 'existência' e 'posse';
- (ii) a ausência da estrutura passiva e a preferência pela diátese lexical (i. é, sujeito paciente + verbo transitivo direto, expressando a ação sofrida pelo sujeito);
- (iii) o emprego de uma preposição multifuncional que exprime 'lugar' e 'direção';
- (iv) emprego da mesma ordem de constituintes nas orações afirmativas e interrogativas.

No caso da segunda categoria - denominada influência de substrato -, teríamos, como exemplo, as estruturas do crioulo de base portuguesa de São Tomé que são diretamente atribuídas ao fato de que os africanos que desenvolveram o crioulo de São Tomé eram falantes de línguas kwa e bantu, que possuem estruturas/formas paralelas, como:

- (i) O emprego de estruturas negativas duplas, em que uma partícula negativa é colocada antes do verbo, e a outra, depois do sintagma verbal: "*ne na tava sici fosa d'ome gletu fa*" - 'eles não se sentiam suficientemente fortes' (apud Holm, 1989);
- (ii) palatalização de /t/ e /d/ antes da vogal fechada /i/: *dwetsi* 'doente'; *dža* 'dia';
- (iii) troca de /l/ e /r/ (que são alófonos nas línguas bantu): *óla* 'hora';
- (iv) o -lh- português, raro nas línguas africanas acima referidas, é [y] em crioulo;
- (v) preposição *ni*, por *em*;
- (vi) regularização do padrão silábico CV: *dotolo* 'doutor'.

2.3. Aquisição de um modelo diluído

A questão que se coloca então é: *como e por que* as estruturas universais e das línguas de substrato entram nas línguas crioulas?

Quando se considera que as populações de escravos eram formadas principalmente por adultos e eram constantemente renovadas pela chegada de novas levas, e quando se pensa que a maioria dos escravos não estava em contato direto com os senhores, é evidente que a disponibilidade de um bom modelo da língua de superstrato era seriamente limitada. Assim, teria faltado um modelo nítido para a primeira geração de falantes do crioulo. O modelo do seu processo de aquisição teria sido de fato uma língua segunda

(L2) que não passava de uma versão defectiva da língua de superstrato, ou seja, uma L2 adquirida unicamente através de um contato precário. Além disso, o modelo de língua transmitido teria sido constantemente enfraquecido e diluído, devido às renovações sucessivas do contingente de escravos africanos.

Derek Bickerton afirma que, quando se contempla o léxico de uma língua crioula, observa-se que o que sobrevive da língua dominante original é basicamente o léxico das palavras referenciais. Os morfemas gramaticais - preposições, artigos, desinências verbais, de gênero, número e caso -, que são exatamente os itens mais difíceis de se adquirir, geralmente não sobrevivem ao processo de transmissão irregular. O Quadro 1 (adaptado de Bickerton, 1988) mostra a gama dos itens gramaticais da língua dominante que geralmente são atingidos pelo processo de diluição:

Quadro 1

Conjunto A:

- a. artigos
- b. indicadores de tempo, modo e aspecto
- c. palavras interrogativas
- d. o indicador de plural
- e. pronomes que indicam pessoa e número
- f. indicadores de caso oblíquo
- g. preposições locativas
- h. partículas de relativização
- i. reflexivos e recíprocos

Conjunto B:

- j. concordância de gênero (não se mantém)
- l. concordância de número (não se mantém)
- m. morfologia verbal flexional (quase não se mantém)
- n. morfologia derivacional (quase não se mantém)
- o. caso e gênero pronominais (quase não se mantém)
- p. a maioria das preposições

Observe-se que, embora todas as áreas da gramática sejam afetadas, Bickerton chama a atenção para o fato dos itens do grupo B serem raramente mantidos nas línguas crioulas, ao passo que os itens do grupo A, se se perdem durante o processo de diluição, são reconstituídos de maneira original. Bickerton propõe ainda que essa reconstituição é orientada por fatores universais, através do que ele chama de *Bioprograma*, um dispositivo inato para a aquisição/ estruturação da língua materna (cf., e. g., Bickerton, 1981, 1984 e 1988).

A idéia de reestruturação que caracteriza o surgimento de uma língua crioula implica (qualquer que seja a explicação teórica) que, por exemplo, se se perde o sistema de artigo original da língua de superstrato, esse sistema é reconstituído dentro de uma nova estrutura formal e funcional na língua crioula. Formas do pronome demonstrativo e do numeral 'um', não raro, podem ser utilizadas para desempenhar a função de artigo definido e artigo indefinido, respectivamente, formando um novo sistema de artigo que marca o grau de definitude dos SN's dentro de uma lógica funcional muitas vezes distinta da lógica observada na língua de superstrato (cf., e. g., Lucchesi, 1993). Da mesma forma, se se perde a morfologia de número dos substantivos, esta será reconstituída empregando-se, por exemplo, um quantificador ou um pronome no plural que exerce a função de indicador de plural, como no crioulo de São Tomé, em que o pronome 'eles/elas' é utilizado como pluralizador. A perda da morfologia passiva dá origem a estruturas do tipo 'a árvore plantou', que substitui a estrutura original da língua de superstrato, geralmente do tipo 'a árvore foi plantada'. A perda das preposições que exprimem destino e lugar conduz geralmente à reconstituição através de uma única preposição com uma função locativa genérica, que nos crioulos de base portuguesa costuma ser derivada da forma contrata da preposição *em* com o artigo (*no, na > ni*). A perda dos reflexivos dará lugar a uma reconstituição analítica que muitas vezes incorpora uma palavra que significa 'igual' ou 'corpo'.

Outros investigadores contemporâneos, como, por exemplo, Thomason e Kaufman (1989), propõem que a homogeneidade no substrato pode facilitar a introdução no crioulo de estruturas específicas do substrato para resolver casos de opacidade resultantes do processo de diluição.

2.4. O processo de crioulição: uma definição atualizada

A partir do que foi dito até agora sobre a diluição do modelo-estímulo e do efeito disso sobre a transmissão da língua, podemos apresentar uma definição do processo de crioulição consoante com as teorias atuais da crioulistica:

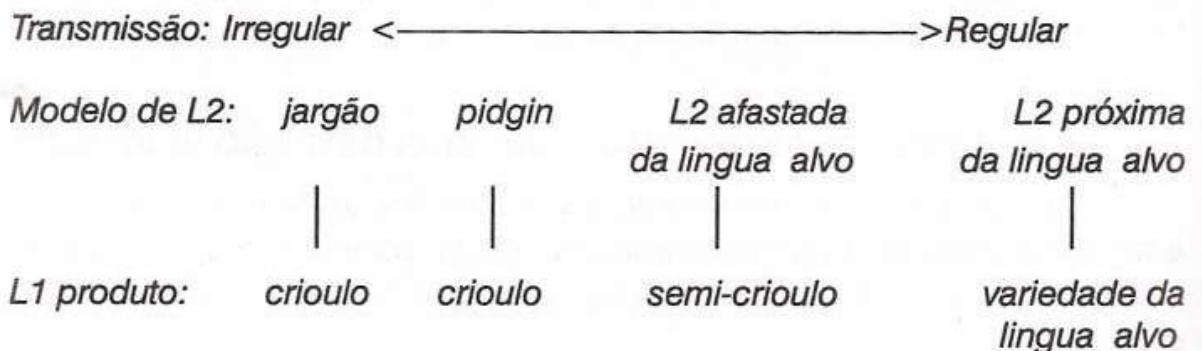
Um processo de transmissão irregular de L2 para L1 em que a L2 foi alterada devido a problemas de acesso à língua alvo (isto é, a língua do grupo dominante) e, possivelmente, à influência das línguas maternas dos falantes desta L2. Nessas circunstâncias, no desenvolvimento, na aquisição/criação da nova L1 (a língua crioula em potencial), acontecem inovações orientadas por universais e pelas outras línguas maternas presentes. As inovações preenchem as lacunas ou opacidades causadas pela diluição do modelo para aquisição. Tal processo é variável.

A idéia de que a crioulação é variável foi proposta de maneira explícita por Baker (1982) e Bickerton (1984). Ambos os autores observaram que uma transmissão crioulação que se serve de um modelo relativamente forte resultará em um produto que se aproxima mais da língua alvo original. Assim, de um lado, na Ilha de Reunião, colônia francesa no Oceano Índico, a proporção de europeus para escravos foi relativamente alta durante as primeiras cinco décadas da colônia. Como resultado, o crioulo de base francesa que se desenvolveu parece um dialeto do francês com estruturas crioulas. De outro lado, a crioulação que ocorre em sociedades quilombolas, como nas comunidades saramacan, do Suriname, ou a comunidade de El Palenque de San Basilio, na Colômbia, em que os escravos se retiraram (através da fuga) do contato com a língua de superstrato, produz resultados radicais: línguas novas estruturalmente bem afastadas da língua europeia.

Nesse sentido, é melhor contemplar a crioulação como um processo que pertence a um contínuo de gêneros de transmissão de L2 para L1, partindo do mais irregular para o mais regular, a depender da qualidade da L2 que serviu de base para o surgimento da língua crioula, e, portanto, das circunstâncias sócio-históricas da transmissão. O Quadro 2 representa esse contínuo:

Quadro 2.

Transmissão/nativização com base em diversos modelos de L2:



3. A crioulização no processo de formação do português do Brasil

Quando se pensa no contexto sócio-histórico do Brasil colonial, é evidente que a aquisição do português como L2 pelos segmentos africanos e indígenas e a transmissão irregular dessas formas precárias de L2 para L1 para os descendentes desses segmentos, em circunstâncias pouco favoráveis, deve ter desempenhado um papel importante na formação de variedades do português no Brasil. Além disso, o segmento afro-brasileiro deve ter sofrido uma influência constante das variedades L2 do português faladas pelos recém chegados. O Quadro 3 mostra, de uma maneira simplificada e incompleta, algumas das variedades de línguas (L1 e L2) presentes e potencialmente em contato no Brasil colonial:

Quadro 3.

Brasil colonial:

algumas variedades de língua presentes e em contato:

Língua indígena	Língua africana	Língua europeia não portuguesa	português metropolitano	português regional
L2 Geral	L2 africana coiné	L2 português dos europeus		
L2 português dos índios	L2 português dos escravos	L1 português dos descendentes dos europeus não lusos		
L1 português de descendentes dos índios	L1 português de descendentes dos africanos			

Quando se consideram as implicações desse quadro, não parece ariscado teorizar que os processos de diluição e transmissão irregular (ou seja, de crioulização, na acepção atual do termo) teriam desempenhado um papel no desenvolvimento do português do Brasil, direta e indiretamente.

Com efeito, existe no português do Brasil, e principalmente nos dialetos rurais, uma extensa gama de propensões estruturais que, aos olhos do lingüista especializado no estudo das mudanças causadas pelo contato entre línguas, apontariam para um processo de transmissão irregular. Algumas dessas características são apresentadas no Quadro 4.

Quadro 4.

Dialetos rurais brasileiros:
algumas propensões morfossintáticas
que apontam para um processo de transmissão irregular:

1. Sintagma Nominal

- a. variação na concordância de número
- b. oração relativa introduzida com *que* complementizador e com cópia pronominal

2. Sintagma Verbal

- a. morfologia de pessoa e número drasticamente reduzida, o que conduz a:
 - (i) significativo enfraquecimento da concordância verbo-nominal;
 - (ii) maior presença do sujeito.
- b. morfologia de tempo e modo:
 - (i) presença variável, e bastante reduzida, do subjuntivo;
 - (ii) ausência do condicional;
 - (iii) futuro perifrástico em vez do futuro sintético
- c. baixa incidência da estrutura passiva, ocorrendo: sujeito (paciente) + verbo transitivo (ação)
- d. o pronome pessoal mantém a forma de sujeito em função de objeto
- e. verbo *ter* indicando posse e existência
- f. baixa incidência de pronomes reflexivos
- g. negação dupla descontínua (preverbal e no final do SV)

Muitas dessas propensões são compartilhadas (embora de maneira mais radical) pelas línguas crioulas, especialmente pelos crioulos de base portuguesa de Cabo Verde, Guiné Bissau e do Golfo da Guiné (das Ilhas de São Tomé, Príncipe e Ano Bom) (cf. Baxter, 1987 e Holm, 1992). Embora não possamos descer aqui a pormenores, é evidente que muitas dessas propensões encontram paralelos no esquema de Bickerton, especialmente no conjunto B (Quadro 1, acima). Além disso, cabe acrescentar que as propensões do Quadro 4 podem ser muito mais radicais em determinados dialetos, principalmente nos dialetos afro-brasileiros isolados.

Considerando que os parâmetros sócio-histórico e lingüístico se conjugam de forma sedutora, não seria o caso de perguntar se a investigação

das vertentes em que a transmissão irregular teria sido mais factível - os segmentos afro-brasileiro e ameríndio - não poderia esclarecer sobre a relevância desse processo na conformação da língua falada no Brasil? Nas seções seguintes, serão considerados alguns aspectos da fala de uma variedade afro-brasileira do português, a partir da perspectiva da transmissão irregular e da aquisição da língua. As finalidades da discussão serão: (i) buscar lançar alguma luz sobre a forma da fala das comunidades afro-brasileiras no passado; (ii) tentar explicar determinados aspectos morfossintáticos dos dialetos rurais atuais.

4. O caso do dialeto afro-brasileiro de Helvécia-Ba

Uma das comunidades estudadas pelo Projeto *Vestígios de dialetos crioulos em comunidades afro-brasileiras isoladas*² fornece dados significativos para o raciocínio que aqui se desenvolve. Trata-se do dialeto afro-brasileiro de Helvécia, no extremo sul do estado da Bahia (veja-se Ferreira, 1969; Baxter, 1992; Baxter e Lucchesi, 1993; e Megenney, 1993). O dialeto é falado por descendentes de iorubás e geges, escravos da antiga Colônia Leopoldina, estabelecida nas primeiras décadas do século XIX por suíços, alemães e franceses.

No dialeto de Helvécia, observam-se traços que apresentam paralelos significativos com crioulos de base portuguesa e reforçam a hipótese de sua criouliização prévia, como se pode observar no Quadro 5, abaixo:

Quadro 5.

Características crioulizantes do dialeto rural de Helvécia:

- No sistema verbal:

- (i) uso de formas do presente para indicar ações e estados situados no passado: 1) "Meu pai é de cativoiro" 'meu pai era escravo'; 2) "Eu pega Mário e mandou ele cortá [lenha]" 'eu peguei Mário e mandei-o cortar lenha';
- (ii) uso de formas finitas em contexto de formas do infinitivo e vice-versa: 1) "Non sê vai lá" 'não sei ir lá'; 2) "Ficou certo de vem um de lá" 'ficou certo de vir um de lá';
- (iii) variação na marca da 1ª pessoa do singular, o que aponta para um sistema anterior desprovido de flexão de número e pessoa: 1) "tomei aquilo, acabô. Acabô... nunca mais eu sintiu." 2) "Eu num cunhece esses povo não."

2 Este projeto é sediado no Departamento de Letras Vernáculas da Universidade Federal da Bahia, e se realiza com a coordenação do Prof. Alan Baxter, da Universidade de La Trobe, de Melbourne - Austrália.

- No nível morfossintático, ainda teríamos:

- (i) uso reduzido do artigo: 1) "eu sô fia de lugá", 'eu sou filha do (deste) lugar'.
- (ii) variação na concordância de gênero no SN: 1) "o meu sobrinha" 2) "umas duas arquerim de terra..."
- (iii) uso da diátese lexical, ao invés de estruturas passivas: 1) "e bicho encontrava lá", 'e bicho era encontrado lá'.
- (iv) regência verbal: 1) "Perguntei o Pedro, ele disse..."; 2) DOC: "E bebê, o sinhô bebe?" INF: "Meu amigo, eu num vô dizê o sinhô que não.", 3) DOC: "E bicho, num gosta de eucalipe, né?" INF: "Diz que não." DOC: "Num gosta não." INF: "Gosta mata virge."; 4) "Eu não queria ficá o jeito que ela queria" 'eu não queria ficar do jeito que ela queria que eu ficasse'.
- (v) Supressão da preposição: 1) "folha mandioca", 'folha de mandioca'; 2) "Eu tenho direito distraí um pouco", 'Eu tenho direito de me distrair um pouco';
- (vi) ausência do verbo copulativo: 1) "Esse aí neto de Casimiro", 'Esse aí é neto de Casimiro'.
- (vii) ausência de concordância de gênero entre o sujeito e o predicativo: 1) "e inté hoje tá istragado minha vista." 2) "A festa aí é bunito." 3) "ela ficô bom." 4) "A minina tá mei duentado.";
- (viii) orações encaixadas sem complementizador: 1) "eu teve cunvessano... foi com quem? Acho que foi cum Pedro mehmo, ele disse a irmã dele veio do Rio.", 'eu estive conversando... foi com quem? Acho que foi com o Pedro mesmo, ele me disse que a irmã dele veio do Rio'; 2) "A pessoa num subé cuzinhá uma carne tá ruim." 'A pessoa que não souber cozinhar uma carne está ruim'; 3) "Acho que seca mehmo porque teve lagoa aí secô." 'Acho que seca mesmo, porque houve lagoa aí que secou'.
- (ix) dupla negação marcada no sujeito: 1) "Ninhum discarado num tá trabaiano não." 3) "Esse tempo lá, ninguem num tava ligano de negóci de falá em fazê procuração."

A maioria dos traços aqui apresentados individualizam o dialeto de Helvécia no panorama dos dialetos rurais brasileiros. Contudo não seria o caso de retirar o dialeto de Helvécia do processo global de formação dos dialetos populares do português do Brasil; pode-se situar essa diferença entre o dialeto de Helvécia e os demais dialetos dentro de uma perspectiva cronológica, estabelecendo-se, a partir de uma observação sincrônica diatópica, um contínuo diacrônico que reúne o dialeto de Helvécia aos demais dialetos rurais, e até mesmo aos dialetos urbanos populares.

Para entender esse quadro atual, é preciso compreender como os padrões lingüísticos predominantes nos segmentos sociais dominados econômica e politicamente vêm interagindo com os padrões dos segmentos sociais dos dominadores ao longo da constituição sócio-histórica do portu-

guês do Brasil. Na seção seguinte, levantaremos algumas hipóteses de pesquisa sobre esse tema.

5. Elementos para uma caracterização sócio-histórica da língua falada no Brasil: a contribuição dos imigrantes e dos segmentos afro-brasileiro e indígena

Como bem observou Celso Cunha (1985), durante os três primeiros séculos da colonização do Brasil, a língua seguiu na zona rural a sua "deriva diversificadora", na medida em que os centros urbanos, em que predominavam as variedades mais próximas da língua padrão, "pouca ou nenhuma influência" exerciam sobre as regiões afastadas do interior. Os dados aqui apresentados buscam dar sustentação à idéia de que, nesse processo de transformação por que a língua passou nas regiões do interior do país, o contato entre o português e as línguas africanas e indígenas teve um papel relevante. As condições sócio-históricas em que se deu esse contato apontam para a ocorrência de transmissão lingüística irregular. Esse processo teria provocado reduções muitas vezes drásticas em certas áreas da gramática da língua falada por esses segmentos da população; podendo ser definido, portanto, como sendo, pelo menos, um processo de crioulização leve, de acordo com a visão atual de contínuo de crioulização (cf. supra).

5.1. A distinção entre norma culta e norma vernácula

Esse processo crioulizante que caracteriza a formação dos padrões de fala dos segmentos rurais e populares opõe-se ao processo de desenvolvimento da língua entre os segmentos das classes médias altas urbanas, que, dentro do espaço institucional, conservariam variedades mais próximas da língua padrão. Desse modo, uma distinção clara entre o que se pode chamar, respectivamente, de *norma vernácula* e *norma culta* é crucial para que se possa compreender as vertentes de constituição da língua falada no Brasil³.

Recentemente, Tarallo (1993) rebateu a hipótese da crioulização prévia do português do Brasil, afirmando que as mudanças por ele observadas no PB indicavam um afastamento em relação à norma padrão, e não a aproximação da "língua alvo", que caracteriza o processo de descrioulização. Atentando para a base empírica dos estudos utilizados por Tarallo - textos escritos (basicamente cartas e peças teatrais) para as fases anteriores da

3 Para uma maior fundamentação (dentro do quadro teórico da Sociolingüística Variacionista) da distinção entre norma culta e norma vernácula na formação da realidade lingüística brasileira, veja-se Lucchesi (1994).

língua, e materiais do *Projeto de Estudo da Norma Culta* (Projeto NURC), para o presente -, pode-se notar que as conclusões a que chegou esse eminente lingüista são válidas para a norma culta, e não para a língua falada no Brasil, em seu conjunto.

5.2. Influências de baixo para cima e de cima para baixo na configuração da língua neste século

Tarallo situa cronologicamente as mudanças por ele analisadas entre os finais do século XIX e as primeiras décadas deste século, período em que se inicia o processo de industrialização e o crescimento dos centros urbanos, com o progressivo deslocamento demográfico do campo para as cidades. Essas mudanças que se operaram na norma culta conduzindo ao seu afastamento da norma padrão poderiam ser enquadradas dentro do que Labov (1972, 1974 e 1982) definiu como mudança de baixo para cima, sendo, portanto, o resultado de uma influência da norma vernácula sobre a norma culta.

Um fator que pode ter sido decisivo para a implementação dessas mudanças seria o significativo contingente de imigrantes europeus e asiáticos que entraram no país nesse período. Uma grande parcela desses imigrantes se dirigiu inicialmente para o campo, deslocando-se posteriormente para as cidades. No plano sócio-econômico, esse contingente de imigrantes acusa rapidamente uma ascensão na escala social. Ora, tendo entrado na base da pirâmide social, o primeiro contato que esses indivíduos tiveram com a língua portuguesa não foi com a norma culta, e sim, com a norma vernácula. Ao subirem na escala social esses indivíduos podem ter transportado padrões lingüísticos, mesmo que esbatidos, que haviam incorporado através do seu contato inicial com os falantes das classes mais baixas.

Ao lado disso, teria também contribuído para a difusão dos padrões lingüísticos na base da pirâmide social o deslocamento de contingentes de escravos que se verifica ao longo do século XIX. As restrições e a proibição da importação de escravos da África teria provocado, principalmente a partir de 1860, um crescente comércio interno de escravos, observando-se um expressivo deslocamento de contingentes de escravos para o sudeste do país que vivia na altura a pujança da lavoura do café.

Por outro lado, em um sentido contrário, pode-se pensar que a norma vernácula, ou seja os padrões lingüísticos que caracterizam a fala dos segmentos da base da pirâmide social, apresentaria até meados deste século profundas alterações e reduções em áreas significativas da gramática, em decorrência do processo de transmissão irregular acima referido. Porém, o

crescimento da influência sócio-econômica e cultural dos grandes centros urbanos sobre as demais regiões do país, devido à constituição da malha rodoviária e da ação dos meios de comunicação de massa, teve por consequência natural a propagação dos padrões lingüísticos da norma urbana culta, neutralizando as marcas mais radicais da gramática dos segmentos vernaculares.

Portanto, é razoável propor, para a análise do desenvolvimento sócio-histórico do português do Brasil neste século, uma interação dos seus dois pólos constituintes, que se atualizaria nas mudanças que se propagaram de baixo para cima e de cima para baixo na estrutura social. Assim, a observação de que a norma culta se estaria afastando da norma padrão não pode ser tomada, como pretendeu Tarallo (1993), como uma contra-evidência para a hipótese da crioulização prévia do português do Brasil.

6. Conclusão

Com base no raciocínio desenvolvido na seção anterior, pode-se contemplar os dados extraídos do dialeto de Helvécia e da tradição oral de outras comunidades brasileiras da seguinte maneira.

Por alguma razão, o dialeto de Helvécia teria se conservado em um estágio anterior em que as marcas crioulizantes de um processo de transmissão irregular pretérito ainda estão presentes. Em outras comunidades afro-brasileiras (em que se podem encontrar, ainda hoje, referências a uma fala "embolada", ou mesmo incompreensível, que teria existido no passado), nos dialetos rurais e principalmente nos dialetos urbanos vernaculares, essas marcas já teriam desaparecido, devido ao contato e à influência dos padrões lingüísticos da norma culta, que se processaram de maneira crescente ao longo deste século. Portanto, o contínuo que se pode observar sincronicamente entre esses dialetos refletiria o processo diacrônico de constituição dos padrões lingüísticos observados nos segmentos da base da pirâmide social no Brasil. Desse modo, o que se observa hoje em Helvécia pode refletir o quadro lingüístico que caracterizou, até pelo menos o início deste século, as variedades vernáculas do português do Brasil, principalmente na área rural.

Por fim, cabe salientar que as proposições contidas neste texto são decorrentes de uma pesquisa que ainda se encontra em desenvolvimento, devendo ser tomadas na discussão como hipóteses de trabalho, e ser, portanto, avaliadas mais por sua coerência argumentativa do que por sua fundamentação empírica. Contudo, acreditamos que a análise aqui desenvolvida apresenta elementos relevantes para a consideração do contato entre línguas no estudo da história sociolingüística do Brasil.

Abstract

This paper emphasizes the major role played by language contact in the formation of Brazilian linguistic reality. It also resumes the discussion on early creolization of Brazilian Portuguese. It introduces the current view of the fundamental concepts of creole studies (creole, pidgin, semi-creole, irregular linguistic transmission etc). Taking structural and socio-linguistic parameters into account, viewing the creolization as a changing process of irregular linguistic transmission is proposed. Eventually, aspects of the social and historical formation of Brazilian Portuguese, as a polarized process in which one can notice the interaction between the standard norm and the vernacular norm, are dealt with. The linguistic facts which are observed in the Helvetic Afro-Brazilian dialect are presented as traces of the irregular linguistic transmission which characterizes the popular variations of the language spoken in Brazil.

6. Bibliografia

- BAXTER, Alan. 1987. **Creole-like traces in rural Brazilian Portuguese dialects**. Manuscrito. Department of Spanish, La Trobe University, Melbourne. 47 pp.
- _____. 1992. "A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a crioulação prévia: um exemplo do Estado da Bahia". In.: d'Andrade, E. e Kihm, A. (orgs.). **Actas do Coloquio sobre Crioulos de Base Lexical Portuguesa**. Lisboa: Colibri. pp. 7-35.
- BAXTER, Alan e LUCCHESI, Dante. 1993. Processos de descrioulização no sistema verbal de um dialeto rural brasileiro, **Papiá**, 2: 59-71.
- BICKERTON, Derek. 1981. **Roots of Language**. Ann Arbor: Karoma.
- BICKERTON, Derek. 1984. The Language Bioprogram Hypothesis. **Behavioural and Brain Sciences**, 7: 173-188.
- _____. 1988. Creole Languages and Bioprogram. In.: F. J. Newmeyer (ed.). **Linguistics: the Cambridge Survey**. Cambridge: Cambridge University Press.
- BORGES, Jorge Luis. 1989. **História Universal da Infância**. Trad.: Flávio José Cardozo. 5 ed. São Paulo: Globo.
- CÂMARA JR., Joaquim Mattoso. 1972. Línguas européias de ultramar: o português do Brasil. In: _____. **Dispersos**. Rio de Janeiro: Fundação Getúlio Vargas, pp. 71-87.
- CAMARA JR., Joaquim Mattoso. 1985. **História e estrutura da língua portuguesa**. Rio de Janeiro: Padrão.
- COELHO, F. A. 1880-86. Os dialetos românicos ou neolatinos na África, Ásia e América. In.: J. MORAIS BARBOSA (org.). **Estudos Lingüísticos Crioulos**. Lisboa: Academia Internacional de Cultura Portuguesa, 1967.
- CUNHA, Celso. 1985. **A Questão da Norma Culta Brasileira**. Rio de Janeiro: Tempo Brasileiro.
- ELIA, Sílvio. 1979. **A unidade lingüística do Brasil**. Rio de Janeiro: Padrão.
- FERREIRA, Carlota da Silveira. 1969. Remanescentes de um falar crioulo brasileiro. **Revista Lusitana**, 5 (1984-5):21-34.
- GUY, Gregory Riordan. 1981. **Linguistic Variation in Brazilian Portuguese: Aspects of Phonology, Syntax and Language History**. PhD dissertation, University of Pennsylvania. Ann Arbor: University Microfilms.
- GUY, Gregory Riordan. 1989. "On the nature and origins of Vernacula Brazilian Portuguese. In: **Estudios sobre Español de América y Lingüística Afroamericana**. Bogotá: Instituto Caro y Cuervo.

- HOLM, John. 1987. Creole influence on Vernácula Brazilian Portuguese. In: Gilbert, Glenn (ed.). **Pidgin and Creole Languages**. Honolulu: University of Hawaii Press, 406-429.
- HOLM, John. 1988-9. **Pidgins and Creoles**. Volumes I e II. Cambridge: Cambridge University Press.
- HOLM, John. 1992. Vernacular Brazilian Portuguese: a semi-creole. In.: d'Andrade, E. e Kihm, A. (orgs.). **Actas do Coloquio sobre Crioulos de Base Lexical Portuguesa**. Lisboa: Colibri. pp. 37-66.
- JEROSLOW, Helen McKinney. 1975. Creole characteristics in Rural Brazilian Portuguese. Comunicação apresentada à **Conferência Internacional sobre Línguas Pidgins e Crioulas**, Universidade do Haváí, 1975.
- LABOV, Willam. 1964. Estágios na aquisição do inglês *standard*. In.: FONSECA, M. S. V. e NEVES, M. F. (orgs.). **Sociolingüística**. Rio de Janeiro: Eldorado, 1974.
- LABOV, Willam. 1972. **Sociolinguistics Patterns**. Oxford: Basil Blackwell.
- LABOV, Willam. 1982. Building on Empirical Foundations. In: LEHMANN, W. P. e MALKIEL, Y. (eds.). **Perspectives on Historical Linguistics**. Amsterdam-Philadelphia: John Benjamins Publishing Company. pp 17-92.
- LUCCHESI, Dante. 1993. The article systems of Cape Verde and São Tomé Creole Portuguese: general principals and specific factors. **Journal of Pidgin and Creole Languages**, Filadelfia (EUA) e Amsterdam (Holanda), John Benjamins Publishing CO., 8, (1): 81-108.
- LUCCHESI, Dante. 1994. Variação e Norma: elementos para uma caracterização sociolingüística do português do Brasil. **Revista Internacional de Língua Portuguesa**, Lisboa, 12: 17-28.
- MEGENNEY, William W. 1993. Helvecian Portuguese: Vernacular dialect or true creole? **Journal of Pidgin and Creole Languages**. 8(2): 237-40.
- MELO, Gladstone Chaves de. 1981. **A Língua do Brasil**. Rio de Janeiro: Padrão.
- MENDONÇA, Renato. 1933. **A influência africana no português do Brasil**. Rio de Janeiro: Sauer.
- NARO, Anthony J. 1981. "The social and structural dimensions of a syntactic change", **Language**, LSA, 57(1): 63-98.
- NARO, Anthony e SCHERRE, M^a Marta Pereira. 1991. "Variação e Mudança Lingüística: Fluxos e Contrafluxos na Comunidade de Fala", **Cadernos de Estudos Lingüísticos**, 20: 9-16.
- PERRONI SIMÕES, Maria Cecília. 1976. **Aspectos da Gramática Portuguesa aos dois anos de idade**. Tese de mestrado, Universidade Estadual de Campinas.
- PERRONI SIMÕES, Maria Cecília e STOEL-GAMMON, Carol. 1979. The acquisition of inflections in Portuguese: a study of the development of person markers on verbs. **Journal of Child Language**, 6: 53-67.
- RAIMUNDO, Jacques. 1933. **O elemento afro-negro na língua portuguesa**. Rio de Janeiro: Renascença.
- SILVA NETO, Serafim da. 1986. **Introdução ao estudo da língua portuguesa no Brasil**. Rio de Janeiro: Presença.
- TARALLO, Fernando. 1993. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In: ROBERTS, Ian e KATO, Mary (orgs.). **Português Brasileiro - uma viagem diacrônica**. Campinas: Editora da UNICAMP.
- THOMASON, Sarah G. e KAUFMAN, Terrence. 1988. **Language Contact, Creolization, and Genetic Linguistics**. Berkeley: University of California Press.

From Latin Weak Pronouns to Romance Clitics

Giampaolo Salvi

Universidade de Budapeste, Hungria

Resumo

Trata-se neste estudo da relação diacrônica entre os pronomes pessoais do latim e das línguas românicas dele derivadas. Nas línguas românicas os pronomes pessoais pertencem a duas diferentes categorias morfológicas e sintáticas, ou seja, formas livres e clítics, enquanto no latim os pronomes pessoais pertenciam a uma única categoria morfológica, com formas tônicas e formas átonas. Dessas últimas se derivam os clítics românicos. Para atingir o objetivo deste estudo serão apresentadas as principais características da estrutura da sentença nas línguas românicas antigas e algumas generalizações sobre a colocação dos clítics; uma análise da estrutura da sentença no latim e da posição dos pronomes pessoais fracos; por fim serão traçadas as principais linhas da evolução do latim para as línguas românicas.

While Romance personal pronouns belong to two different morphological and syntactic categories: free forms and clitics, Latin had only one morphological type of personal pronouns. I intend to show here 1) that the Latin forms could be tonic (or *strong*) or atonic (or *weak*) and had accordingly a different syntactic behaviour and a different pragmatic use, and 2) that the weak forms are the source of Romance clitics and can explain most of their syntactic properties (e.g. their position). For this purpose, I will first briefly review the principal features of Old Romance sentence structure and the main generalizations about clitic placement (section 1), then propose an analysis of Latin sentence structure and of the position of those pronouns that can be considered *bona fide* weak (section 2); finally, I will trace the main lines of the diachronic evolution from Latin to Romance (section 3).

1. Old Romance

Old Romance languages were V2-languages (Salvi 1991, with extensive bibliography; Battye-Roberts 1995; Ribeiro 1995).

1.1. In the main clause, the inflected verb occupied the second position in the sentence and was preceded by one constituent with the pragmatic function of theme (1a) or focus (1b-c). The grammatical function of the constituent in first position was indifferent: it could be the subject (1a), the direct

object (1b), a prepositional complement (1c), an adverb (1d). The position preceding the verb could remain empty in particular constructions (yes-no questions, imperative sentences, beginning of narrative passages, etc. — (1e)):

- (1) X (cl) V ... / V (cl) ...¹
- a. S (cl) V ...: questa mi sembra più bella e di
this-one to-me seems more beautiful and of
maggiore valuta (Nov. II.33)
greater value
 - b. O (cl) V ...: cotanto vi dico (Nov. II.18)
so-much to-you I-say
 - c. PP (cl) V ...: di grande scienza ti credo (Nov. II.45)
of great learning you I-believe (to be)
 - d. Adv (cl) V ...: manifestamente l' ho veduto nelle cose in
manifestly it I-have seen in-the things in
ch'... (Nov. II.45)
which
 - e. V (cl)...: fu li detto che... (Nov. II.11)
it-was to-him said that

(Examples are from Old Italian, but are meant to be representative of all Medieval Romance languages; for the more archaic construction types, not or scarcely attested in Italian, Old Portuguese examples are used. Notice that the common Old Romance syntax we are referring to here was not actually realized in all languages at the same time: so Italian texts generally show a more archaic character than the contemporary French ones, but more modern features than the Spanish or Portuguese ones.)

As far as clitics are concerned, they were placed after the first realized element of the sentence, as the examples in (1) show.

1.2. The core part of the sentence described in the previous section could be preceded by peripheral (left dislocated) constituents, as in (2):

¹ The abbreviations not explained in the text are as follows:

Adv	adverb	P	preposition
C	subordinate clause	PP	prepositional phrase
cl	clitic	que	complementizer or SpecC"-
de	peripheral element introduced by <i>de</i>	REFL	reflexive
Disl	peripheral element	S	subject
O	direct object	V	verb
p	weak pronoun	X	constituent

(2) Disl | X (cl) V... / Disl | V (cl)...

- a. dopo non molti giorni, lo re si pensò di... (Nov.
after not many days the king REFL thought of II.42)
- b. O_i | X cl_i V...: tutti coloro de la terra ch' erano
all those of the country who were
colpevoli, il Grande Cane li fece uccidere (Vanelli 1986: p.
guilty the Great Khan them made kill 262)
- c. C | X (cl) V ...: se no 'l mi dirai, io ti far[~] di
if not it to-me you-will-say I you will-make of
villana morte morire (Nov. II.53)
infamous death die
- d. C | V (cl) ...: essendo poveramente ad arnese, misesi
being miserably equipped he-began-REFL
ad andare ad Alexandro (Nov. III.6)
to go to Alexander

The fact that the constituents separated by a vertical bar from the rest of the sentence are external with respect to the latter, appears clearly from example (2b): here, the direct object does not immediately precede the verb and the use of a resumptive clitic is compulsory. We can explain this fact in the following way: as in V2-languages only one constituent can precede the verb, the elements that do not precede it immediately are peripheral; on the assumption that peripheral elements cannot satisfy the Projection Principle, a resumptive clitic is inserted in the core part of the sentence for this purpose.

Peripheral elements are often represented by subordinate clauses (2c-d). This will be of great importance for our study of Latin sentence structure in section 2: since Latin does not have the construction in (2b), we will mainly use subordinate clauses as clear-cut instances of peripheral elements.

As for clitics, peripheral elements do not influence their position: as the examples in (2) show, clitics appear there where they would appear if no peripheral element were present.

1.3. Subordinate clauses have mostly the same structure and the same word order as main clauses, except for the presence of a complementizer (or a SpecC"-filling element):

(3) que | Disl | X (cl) V .../ que | Disl | V (cl) É

- a. que (cl) V ...: ché m' è fatto conto che... (Nov. II.15)
because to-me it-is told that
- b. que S (cl) V ...: che un pane intero li fosse dato
that a loaf whole to-him were given
per giorno (Nov. II.40)
daily
- c. que PP (cl) V ...: che di tutte le cose t' intendi (Nov.
that in all the things REFL you-are-an-expert II.27)

- d. que Adv (cl) V ...: che allora m' avidi cui figliuolo
that then REFL I-realized whose son
voi foste (*Nov. II.76*)
you were
- e. que | Disl | V (cl) ...: (port.) que nembrando-lhe a
that occurring to-him the
criaçom que em elles fezera... viinham-lhe as lagrimas
education that to them he-had-given came to-him the tears
aos olhos (*D. Pedro VI.60*)
to-the eyes

Clitics occupy the same position as in the corresponding main clauses (3b-e). The only exception is when the verb is the first element after the complementizer in the subordinate clause: in this case clitics always follow the complementizer (3a), which apparently counts here as the first element of the clause domain.

1.4. In Old Romance there existed a rarer and (as we shall see later) more archaic type of sentence, attested mainly in Old Portuguese and Old Spanish. Except for possible peripheral elements, it was opened by a focused constituent immediately followed by the clitics, while the verb occupied a position farther in the sentence and could be separated from the clitics by one or more elements:

(4) Disl | X cl... V...

- a. X (cl) S V...: (port.) logo lhe el- rrei taxava que
at once to-him the king ordained that
ouvesse por dia quatro soldos, e mais nom (*D. Pedro IV.64*)
he-had daily four shillings and more not
- b. C | X (cl) Adv V...: (port.) se nō, logo vos agora matarei
if not at once you now I-will-kill
(*Crônica CDLXXI.14*)
- c. X (cl) S Adv V...: (port.) tanto vos eu mui más precei
so-much you I much more prized
(Joan Airas 120)

The same separation of the clitics from the verb was possible in subordinate clauses, where the clitics could immediately follow the complementizer, while the verb was located farther in the sentence with the interpolation of one or more elements:

(5) que cl... V...

- a. que (cl) S V...: (port.) se se essa Stevaya Periz, nossa
if REFL that S. P. our
fila, cassar (*Doc. 138.12*)
daughter will-marry

- b. que (cl) Adv V...: *dacché vi pur piace* (Renzi 1989:p
since to-you yet pleases 369n)
- c. que (cl) S Adv V...: (port.) *como quer que o el-rei muito amasse*
although him the king much loved
- (D. Pedro VIII.47)

But note that the type exemplified in (4) is found in subordinate clauses, too: in this case clitics followed a focussed element and we have interpolation of one or more elements between the clitics and the subject:

- (6) que | Disl | X cl... V ...
- a. que X cl S V ...: (port.) *que sempre vos eu amey* (*Crônica CDLXIX.21*)
that always you I loved
- b. que | Disl | X cl S V ...: (port.) *que todollos filhos dos seus altos*
that all-the sons of-the his high-born
homees, depois que (...), logo os elle tomava (*Crônica CDXLVI.2*)
men after (...) at once them he took

(In the absence of peripheral elements, clitics could follow the complementizer, too, as appears from the contrast between (6a) and (7b) below.)

The syntactic type represented by main clauses in (1)-(2) and subordinate clauses in (3) (*type I*) and the one represented by main clauses in (4) and subordinate clauses in (5)-(6) (*type II*) show very different properties, e.g.:

- a) in type I adverbs are postverbal (7a), in type II they are preverbal (7b);
- b) Type II sentences admit free preposing of different elements in preverbal position without the features characteristic of left dislocation: in (7c) the direct object *aquello* does not immediately precede the verb, but there is no resumptive pronoun (contrast with (2b)):

- (7) a. (port.) *A raynha foi logo veer el rey* (*Crônica CCLXXV.7*)
the queen went at once to-see the king
- b. (port.) *que lhe logo el-rei nom mandou cortar a*
that to-him at once the king not made cut the
cabeça (D. Pedro IX.33)
head
- c. (Port.) *Quando lhe aquello el rey ouvyo dizer* (*Crônica CCLXXV.15*)
When to-him that the king heard say (When the
king heard him say that)

1.5. Table I summarizes the principal types of sentence structure and the possibilities for clitic placement in Old Romance, with a tentative (minimal) structural analysis (I leave the elements in peripheral position out of consideration).

Table I: Old Romance

sentence structure		clitics after	
I / main /	a. [FX V [S... ...]]	X	(1a-d)
	b. [FV [S... ...]]	V	(1e)
sub. /	a. [Cque [FX V [S... ...]]]	X	(3b-d)
	b. [Cque [FV [S... ...]]]	que	(3a)
II / main	[FX [S... V ...]]	X	(4)
sub. /	a. [Cque [S... V ...]]	que	(5)
	b. [Cque [FX [S... V ...]]]	que/X	(7b)/(6a)

We can interpret the different positions of the adverbs in type I and type II sentences as evidence that in type I sentences the verb moves into a functional position (F), while in type II sentences the verb remains in its base position within the sentence (S). In type I, Verb Movement could be accompanied by preposing (thematization or focalization) of a constituent (I/main/a, I/sub./a); in type II, focalization of a constituent was possible independently from Verb Movement (II/main, II/sub./b).

In type I sentences, clitics were placed after the constituent preposed in the F domain (I/main/a, I/sub./a); in clauses without constituent preposing, clitics were placed after the complementizer (or after the SpecC"-filling constituent), i.e. in the next upper domain of sentence structure (I/sub./b); if no such upper domain existed, clitics appeared after the verb (I/main/b). Type II sentences were possible only in two marked cases: in subordinate clauses and in main clauses with a focussed element; in these sentences, clitics were placed after the focussed element (II/main, II/sub./b) or after the complementizer (II/sub./a-b).

The position of clitics in Old Romance seems to obey some form of Wackernagel's Law (cf. Wackernagel 1892; for Romance cf. Renzi 1987, Salvi 1991).

2. Latin

As noted in the introductory section, while Romance personal pronouns belong to two different morphological and syntactic categories: free forms and clitics, Latin had only one morphological type of personal pronouns. But we may wonder whether, under this morphological uniformity, a difference

existed between two uses of the same forms, a difference that could have been the basis for the later evolution that resulted in the formation of the two different series in Old Romance. Such a question is all the more legitimate because other languages with only one morphological type of personal pronouns admit two uses of them (e.g. modern German; cf. Cardinaletti 1992, Lenerz 1992): they may be stressed or unstressed and this prosodic difference goes together with a different syntactic behaviour (e.g. possible positions) and with a different pragmatic use.

The existence of two *prosodically* different uses of Latin personal pronouns is demonstrated by the very existence of two different Romance outcomes of the same Latin personal pronoun, where the difference in the phonetic evolution can be explained on the basis of the stressed or unstressed nature of the form: so, e.g., from Latin *me* we have in Old French free *mei/moi* and clitic *me*, with the normal evolution of (Vulgar) Latin stressed *e* to the diphthong *ei/oi* and of unstressed *e* to [ə] (Schwan-Behrens 1932: 13). (Evidence from versification procedures is difficult to evaluate due to the quantitative nature of Latin verse.)

Latin pronouns could serve for different *pragmatic* uses: they could have the textual function of focus, of contrastive topic, of new topic and they could be simply anaphoric. These differences in pragmatic function could well be associated to the prosodic difference just noted, as in modern Romance languages free pronouns and clitics have different pragmatic functions (or semantics, as in Calabrese (1980)'s analysis, where clitics have expected referents and free pronouns have unexpected referents). It is interesting to note that in modern Romance translations of Latin texts, Latin personal pronouns can be rendered with free pronouns as well as with clitics, indirectly reflecting in this way the different pragmatic functions they have in the discourse.

As for the different *syntactic* uses of Latin personal pronouns, I will assume that differences in stress, in meaning and pragmatic function and in syntactic behaviour go together and distinguish two principal uses of Latin personal pronouns: *strong* (use of) pronouns and *weak* (use of) pronouns. Strong pronouns were stressed and had the function of focus, contrastive topic or new topic (unexpected referent); weak pronouns were unstressed and had a purely anaphoric function (expected referent). Syntactically, strong pronouns could occupy the same positions as nominal NPs, but weak pronouns, I will argue, were limited to a number of well definable positions (as in Old Romance, those described by a form of Wackernagel's Law).

Strong pronouns appeared in all possible positions: e.g., in absolute initial position, as focus (8a) or as contrastive topic (8b):

- (8) a. Mihi crede (*Fam.* IX.16.8)
 'Believe me'
 b. Tibi autem idem consili do quod mihimet ipsi (*Fam.* IX.2.2)
 'To you, I give the very same advice I give myself'

To determine which were the possible positions for weak pronouns is the task of the following sections.

For this purpose I examined a little *corpus* taken from Cicero's letters (numbers CCCCLXXVIII-DVI of Beaujeu's edition), a text chosen as a good representative of colloquial Latin. For lack of direct prosodic evidence, the only basis for the identification of weak pronouns is their semantic/pragmatic function: indeed a dangerous ground to go on, because we can never be absolutely sure that we have grasped the author's intention about the informative value of a form. Nevertheless, even if we adopt a very conservative approach and leave out of consideration only those examples in which the context makes us sure that the pronoun is focussed or contrasted, the result is very encouraging: the overwhelming majority of the remaining forms takes an absolutely regular position within the sentence and, what is more, this regularity is easily comparable to the one observed in Old Romance.

We will consider first only the non-nominative, non-prepositional (mainly accusative and dative) forms of I and II person pronouns, of reflexives and of the anaphoric *is* (although the ancestor of Romance III person pronouns is Lat. *ille*, it was *is* that had the anaphoric function typical of weak pronouns in Classical Latin Ñ de Jong 1993); nominative and prepositional forms will be discussed in sections 2.7.1 and 2.7.2, respectively. Only pronouns depending on tensed verbs will be considered (for pronouns depending on non-finite forms cf. Salvi 1993: 5.4).

Results partially similar to my own were independently reached by Adams (1994a,b) and Janse (1994); cf. also Kiparsky (1995) on Germanic.

2.1. In the Latin sentence, the verb normally occupied the final position. As a matter of fact, some constituents could appear after the verb: this was the normal case when the constituent was a clause, or with rhematic elements (Panhuis 1982; Pinkster 1988: ch. 9), but I will not dwell on this subject, since it does not affect the problem of weak pronouns' placement.

In main clauses, the first position could be held by a constituent with the pragmatic function of theme (9) or of focus (10), independently of its grammatical function; under particular conditions (e.g. emphasis), also the verb could come first (11):

(9) X_{theme} (p) ... V

- a. S (p) ... V: Caninius noster me tuis uerbis admonuit ut (...) (*Fam.* IX.6.1)
'Our friend Caninius gave me your message, reminding me to (...)'
- b. O (p) ... V: Haec tibi antea non rescripsi, non quo (...) (*Fam.* IX.17.3)
'If I did not send you this reply before, it is not that (...)'
- c. PP (p) ... V: Ex Syria nobis tumultuosiora quaedam nuntiata sunt
(*Fam.* XII.17.1)
'We have reports from Syria of some rather serious disturbances there'
- d. Adv (p) ... V: ante te certiores faciam (*Fam.* IX.5.3)
'(I shall) give you notice beforehand'

(10) X_{focus} (p) ... V

- a. S (p) ... V: nihil te omnino fefellit (*Fam.* IX.2.2)
'nothing whatever escaped your notice'
- b. PP (p) ... V: in Epicuri nos, aduersarii nostri, castra coiecimus
(*Fam.* IX.20.1)
'I have thrown myself into the camp of my former adversary Epicurus'
- c. Adv (p) ... V: ita se cum multis conligauit (*Fam.* IX.17.2)
'so inextricably has he tied himself up with his multitude of counsellors'

(11) V (p) ...

- Delectarunt me tuae litterae (*Fam.* IX.16.1)
'I was delighted with your letter'

As it is clear from the above examples, weak pronouns are placed after the first constituent of the sentence or after the first (independent) word of the first constituent (10b) (we will not treat this peculiarity of Latin word order here; about the non-independent word status of prepositions, see below section 2.7.2).

2.2. As in Old Romance, the core sentence could be preceded by peripheral elements: constituents introduced by *de* 'as for' (12a), embedded clauses (12b-c), etc.:

(12) Disl | X (p) ... V / Disl | V (p) ...

- a. *de* | X (p) ... V: De Aufidiano nomine nihil te hortor (*Fam.* XVI.19)
'In the matter of Aufidius's debt, I put no pressure upon you'
- b. C | X (p) ... V: si proficiscerer ad bellum, periculum te meum commouebat (*Fam.* VII.3.1)
'if I set out for the scene of war, you were appalled at the thought of my danger'
- c. C | V (p) ...: cum enim salutationi nos dedimus amicorum ..., abdo [V] me in bibliothecam (*Fam.* VII.28.2)
'after devoting myself to the reception of my friends (...), I hide myself away in my library'

As the examples show, weak pronouns occupy the same position they would if peripheral constituents were not there.

2.3. Latin possessed a weak verb form, too: the copula *esse* 'to be' (or at least some of its forms; cf. Adams 1994b). Besides the normal clause final position (13), *esse* could also follow the first element of the clause, especially if that was focussed (14): the copula position is then the same as for weak pronouns.

- (13) hoc etiam κατὰ; Χρυσησιππον δυνατὸν est (*Fam.* IX.4)
 "for this also is possible according to Chrysippus"
- (14) X_{focus} V_{weak} É
- a. Quis est tam Lynceus qui (...)? (*Fam.* IX.2.2)
 "who [...] is such an absolute Lynceus as (E)"
- b. Itaque nullum est ἀφοπρηγμένον quod non verear (*Fam.* IX.7.2)
 'so there is every possible objectionable factor for me to fear'
- c. ita est a me consulatus peractus ut (...) (Adams 1994b: p. 38)
 "Indeed my consulship was so conducted from its beginning to its end, that (...)"

2.4. In subordinate clauses, weak pronouns appear after the complementizer (or the SpecC"-filling constituent) (15) or, as in main clauses, after the first constituent of the clause, especially if that constituent is focussed (16). In this case, as in main clauses, the presence of a peripheral element between the complementizer and the core part of the subordinate clause does not affect the position of weak pronouns (17):

- (15) que (p) ... V
- a. quamquam me non ratio solum consolatur (...), sed etiam(...)
 (*Fam.* VII.28.3)
 'although I am comforted not only by rational reflection (...), but also (...)'
- b. quo me coniectura ducat (*Fam.* IX.2.4)
 'where my conjectures lead me'
- c. qui se domo non commouerunt (*Fam.* IX.5.2)
 'who never stirred from home'
- (16) que X (p) ...V / que V (p) ...
- a. que O (p) ... V: qui nulla sibi subsidia ad omnis uitae status parauerunt (*Fam.* IX.6.4)
 'who have provided themselves with no resources against any vicissitudes of existence'
- b. que Adv (p) ... V: quod non ultro mihi Caesar detulerit
 (*Fam.* IV.13.2)
 'which Caesar has not spontaneously bestowed upon me'
- c. que V (p) ...: Etsi enim et audio te et uideo lubenter (*Fam.* XVI.22.1)
 'though it is a pleasure to me to hear and see you'
- (17) que | Disj | X (p) ...V / que | Disj | V (p) ...
- a. que | C | Adv (p) ... V: quae, quia tibi sunt propiora quam nobis, tua me causa magis mouent quam mea (*Fam.* II.17.1)

'and [lit.: which] as they touch you more nearly than they do me, they cause me more agitation on your account than my own'

- b. que | C | V (p) ...: ut, quomodo in tali re atque tempore, aut liberarem te ista cura aut certe leuarem (*Fam.* IX.16.1)
'that I might, seeing how the matter stands, considering the circumstances and the time, free you from your anxiety, or at any rate alleviate it'

Notice that, if there was a focussed constituent in the subordinate clause, it was possible for weak pronouns to follow it (16) or to follow the complementizer, too (15a) (where *non ratio solum* is contrastive focus).

The weak forms of the copula *esse* have the same possibilities of placement in subordinate clauses as the weak pronouns do: after the complementizer (18a) or after the first constituent, especially if focussed (18b), the presence of peripheral elements between the complementizer and the core part of the clause being irrelevant (18c) (cf. Adams 1994b):

- (18) que V_{weak} ... / que X_{focus} V_{weak} ... / que | Disj | X_{focus} V_{weak} ...
- a. que V_{weak} ...: quid sis acturus (*Fam.* IX.2.5)
'what you intend to do'
- b. que Pred V_{weak} ...: cum plena sint monumenta Graecorum quemadmodum (...) (*Fam.* IX.16.6)
'seeing that the annals of the Greek abound in examples of how (...)'
- c. que | Disj | Adv V_{weak} ...: qui una cum hoc Furio semper erat in hac Allobrogum sollicitatione uersatus (Adams 1994b: 39)
'who had always been associated with this Furius in tampering with the Allobroges'

2.5. Finally, in subordinate clauses, Latin permitted a constituent with theme or focus function in the position preceding SpecC". In this construction weak pronouns followed the preposed element (19a) or the complementizer (19b):

- (19) X (p) que ... / X que (p) ...
- a. Mortem mihi cur consciscerem causa non uisast, cur optarem multae causae (*Fam.* VII.3.4)
'Why I should contrive my own death there seemed no reason; why I should pray for it there were many'
- b. Ego quanti te faciam semperque fecerim (*Fam.* VI.10.1)
'How highly I esteem and always have esteemed you'

2.6. Table II summarizes the principal types of sentence structure and the possibilities for the placement of weak pronouns in Latin, with a tentative (minimal) structural analysis (as for the Old Romance cases in section 1.5, I leave the elements in peripheral position out of consideration).

Table II: Latin

	sentence structure	weak pronoun after	
main	a. [SX ... V]	X	(9)
	b. [FX [S ... V]]	X	(10)
	c. [FV [S ...]]	V	(11)
sub.	a. [Cque [S ... V]]	que	(15b-c)
	b. [Cque [FX [S... V]]]	que/X	(15a)/(16a-b)
	c. [Cque [FV [S...]]]	que/V	?(16c)
	d. [GX [Cque [S ... V]]]	X/que	(19a)/(19b)

We may assume that the Latin sentence had a V-final structure (main/a) and a rule of focussing/emphasis that preposed a constituent (including the verb) into an initial functional position (main/b-c) (cf. Panhuis 1982). We have no positive evidence for postulating a parallel rule of topicalization, as in Old Romance: the expression of the topic was the task of peripheral elements. Subordinate clauses presented the same possibilities as main clauses (sub./a-c) and in addition constituents (but not the tensed verb) could be moved into a position preceding the C-domain (position labeled here G: sub./d).

Weak pronouns were freely placed after the preposed constituent in the F/G-domain (main/b-c, sub./b-d) or after the complementizer (sub./a-d). In main clauses without focussing, weak pronouns appear after the first element of the S-domain (main/a).

We could imagine that the position of the weak pronouns was bound to the realization of one of the functional domains F/C/G for reasons of semantic interpretation, e.g. because, being thematic elements, they had to have S in their domain. We could also conceive that, being unstressed, they could not appear in absolute initial position in the sentence at the level of Phonetic Form. So, weak pronouns always raised to a functional domain higher than S; if this domain was phonologically realized, weak pronouns did not violate the filter forbidding them to appear in absolute initial position; if this higher domain was not phonologically realized, a Last Resort rule applied that postposed the weak pronoun, avoiding the application of the filter.

2.7. Independently of the details of the analysis, at a descriptive level, it is evident that the position of weak pronouns in Latin and the position of clitics in Old Romance were very similar, as Tables I and II show. But before I trace the lines of the diachronic evolution from the older system to the newer one, I must clarify why I insist in keeping weak pronouns and clitics distinct, or, in other words, what the reason is for claiming that Latin weak pronouns

were not clitic (yet). (As for the copula *esse*, cf. the discussion in Adams 1994b: 12.2).

2.7.1. It was not only the non-nominative forms of Latin personal pronouns that had a weak use, but also the nominative ones, as their placement in the examples in (20) shows (cf. Adams 1994a: section 8):

- (20) a. Hunc ego mihi belli finem feci (*Fam.* VII.3.3) [cf.. (9)]
'This I resolved should be for me the end of the war'
- b. Quem tu mihi Popilius, quem denarium narras (...)?
(*Fam.* IX.16.7) [cf.. (10)]
'What is this Popilius you prate to me about, and this penny [...]?'
- c. quoniam tu secundum 'Oenomaum' Acci non (É) Atellanam, sed (...) mimum introduxisti (*Fam.* IX.16.7) [cf.. (15)]
'since you now, following the lead of Accius's *Oenomaus*, have staged not (...) an Atellan play, but (...) a farce'
- d. quanto semper tu et studio et officio in meis rebus fuisti (*Fam.* VI.22.3) [cf.. (16)]
'as you have consistently shown, and effectively too, in dealing with my affairs'
- e. quod ego cur nolim nihil uideo (*Fam.* IX.6.2) [cf.. (19)]
'and I see no reason why I should be indisposed to do so'

Now, if the non-nominative pronouns had been clitic, the nominative ones should have been too. But in the early stages of Romance languages only accusative and dative clitics existed: it would be very strange if a well-developed clitic system with forms for all grammatical cases had been reduced to only non-nominative forms, especially in view of the more generalized use of pronouns in Romance compared to Latin.

It is more natural to assume that in Latin weak pronouns were not clitic and that they became clitic only when the use of pronominal anaphora became obligatory. In Latin pronominal anaphora freely alternated with zero anaphora. Consider the following text excerpt:

- (21) Demetrius uenit ad me (...). Tu eum uidelicet non potuisti uidere; cras aderit; uidebis igitur (*Fam.* XVI.17.2)
'Demetrius came to see me (...). You evidently could not have seen him. He will be in Rome tomorrow, so you will see him then'

The English translation shows that the text refers to Demetrius anaphorically three times (with the underlined pronouns *him/he/him* respectively): in the Latin text, only the first reference is realized with an anaphoric pronoun (*eum*), the other two have zero anaphora. In Old Romance, pronominal anaphora was obligatory in the case of direct and indirect objects and was realized by a clitic; in the case of the subject, pronominal anaphora

was not obligatory and subject clitics did not exist (Vanelli-Renzi-Benincà 1985). It was only later, in some Romance languages as French and the Northern Italian dialects, that the expression of the subject became obligatory and these languages developed subject clitics, too (Vanelli 1987).

2.7.2. Not only personal pronouns used without a preposition had a weak use, but also those which were preceded by a preposition, as their placement in the following examples show (for a similar phenomenon in Old Czech cf. Esvan 1992):

- (22) a. Haec ad te scripsi uerbosius (*Fam.* VII.3.5) [cf. (9)]
'I have written all this to you with some verbosity'
- b. disce a me prolegomevna", quas quaeris (*Fam.* IX.18.3) [cf. (11)]
'let me at once teach you the first principles, since you ask for them'
- c. At tibi repente paucis post diebus, cum minime expectarem, uenit ad me Caninius mane (*Fam.* IX.2.1) [cf. (12)]
'But, lo and behold, a few days afterwards, when I least expected it, Caninius paid me a sudden visit one morning'
- d. cum ad me peruesperi uenisset (*Fam.* IX.2.1) [cf. (15)]
'[as he] came to me very late one evening'
- e. Si quid ad me scripseris (*Fam.* VI.22.3) [cf. (16)]
'If you send me any message'
- f. (Non ea res me deterruit) quo minus, postea quam in Italiam uenisti, litteras ad te mitterem (*Fam.* VI.22.1) [cf. (17)]
'(It was not the fact of (...)) that put me off) writing to you after your arrival in Italy'

The group *P + pronoun* is a phrase and this cannot be reconciled with the widely accepted view that clitics are heads (or rather, at least in their enclitic forms, a sort of affixes — cf. the following section).

To explain these facts, we must assume that prepositions were proclitic words in Classical Latin and that they could be adjoined to nominal NPs and to strong as well as to weak pronouns, not altering the character of their host. The cliticity of prepositions is proved by the fact that they could not be separated from their complement (perhaps the unique impossible scrambling in Latin word order) and that, differently from other types of phrases (cf. ex. (10b)), they did not admit a weak pronoun between them and their complement (examples as *Per ego uobis deos atque homines dico, ut (É)* (Plautus *Menaechmi* 990), where the weak pronouns *ego* and *uobis* are inserted between the preposition *per* and its complement *deos atque homines*, are limited to the archaic language).

2.7.3. In Latin, examples with a weak pronoun following and depending on two coordinated verbs are possible (23a); the same holds in German

(23b), while in Romance this construction is impossible with an enclitic pronoun (23c-d) (cf. Benincà-Cinque 1993):

- (20) a. sed cum consilii tui bene fortiterque suscepti eum tibi finem statueris quem ipsa fortuna terminum nostrarum contentionum esse uoluisset, oro optestorque te, pro uetere nostra coniunctione ac necessitudine proque summa mea in te beneuolentia et tua in me pari, (...) ut (...) (*Fam.* VI.22.2)
'But seeing that you have resolved that the policy you so honourably and gallantly adopted should cease from the very moment when it had pleased fortune herself to put an end to our struggles, I beg and implore of you in the name of our old and intimate connection and of my most sincere affection for you, and of yours, just as sincere, for myself, to (...)'
- b. Jeder von ihnen verbirgt und versteckt sich.
each of them conceal and hide himself
- c. (Port.) *Oiço e vejo-te
I-hear and I-see you
- d. (Port.) Oiço- te e vejo-te (Barreto 1980: p. 237n)
I-hear you and I-see you

If we assume that clitics, and especially enclitics, form a sort of compound with their host, the behaviour of Romance enclitics is straightforward: two coordinated verbs are not a word, so composition is impossible. The very fact that in Latin and in German this construction is possible, shows that these pronouns are not clitics, but independent words (though weak), to which no such restriction is applicable.

2.7.4. Weak pronouns might be separated from their host by other weak words: besides copula (cf. (14c)), mainly adverbs with connective value, as in the following examples:

- (24) a. Sollicitum autem te habebat cogitatio cum officii tum etiam periculi mei (*Fam.* VII.3.1) [cf. (9)]
'But your anxiety was due to your brooding over the double problem of my duty and my danger'
- b. illud tamen tibi polliceor (*Fam.* VI.22.3) [cf. (10)]
'this much, however, I promise you'
- c. erit enim nobis honestius (...) (*Fam.* IX.2.5) [cf. (11)]
'for it will be more creditable to us (...)'
- d. cum enim te semper magnum hominem duxerim (*Fam.* IX.6.4) [cf. (15)]
'while I ever deemed you a great man'

Although the syntactic behaviour of these adverbs was not exactly the same as in the case of personal pronouns (they always followed the first word, not the first constituent of the clause), they were clearly weak (they could not occupy the absolute initial position in the sentence). Now, if these

groups of weak words had been enclitic in the same sense as Romance clitics may be, we would have to assume the formation of very complex compounds, which seems unnatural. Other solutions are surely imaginable, but it is clear that the Latin facts are very different from the Romance ones and deserve a different analysis. (Note that, if weak pronouns had been enclitic in Latin, this would have had some consequences in the accentual pattern of the group *host* + *enclitic* due to the nature of Latin stress; Latin grammarians do not report any fact of this type.)

3. From Latin to Old Romance

I will now sum up the main similarities and differences between Latin and Old Romance in the domain of sentence structure and of weak pronoun vs. clitic placement and then I will trace the main lines of the diachronic evolution.

3.1. If we compare Table I (section 1.5) and Table II (section 2.6), it turns out that:

- a. if we disregard the position of the verb (final in Latin, not final in Old Romance), *Old Romance Type II sentences* preserved the Latin sentence structure quite well, namely:
 - the main clause structure corresponds to Latin pattern *b*;
 - the subordinate clause structures correspond to Latin patterns *a* and *b*;
 - clitics occupy the same position as Latin weak pronouns.
- b. *Old Romance Type I sentences* represent the innovative trend:
 - they generalize Verb Movement into a pre-S functional position (in Latin, the verb could move as an option of the general rule of focussing/emphasis);
 - they generalize Constituent Preposing to thematization/topicalization (in Latin, it was limited to focussing/emphasis);
 - the position of clitic pronouns is fixed more strictly: in principle they follow the preposed constituent (pattern *a* of main and subordinate clauses); in the absence of such a constituent, they are cliticized in the upper domain C (pattern *b* of subordinate clauses); in the absence of an upper domain, they are enclitic to the verb (Law Tobler-Mussafia — pattern *b* of main clauses). In Latin, weak pronouns might be freely placed after the focussed element (constituent or verb) or in the upper domain C (patterns *b* and *c* of main clauses and all the patterns of subordinate clauses); in the absence of these contexts, they were placed after the first element of the S-domain.
- c. the possibility of preposing a constituent into a pre-C position (Latin pattern *d* of subordinate clauses) is not preserved in Romance.

3.1. The main innovations in the evolution of sentence structure from

Latin to Old Romance are then (disregarding the more archaic Type II):

- a. the different basic position of the verb (non final);
- b. the generalization of Verb Movement into a functional position;
- c. the extension of Constituent Preposing to the cases of thematization/topicalization.

As for the weak pronouns:

- d. they had become clitic;
- e. their position had been fixed.

I will conclude with some remarks on these points.

To point *a*: Verb Movement being generalized, the basic position of the verb can be established only on the basis of Type II sentences, in which the verb did not raise into a pre-S functional position. It is not clear to me what the relation is between this diachronic change and the following one.

To point *b*: as for the origin of this construction, we may think that it is based on an preexisting Latin structure and we may wonder whether it is a generalization of the rule that focussed/emphasized a constituent and that could focus/emphasize the verb too (Lernerz 1985) or it is the extension to all the other verbs of the placement of the (weak) copula (Hock 1982). Both hypotheses entail some difficulties: while in Latin the focussing of the verb was in complementary distribution with the focussing of another constituent, in Old Romance Verb Movement always accompanies Constituent Preposing. We would thus have to explain the cooccurrence of the two movements and also the fact that the preposed verb had no emphatic value anymore. The copula hypothesis does not face this last problem, but comes up against difficulties when it has to explain V1-sentences (cf. (1e) — being weak, the copula could not appear initially in the sentence). For some discussion cf. Kiparsky (1995: 3.3).

To point *c*: the fact that a preposed constituent could have the pragmatic function of theme/topic, had some repercussion on the possible functions of peripheral elements; in some Romance languages the peripheral position tended to be limited to circumstantials (to arguments only when they were heavy — cf. de Kok (1985: 2.1) on Old French, Vanelli (1986) on Old Italian, Salvi (1993a: 4.1) on the differences between Old Portuguese and Old Spanish).

To point *d*: cf. the discussion in section 2.7, especially 2.7.1.

To point *e*: adverbality is not an attribute Romance clitics had from the beginning, as examples (4)-(6) show and as is natural, since Latin weak pronouns were not bound to the verb, but to the first element of a syntactic

domain. But the adverbiality of clitics was typical of Type I sentences and this is an indirect consequence of the obligatory rule of Verb Movement: in main clauses the verb was always the first or the second element of the clause and the clitics were placed after the first element of that domain, so they were always adjacent to the verb (a similar reasoning is applicable to subordinate clauses).

We may conclude that Romance clitics are the descendants of Latin weak pronouns and that their position is essentially the position weak pronouns occupied in the Latin sentence: the differences in this domain are only apparent and are due to the changes that had independently occurred in the evolution of sentence structure².

Résumé

On aborde dans ce travail le rapport diachronique entre les pronoms personnels du latin et des langues romanes qui en dérivent. Dans les langues romanes, les pronoms personnels appartiennent à deux catégories morphologiques et syntaxiques différentes, c'est-à-dire, à des formes libres et à des clitiques, tandis qu' en latin les pronoms personnels appartenait à une seule catégorie morphologique, avec des formes toniques et des formes atones. De celles-ci sont dérivés les clitiques romans. Pour atteindre le but de cette étude sont présentées les principales caractéristiques de la structure de la phrase dans les langues romanes anciennes et quelques généralisations sur l'emploi des clitiques; une analyse de la structure de la phrase en latin et de la place des pronoms inaccentués; enfin on retrace les principales lignes de l'évolution du latin vers les langues romanes.

2 *Acknowledgments*. The material presented here develops parts of an unpublished paper (Salvi 1993b), which I had the opportunity to present in various forms to several audiences (University of Padua 1993; VIIth International Colloquium on Latin Linguistics, Hebrew University, Jerusalem 1993, University of Hamburg 1993; University of Venice 1993; Ist International Congress of the Brazilian Linguistic Association, Federal University of Bahia, Salvador 1994, Linguistic Institute of the Hungarian Academy of Sciences, Budapest 1995; VIIIth International Colloquium on Latin Linguistics, University of Eichstätt 1995; XXIst International Congress of Romance Linguistic and Philology, University of Palermo 1995): my thanks to the participants, for my ideas on these topics benefited a lot from the discussion. I am grateful also to Mark Janse, who read and punctually commented on a first draft of Salvi (1993b).

References

- Adams, J.N. (1994a): Wackernagel's Law and the Position of Unstressed Personal Pronouns in Classical Latin, in *Transactions of the Philological Society*, 92, 103-178.
- Adams, J.N. (1994b): *Wackernagel's Law and the Placement of the Copula Esse in Classical Latin*, The Cambridge Philological Society.
- Barreto, M. (1980): *Novos Estudos da Língua Portuguesa*, Rio de Janeiro, Presença/INL-FCRB-MEC.
- Battye, A.-I. Roberts (eds.) (1995): *Clause Structure and Language Change*, New York—Oxford, Oxford University Press.
- Benincà, P. - G. Cinque (1993): Su alcune differenze fra enclisi e proclisi, in *Omaggio a Gianfranco Folena*, Padova, Editoriale Programma, 2313-2326.
- Calabrese, A. (1980): Sui pronomi atoni e tonici dell'italiano, in *Rivista di Grammatica Generativa*, 5, 65-116.
- Cardinaletti, A. (1992): On Cliticization in Germanic Languages, in *Rivista di Grammatica Generativa*, 17, 65-99.
- de Jong, J.R. (1993): *The Borderline between Deixis and Anaphora in Latin*, paper presented at the VIIth International Colloquium on Latin Linguistics, Jerusalem.
- de Kok, A.C. (1985): *La place du pronom personnel régime conjoint en français. Une étude diachronique*, Amsterdam, Rodopi.
- Esvan, F. (1992): *Quelques observations sur l'ordre des mots en tchéque d'époque préhussite*, in 'Europa Orientalis', 11, 163-179.
- Hock, H.-H. (1982): *AUX-Cliticization as a Motivation for Word-order Change*, in *Studies in the Linguistic Sciences*, 12, 91-101.
- Janse, M. (1994): *La loi de Wackernagel et ses extensions en latin: à propos de la collocation pronominale chez Pétrone*, in *Tema. Techniques et Méthodologies Modernes Appliquées à l'Antiquité*, 1, 107-146.
- Kiparsky, P. (1995): *Indo-European Origins of Germanic Syntax*, in Battye—Roberts (eds.) (1995), 140-169.
- Lenerz, J. (1985): *Diachronic Syntax: Verb Position and COMP in German*, in J. Toman (ed.): *Studies in German Grammar*, Dordrecht, Foris, 103-132.
- Lenerz, J. (1992): *Zur Syntax der Pronomina im Deutschen*, in *Sprache und Pragmatik*, 1-54.
- Panhuis, D.G. (1982): *The Communicative Perspective in the Sentence: A Study of Latin Word Order*, Amsterdam, Benjamins.
- Pinkster, H. (1988): *Lateinische Syntax und Semantik*, Tübingen, Francke.
- Renzi, L. (1987): *Essor, transformation et mort d'une loi: la loi de Wackernagel*, in *Mélanges offerts à Maurice Molho*, vol. III: *Linguistique, Les Cahiers de Fontenay*, 291-302.
- Renzi, L. (1989): *Two Types of Clitics in Natural Languages*, in *Rivista di Linguistica*, 1, 355-372.
- Ribeiro, I.M.O. (1995): *A sintaxe da ordem no português arcaico; o efeito V2*, PhD dissertation, Campinas, SP.
- Salvi, G. (1991): *Difesa e illustrazione della legge di Wackernagel applicata alle lingue romanze antiche: la posizione delle forme pronominali clitiche*, in L. Vanelli - A. Zamboni (eds.): *Per Giovan Battista Pellegrini. Scritti degli allievi padovani*, Padova, Unipress, pp.439-462.
- Salvi, G. (1993a): *Ordine delle parole e struttura della frase nelle lingue romanze antiche*, in Chr. Schmitt (ed.): *Grammatikographie der romanischen Schprachen*, Bonn, Romanistischer Verlag, 455-477.
- Salvi, G. (1993b): *La posizione dei pronomi in latino*, unpublished manuscript, Loránd Eötvös University, Budapest.

- Schwan, E.-D. Behrens (1932): **Grammaire de l'ancien français**, Leipzig.
- Vanelli, L. (1986): *Strutture tematiche in italiano antico*, in H. Stammerjohann (ed.): **Tema-Rema in Italiano**, Tübingen, Narr, pp. 249-273.
- Vanelli, L. (1987): *I pronomi soggetto nei dialetti italiani settentrionali dal Medio Evo a oggi*, in **Medioevo Romano**, 12, 173-211.
- Vanelli, L.-L. Renzi - P. Benincà (1985): *Typologie des pronoms sujets dans les langues romanes*, in **Actes du XVIIe Congrès International de Linguistique et Philologie Romanes, vol. 3: Linguistique descriptive, phonétique, morphologie et lexique**, Aix-en-Provence, Université de Provence, 163-176.
- Wackernagel, J. (1892): *Über ein Gesetz der indogermanischen Wortstellung*, in **Indogermanische Forschungen**, 1, 333-436.

Texts

- Cicéron: *Correspondance*, t. VII, ed. J. Beaujeu, Paris, 'Les Belles Lettres', 1980 (translations from Cicero: **The Letters to His Friends**, ed. W. Glynn Williams, Cambridge, Mass. - London, Harvard University Press/Heinemann, 1965 (1927), I-III).
- Crônica* = **Crônica Geral de Espanha de 1344**, ed. L.F. Lindley Cintra, Lisboa, 1951-61 (Lisboa, IN-CM, 1983).
- Doc.* = *Documentos* in C. de Azevedo Maia: **História do Galego-Português**, Coimbra, INIC, 1986, pp. 41-295.
- D. Pedro* = Fernão Lopes: **Crônica de D. Pedro**, ed. G. Macchi, Roma, Ateneo, 1966. *Fam.* see Cicéron
- Joan Airas = J.L. Rodr'guez: **El cancionero de Joan Airas de Santiago**. Edición y estudio, Universidade de Santiago de Compostela, 1980.
- Nov.* = **Il Novellino**, ed. G. Favati, Genova, Bozzi, 1970.

Do português clássico ao português europeu moderno: uma análise minimalista.

Charlotte Galves

Universidade Estadual de Campinas

Resumo:

Este artigo descreve e analisa a mudança ocorrida na colocação de clíticos no português europeu no séc 19. Nas orações raízes afirmativas com sujeito referencial, próclise e ênclise eram lícitas no português clássico (PCI). No mesmo contexto, só a ênclise é permitida no português moderno (PE). Propõe-se uma análise no quadro do Programa Minimalista de Chomsky, baseada na idéia de que as formas enclíticas correspondem a palavras construídas no léxico enquanto que as formas proclíticas são geradas na sintaxe. As formas enclíticas são então sujeitas a um processo de verificação de traços que implica a existência de traços correspondentes em outro lugar da estrutura. Tanto no PCI quanto no PE, é a categoria Comp que pode conter traços de concordância capazes de verificar os traços do clítico em ênclise. A diferença entre as duas gramáticas é que o verbo se move para Comp na primeira, mas não na segunda. Essa afirmação é justificada pelo caráter "V2" do português clássico.

O português europeu moderno (doravante PE) distingue-se claramente das outras línguas românicas pela sua sintaxe de colocação de clíticos. Quando as outras línguas românicas perderam a ênclise com verbo inicial já na sua fase clássica, o português europeu não só a manteve, como estendeu a sua obrigatoriedade às orações afirmativas com sujeitos referenciais. É dessa extensão, tomada como o índice de uma mudança paramétrica, que se tratará aqui, no quadro teórico do Programa Minimalista para a Teoria Lingüística desenvolvido recentemente por Chomsky, em particular na sua última versão (Chomsky 1995). Propor-se-á uma formulação dessa mudança baseada numa concepção das formas enclíticas como palavras já formadas no léxico, em que a legitimação do clítico se faz pelo intermédio de uma verificação de traços análoga àquela a que são submetidos os morfemas flexionais.

O texto se organiza como segue. A primeira parte apresentará os fatos relevantes. A segunda parte explicitará os aspectos da teoria envolvidos na análise. A terceira enfim mostrará como os fatos derivam de uma parametrização diferente de uma categoria funcional identificada como Comp.

1. Os fatos.

1.1 O Português europeu moderno

A sintaxe de colocação de clíticos do PE nas frases com tempo é

extremamente complexa, e não é o objetivo desse texto apresentá-la exaustivamente. Para diversas apresentações e análises, remete-se o leitor a numerosos artigos recentes: Barbosa (1991, 1993), Duarte e Matos (1995), Galves (1992, 1994) Madeira (1992), Manzini (1992, 1994), Martins (1993, 1995), Raposo (1995), e Rouveret (1993, 1994-1995). A rápida apresentação que segue será suficiente para os propósitos deste artigo.

1.1.1 Estruturas de ênclise obrigatória

- a) o verbo está em primeira posição absoluta:
- (1) Parece-me que choverá amanhã
 - (2) * Me parece que choverá amanhã
- b) nas orações raiz afirmativas, quando o XP antecedendo o verbo é um sujeito referencial ou um tópico, retomado ou não por um pronome:
- (3) A Maria deu-lhe esse livro ontem
 - (4) *A Maria lhe deu esse livro ontem
 - (5) Essa rapariga, conheço-a muito bem
 - (6) * Essa rapariga, a conheço muito bem
 - (7) Na semana passada, encontrei-me com os meus antigos colegas de escola
 - (8) *Na semana passada, me encontrei com os meus antigos colegas de escola
- c) Nas estruturas encaixadas de deslocação clítica à esquerda, ou de topicalização, quando a oração é afirmativa e o sujeito é nulo ou referencial
- (9) Sabes que, a Maria, vi-a ontem
 - (10) * Sabes que, a Maria, a vi ontem
 - (11) Eu sei que, a Maria, o João viu-a ontem
 - (12) *Eu sei que, a Maria, o João a viu ontem

As frases (1) e (2) ilustram a aplicação da lei que proíbe um clítico em primeira posição, conhecida na literatura sob o nome de *lei de Tobler Mussafia*, do nome de dois gramáticos históricos do final do século passado que enunciaram essa generalização observada nas línguas românicas antigas. A estabilidade absoluta da aplicação dessa lei ao longo da sua história é um dos fatos marcantes do PE.

(3)-(8) mostram um aspecto mais recente da sintaxe portuguesa, que a lei de Tobler Mussafia não prevê. Trata-se do carácter obrigatório da ênclise mesmo quando um sintagma de certo tipo, sujeito ou não, antecede o verbo nas orações raízes afirmativas¹.

1 Uma das questões fundamentais da gramática do PE, uma vez que essa definição é crucial para a compreensão dos processos de legitimação da ênclise e da próclise em português, consiste em definir adequadamente esse tipo de sintagma, que caracterizamos acima como referencial. Para entender melhor o que isso significa, é interessante considerar os casos em

Enfim, o terceiro grupo de dados, que diz respeito às orações encaixadas mostra que a presença de um tópico torna a oração encaixada numa oração com características de matriz².

1.1.2 Estruturas de ênclise impossível

- a) nas orações encaixadas (menos nos casos considerados acima)
 - (13) Sei que a Maria lhe deu esse livro ontem
 - (14) *Sei que a Maria deu-lhe esse livro ontem³
- b) nas orações interrogativas polares, com o elemento interrogativo precedendo o verbo.
 - (15) Que livro a Maria lhe deu ontem?
 - (16) * Que livro a Maria deu-lhe ontem?
- c) nas orações negativas
 - (17) A Maria não lhe deu esse livro ontem
 - (18) *A Maria não deu-lhe esse livro ontem

que uma aparente escolha entre próclise e ênclise aparece. Martins (1992) discute a interpretação das duas frases seguintes que só diferem pela posição do clítico [a numeração é de Martins]:

(28a) Muitos amigos meus queixaram-se às autoridades

(29a) Muitos amigos meus se queixaram às autoridades

"a frase (28a) é verdadeira se um grande número de pessoas que são meus amigos se queixaram às autoridades. Mas (29a) é verdadeira somente se uma grande proporção dos meus amigos se queixou às autoridades. O número de amigos que eu tenho é irrelevante para avaliar a verdade de (28a), mas indispensável para avaliar a verdade de (29a)... Só (29a) tem a leitura proporcional característica dos quantificadores." Martins acrescenta que essa análise é reforçada pelo fato de que uma relativa não restritiva pode qualificar o sujeito na oração com ênclise, mas não na oração com próclise. Isso é devido ao fato de que as relativas não restritivas exigem um antecedente referencial.

i) Muitos amigos meus, que são antropólogos, queixaram-se às autoridades

ii) * Muitos amigos meus, que são antropólogos, se queixaram às autoridades

Barbosa (1993, nota 2) observa também o quantificador universal *todos* desencadeia a próclise embora permita a ênclise quando tem uma leitura de "grupo". A leitura de grupo é uma leitura referencial, por oposição à leitura quantificacional.

2. Raposo (1995) dá como possível a ênclise em orações encaixadas com sujeito explícito como no exemplo a seguir, que contrasta com a sua contrapartida com sujeito nulo:

i) Penso que eu vou-lhe dar um livro

ii) * Penso que vou-lhe dar um livro

Pode-se considerar que na primeira o sujeito está de fato em posição de tópico, com uma categoria vazia na posição do sujeito. Note-se todavia que nessas frases, a ênclise é ao máximo possível, mas nunca obrigatória, contrariamente a frases contendo um tópico não sujeito. Rouveret (1994-95) observa que isso mostra que o sujeito não é totalmente assimilável a um tópico em PB, contrariamente ao que afirmam linguistas como por exemplo Benincà (1995), Salvi (1990) e Barbosa (1991).

3. cf porém a nota anterior. É preciso ressaltar que os outros autores são unânimes quanto à impossibilidade de (14).

- d) quando um sintagma quantificado ou focalizado precede o verbo, seja ele sujeito ou não
- (19) Alguém o tinha avisado
 - (20) * Alguém tinha-o avisado
 - (21) Todos os alunos se riram
 - (22) * Todos os alunos riram-se⁴
 - (23) A todos o leram
 - (24) * A todos leram-no
 - (25) Muito trabalho me deu essa descrição
 - (26) *Muito trabalho deu-me essa descrição
 - (27) Até a ele lhe contaram mentiras
 - (28) * Até a ele contaram-lhe mentiras
- e) quando o verbo está no escopo de certos advérbios: *já, sempre, ainda, também, raras vezes...*

A propósito dos casos em que o sintagma preverbal desencadeador da próclise não é o sujeito, convém ressaltar que este pode preceder ou seguir o verbo sem que isso interfira na colocação proclítica. Em outros termos o sintagma interrogado, quantificado ou focalizado não precisa anteceder imediatamente o verbo para desencadear a próclise. Isso se verifica em (15) e nas orações seguintes, e mostra que a variação ênclise/próclise é independente do movimento do verbo⁵:

- (29) Muito trabalho essa descrição me deu
- (30) * Muito trabalho essa descrição deu-me
- (31) Até a ele as crianças lhe contaram mentiras
- (32) * Até a ele as crianças contaram-lhe mentiras

Enfim, é importante lembrar que a natureza interpretativa dos sintagmas pós-verbais não interfere na colocação de clíticos de maneira alguma. Nesse caso a ênclise é sempre obrigatória quando não há nenhum desencadeador da próclise em posição preverbal, como mostrado nos exemplos seguintes de Duarte e Matos (1995), onde o elemento que se encontra à direita do verbo tornaria a próclise obrigatória se ele estivesse à esquerda:

- (33) Contaste-lhe que mentira?
- (34) Eles leram-no a todos
- (35) Ele leu-o também

Isso significa que a alternância próclise/ênclise depende do que precede o verbo em sintaxe visível⁶.

4 Esse exemplo é de Duarte e Matos, deve-se considerar contudo a citação de Barbosa acima.

5 Parece aliás que o movimento do verbo tem tendência em desaparecer na língua coloquial nas estruturas interrogativas e focalizadas. Nos termos de Rizzi (1990b) o PE estaria assim perdendo o fenômeno V2 "residual".

6 cf nota 23 para uma proposta de análise.

1.2 O português clássico

Um ponto crucial aqui é que os contextos de próclise obrigatória não sofreram modificações na história da língua, pelo menos na sua versão europeia, desde o português arcaico (PA), cf Ribeiro (1995). A mudança que nos interessa aqui diz respeito aos contextos de ênclise obrigatória. Com efeito, no PCI, um sujeito referencial era compatível tanto com próclise quanto com ênclise. Os dados que seguem são tirados de Torres Moraes (1995, pp. 280- 284) ⁷:

- (36) O amigo Pedro Antonio **me** disse ultimamente que V.M me remetia mais tabaco... (Gusmão)
- (37) ... mas o papa, que presidiu à elaboração dele, **o** consolou, declarando-o ex. Geral. (Gusmão)
- (38) A familiaridade terna que temos com a nossa querida mãe **nos** deu lugar a tratar esse ponto com a delicadeza necessária (Marquesa de Alorna).
- (39) Dom Francisco foi muito bem recebido. El-Rei **lhe** fez muitas honras, e **lhe** deu uma larga audiência; ele **me** escreve muito satisfeito, e eu **o** tenho estimado muito pelos grandes talentos que conheço naquele fidalgo. (Marquês do Lavradio)
- (40) O ar parece-**me** bastante benigno. (Marquês do Lavradio)
- (41) A última carta que recebi de V.Exma deu-**me** uma consolação inexplicável (Marquesa de Alorna)
- (42) E eu, pelo contrário, sustentei-**lhe** sempre que era mentira (Marquesa de Alorna)

Esses exemplos mostram que a ênclise e a próclise parecem alternar livremente, independentemente do sujeito ser pronominal, nominal, longo ou curto, separado ou não do verbo por um adjunto. De fato, nos autores do

⁷ Os autores dos séc 18 e 19 estudados por Torres Moraes são:

(Camilo) Castelo Branco, Camilo, Cartas de Camilo Castelo Branco a Thomaz Ribeiro. (1873-1890),

O morgado de Fafe em Lisboa, comédia em dois atos.

(Judeu) Da Silva, Antonio José. Esopaida ou a vida de Esopo.

De Castro, João Batista, De hora de recreio.

De Gusmão, Alexandre, Cartas (1734-1752)

Dinis, Júlio, As duas cartas. Comédia original em dois atos (1857), Cartas e esboços literários (1865-1870).

Garrett, João Batista de Almeida, Falar verdade a mentir (1845), O Camões do Rocio (1842), Cartas íntimas (1824-1854).

(Marquês) D. Luis de Almeida Portugal, Marquês do Lavradio Cartas da Bahia (1768-1769), Cartas do Rio de Janeiro (1769-1776).

(Marquesa) , D. Leonor de Almeida, Marquesa de Alorna, Inéditos, cartas e outros escritos. (1775-1808)

Oliveira, Francisco Xavier, Cartas familiares.

Verney, Luis Antonio, Verdadeiro método de estudar (1746)

século 18 estudados por Torres Moraes, a próclise é geralmente dominante nesse contexto, mas só exclui a ênclise em um deles, como se vê na seguinte tabela (Torres Moraes, op. cit. p.280):

Autor (ano nasc.)	próclise	enclise	% enclise
Gusmão (1695)	29	0	0%
Castro (1700)	15	1	7%
Oliveira (1702)	39	7	16%
Judeu (1705)	30	6	17%
Verney (1713)	14	11	44%
Marquês (1728)	61	14	19%
Marquesa (1750)	34	23	40%

No começo do século 19, essa relação se inverte e a próclise passa a ser minoritária. Veja-se a tabela II, também baseada em Torres Moraes (op. cit. p. 309):

Autor (ano nasc.)	próclise	ênclise	% ênclise
Garrett (1799)	14	47	77%
Camilo (1825)	6	67	92%
Dinis (1839)	8	70	90%

Encontramos a mesma variação na posição do clítico quando o sintagma antecedendo o verbo não é o sujeito mas um outro elemento da oração, um argumento ou adjunto topicalizado, um advérbio, uma conjunção ou uma oração adjunta, como nos exemplos a seguir:

- (43) A este fim **se** encaminham as minhas diligências (Gusmão)
- (44) e, por isso, a ela **se** aplicavam homens grandes (Verney)
- (45) e em igual desordem **se** acham os corpos auxiliares (Marquês do Lavradio)
- (46) e depois **o** irei repondo por ser dinheiro dotal (Gusmão)
- (47) verdadeiramente **me** achava necessitado (Gusmão)
- (48) e se tal suceder, **a** renegarei deles para sempre (Marquês do Lavradio)
- (49) Aos acentos, **seguem-se** as linhas que se escrevem entre as dicções para as juntar ou dividir na pronúncia (Verney)
- (50) depois segue-**se** a sintaxe (Verney)
- (51) finalmente recolhi-**me** à casa com o mesmo triunfo (Marquesa de Alorna)
- (52) Ao voltar Temístocles a cara, dá-**lhe** Esopo uma cutilada (Judeu)

No século 18, próclise e ênclise alternam com preferência para a próclise na maioria dos autores. Ao contrário, no séc. 19, a tendência à ênclise observada com o sujeito se verifica também com os não sujeitos. No *corpus* estudado por Torres Moraes, só Garrett ainda apresenta casos de próclise nesse contexto⁸, enquanto que todos os outros autores usam a ênclise categoricamente. O português do século 18 encontra-se assim numa fase intermediária entre o português do século 16, que apresenta uma próclise quase que categórica sempre que o verbo é precedido por algum elemento na oração, como mostrado por Lobo (1992) e Martins (1995), e o PE, categoricamente enclítico, como mostram os exemplos acima⁹.

1.3 Uma explicação da alternância e da mudança

É consenso entre os autores que analisaram esses fatos¹⁰ de que a variação entre próclise e ênclise no PCI se deve à disponibilidade de duas posições para o sintagma preverbal, sujeito ou não: uma posição interna à oração e uma posição externa à oração. A próclise corresponde ao primeiro caso, e a ênclise ao segundo. Quanto à natureza dessas posições, também há consenso em considerar que a primeira é o especificador de uma categoria funcional mais alta do que Infl (Comp para muitos), enquanto que a segunda é uma posição de adjunção a essa categoria. Essas duas construções são representadas a seguir, onde XP corresponde a qualquer sintagma pre-verbal, inclusive o sujeito, e F à categoria funcional relevante:

- a) [_{FP} XP [_F cl V] ...]
- b) XP [_{FP} [_F V-cl] ...]

Essa análise permite enunciar de maneira bastante geral e simples a

8 Para uma argumentação no sentido de caracterizar Garrett como um falante de português clássico, cf Galves e Galves (1995).

9 Falta contudo fazer um reparo importante. Nos enunciados acima, o pronome clítico nunca é co-referente do sintagma topicalizado. Quando isso acontece, configurando a construção comumente chamada de deslocamento à esquerda clítico, a ênclise é obrigatória em todas as fases do português quando o verbo segue imediatamente o elemento deslocado; cf Ribeiro (1995) para o PA, e Torres Moraes (1995) para o PCI. Note-se que nessas construções, algum elemento interno à oração, como a negação, um sujeito quantificado ou certos advérbios, pode forçar a próclise. No PCI, ela pode ser desencadeada por qualquer sujeito preverbal. Vejam-se esses dois exemplos de Torres Moraes (op. cit. p. 317)

- i) O sangue dele derramaram-no como água nos vales de Jerusalém (Garrett)
- ii) A quietação em que vivo, ella ma trouxe (Garrett)

10 cf Salvi (1990), Galves (1994), Galves e Galves (1995), Torres Moraes (1995), Benincà (1994, 1995), e Ribeiro (1995) para o mesmo fenômeno em PA.

colocação de clíticos no PCI. Se o verbo está em primeira posição absoluto dentro de FP, a ênclise é obrigatória. Se o verbo não está em primeira posição, a proclise é obrigatória. Note-se que voltamos a uma formulação bem próxima da lei de Tobler Mussafia, só que com uma caracterização mais precisa e abstrata da fronteira da oração, que permite estender a noção de verbo inicial a casos em que um sintagma precede o verbo. Desse ponto de vista, a liberdade do falante não está na colocação dos clíticos mas na posição interna ou externa atribuída ao sintagma preverbal. Note-se também que a representação a), em que o sujeito é interno, é assimilada a uma estrutura V2 na qual o verbo está em Comp e o XP inicial no especificador de Comp.

Dada essa análise da alternância, uma análise da perda da próclise nesses contextos parece impor-se: a posição que, nos sistemas em que as duas construções co-existem, é associada à próclise, deixa de ser disponível, tanto para os sujeitos quanto para outros sintagmas preverbiais tópicos. Essa é a posição de Salvi (1990) quando ele afirma: "A ênclise obrigatória com o sujeito se explica pelo fato de que um sujeito não focalizado não pode achar-se no especificador de C; ele deverá portanto estar deslocado à esquerda... O primeiro elemento no interior de CP será portanto V, e teremos ênclise. A mesma explicação vale para os constituintes não sujeitos antepostos ao verbo mas não focalizados" (op. cit. p.194).

Proporemos a seguir uma formulação dessa mesma idéia no quadro do programa minimalista de Chomsky, acrescido de algumas hipóteses complementares a respeito da ênclise e do papel de Agr nas derivações.

2. A teoria

2.1. O programa minimalista de Chomsky

2.1.1. Traços e verificação.

Um dos aspectos mais inovadores do programa minimalista é o desenvolvimento de uma teoria da verificação¹¹ dos traços. Contrariamente ao modelo anterior em que a morfologia flexional era derivada na sintaxe, os itens lexicais de categoria V, N, A são flexionados no léxico, sendo inseridos na derivação já com seus traços morfológicos. A necessidade de verificação

11 Não traduzirei aqui o termo *checking* por *checagem*, disponível e já em uso na comunidade gerativista brasileira, pela preocupação de trabalhar com um termo de raiz latina, mais próxima das palavras francesas, italianas ou espanholas, do que do original inglês. A tradução mais exata em português é certamente "conferência". Mas talvez seja uma palavra polissêmica demais para ser a mais adequada.

desses traços com as categorias funcionais correspondentes é que força o movimento. Esse movimento pode ser visível se se aplica antes da operação *Spell-Out* que liga a derivação à interpretação fonética, ou invisível se se aplica depois. Na última formulação que Chomsky fez até agora da teoria da verificação (Chomsky 1995b), o movimento é visível quando é forçado por traços chamados fortes. Esses, com efeito, não são tolerados pela derivação, e devem ser apagados o quanto antes. O apagamento resulta da operação de verificação. Chomsky (1995b) define assim os traços fortes:

"Se F é forte, então F é o traço de uma categoria não substantiva e F é verificado por um traço categorial".

Isso significa que os traços fortes são associados às categorias funcionais. Nomes e Verbos não têm. Por exemplo, a categoria funcional Agr terá, numa língua como o francês, um traço-V forte que força o movimento do verbo em sintaxe visível. Já em inglês, Agr não é associado a tal traço, e o verbo não se move para Agr em sintaxe visível¹². O outro traço forte é aquele que Chomsky chama traço-NP ou D, e é responsável pelo movimento dos sintagmas nominais em sintaxe visível. Como enfatizado por Chomsky, o traço-D forte desempenha no Programa Minimalista o papel do Princípio de Projeção Estendido nos modelos anteriores. A posição de verbos e sintagmas nominais na sintaxe visível depende portanto desses traços fortes que estão assim na origem das diferenças entre as línguas, e constituem a versão minimalista dos parâmetros. De maneira interessante, e contrariamente a versões anteriores, os outros traços, como caso e concordância, que não são fortes, não são nunca diretamente responsáveis por diferenças paramétricas.

A noção de traço forte introduz uma primeira dimensão em torno da qual se articula uma tipologia dos traços. Uma outra dimensão é a noção de *interpretabilidade*. Certos traços são interpretáveis em Forma Lógica, outros não são. Nos primeiros, Chomsky conta os traços categoriais e os traços- ϕ dos nominais, nos segundos, o caso dos nominais, e os traços- ϕ dos verbos e dos adjetivos, bem como todos os traços das categorias funcionais: traços fortes, traços- ϕ , traços atribuidores de caso¹³.

12 Vale ressaltar que também não se move em sintaxe invisível, uma vez que nesse modelo, o contrário de traço forte não é traço fraco, como no primeiro minimalismo, mas ausência de traço. Isso é possível porque os traços categoriais são interpretáveis (ver mais abaixo), e por isso não necessitam entrar numa relação de verificação.

13 Chomsky nota que "o caso difere dos traços-f pelo fato de ser sempre [-interpretável], para ambos os termos da relação de verificação. O caso é assim o traço formal por excelência, e não é surpreendente que essa linha de pesquisa tenha as suas origens no filtro do caso de Vergnaud."

Do ponto de vista da teoria da verificação, a interpretabilidade tem duas consequências para um traço. Por um lado, ele não precisa entrar em nenhuma relação, já que não precisa ser apagado, uma vez que é um objeto interpretável na interface. Mas Chomsky argumenta também que, não precisando ser apagado, pode entrar em mais de uma relação de verificação¹⁴. Aqui, só o primeiro aspecto será relevante.

Enfim, toda verificação implica que o elemento a ser checado esteja em algum ponto da derivação no *domínio de checagem* da categoria verificadora. Esse domínio corresponde a duas configurações:

- Adjunção a um núcleo para as projeções mínimas (heads).
- Posição de especificador de um núcleo para as projeções máximas.

2.1.2 Mover- α e Mover-traços

No programa minimalista, movimento e verificação são intimamente ligados, uma vez que o primeiro é provocado pela necessidade da segunda. Chomsky (1995b) enfatiza essa relação ao propor que depois de Spell-Out só os traços se movem, deixando no lugar a categoria e o item lexical ao qual estão associados. Segundo ele, só não é assim antes de Spell-Out porque as regras de interpretação fonética interpretam os itens lexicais, e não os traços. Esses devem portanto arrastar com eles toda a categoria. Mas depois de Spell-Out, essa necessidade desaparece, e só os traços têm razões de se mover. Esse movimento corresponde então sempre a uma adjunção a um núcleo. Em outras palavras, o movimento abstrato nunca cria posições de especificador.

2.2 Elementos para uma teoria da concordância

2.2.1 Agr: conjunto de traços- ϕ pivô da predicação

Pollock (1989) representa um marco essencial na teoria das categorias tal como se define no primeiro minimalismo. É nesse artigo, com efeito, que Agr passa a ter estatuto de categoria, nascida, junto com Tempo (doravante T), da explosão da categoria Infl. Uma vez separadas, essas duas categorias devem receber uma posição respectiva na estrutura. Começam então longas discussões para se definir se T domina Agr, como propõe o próprio Pollock no seu artigo, ou se Agr domina T, como foi proposto, com bastante

14 É o que acontece, por exemplo, na relação de concordância do nome com o verbo e o adjetivo numa frase como "As meninas são bonitas", onde o DP *meninas* entra com uma relação de verificação primeiro com o adjetivo e depois com o verbo.

êxito para as línguas românicas, por Belletti (1990). Chomsky (1989) resolve o debate propondo uma estrutura oracional na qual T ao mesmo tempo domina e é dominado por Agr. Mas este segundo Agr, mais baixo do que T, é de natureza diferente do Agr que domina T: os seus traços correspondem não ao sujeito, mas ao objeto. Chomsky propõe assim a distinção entre Agr_o e Agr_s. É o ponto de partida de uma inflação de Agrs. Agr se torna uma categoria incontornável, que aparece também internamente ao DP, e que se multiplica na oração¹⁵.

O primeiro texto minimalista (Chomsky, 1993) continua nesse sentido ao considerar que, além da relação de concordância, toda verificação casual implica a intermediação de uma categoria Agr. Contudo, Chomsky opera uma reviravolta importante na última seção do último capítulo do seu livro de 1995. Baseado na idéia de que o sistema computacional só contém categorias interpretáveis nas interfaces, Chomsky denega a Agr o estatuto de categoria. Com efeito, a presença ou ausência de marcas de concordância carece de correlato semântico nas línguas. A noção de concordância volta assim a ser simplesmente relacional, em contradição com o que se assumia desde Pollock (1989).

Entre uma noção puramente relacional, e a definição de uma categoria sempre que uma relação de concordância se estabelece, existe uma terceira via, já subjacente nos modelos anteriores, mas que parece particularmente pertinente explicitar no quadro minimalista. Consiste em definir Agr, antes de mais nada, como um conjunto de traços-f. Esses traços-f podem constituir o conteúdo de uma categoria funcional independente, ou ao contrário estar associados a uma outra categoria funcional¹⁶. De fato, essa proposta já está em Rizzi (1990a, p. 52) quando afirma:

"Continuamos a assumir que Agr pode ser tanto um núcleo independente com a sua projeção flexional autônoma (AgrP) quanto um traço ou um conjunto de traços atribuído a um outro núcleo."

2.2.2 A parametrização de Agr

A partir dessa concepção de Agr, trabalharemos com a hipótese de que os sistemas de concordância das línguas diferem em relação a dois parâmetros:

15 Veja-se por exemplo o artigo de Cardinaletti e Roberts (1991), que teve um grande impacto sobre a análise do português (cf Ribeiro 1995, Torres Morães 1995, Galves 1991, 1992a, b).

16 Na secção seguinte, bem como em Galves (em progresso), também se propõe que esse conjunto de traços-f pode ser gerado em posição argumental. Seria o caso dos pronomes fracos em geral. A possibilidade de ser tanto functor quanto argumento deriva provavelmente da especificidade de Agr em relação às outras categorias funcionais.

- a associação de traços- ϕ (Agr) às diversas categorias funcionais, e em particular a Comp.

- a associação de um traço-V forte a esses traços- ϕ .

A análise proposta mais abaixo visará mostrar que a diferença paramétrica entre PCI e PE diz respeito ao traço-V dos traços- ϕ associados a Comp: presente em PCI, ausente no PE.

2.3 A teoria dos clíticos

Muita literatura gerativista tem sido devotada aos pronomes clíticos. Questões como sua natureza categorial, a razão pela qual se movem, para onde se movem, a sua relação com outros aspectos da sintaxe das línguas, a natureza da ênclise, têm sido debatidas sem que se chegue verdadeiramente a um consenso (cf Uriagereka 1992, 1995, Corver e Delfitto 1993, Cardinaletti e Starke 1993, entre outros).

A análise proposta aqui se baseia nas três hipóteses seguintes¹⁷.

- 1) Os clíticos são feixes de traços pronominais (traços- ϕ), ou seja, em termos categoriais, correspondem à categoria Agr.
- 2) Eles podem ser gerados diretamente em posições argumentais, sofrendo adjunção ao verbo em seguida, e nesse caso, temos a próclise. Eles podem também ser morfologicamente associado ao verbo antes do início da derivação, e temos então a ênclise.
- 3) A ênclise no tempo finito põe em jogo a categoria Comp.

Justificaremos agora cada uma dessas hipóteses.

A primeira se opõe à idéia bastante comum na literatura de que o clítico é um determinante (D). Apesar de ser atrativa, em particular por causa da identidade morfológica do clítico de 3a pessoa e do artigo nas línguas românicas, essa hipótese não deixa de ser problemática. Se D é por excelência a categoria associada à referência, como dar conta da ausência de qualquer interpretação referencial associada aos clíticos em construções como as do francês exemplificadas a seguir onde o pronome clítico *le* está no lugar de uma proposição ou de um adjetivo?

(53) Je le sais
Eu o sei

(54) Belle, elle ne l'a jamais été
Bela, ela neg o aux nunca sido
"Bela, ela nunca foi"

17 Essas hipóteses são discutidos de maneira mais detalhada e aprofundada em Galves (em progresso), onde também se argumenta que os clíticos se movem em sintaxe visível porque são dotados de um traço-V forte.

Ao contrário, se admitirmos que eles funcionam como elementos de concordância, marcando simplesmente uma posição com alguns traços gramaticais que permitem recuperar um argumento ou um predicado antecedente, independentemente de qualquer interpretação referencial, todo o seu comportamento se torna coerente¹⁸. Note-se que a assimilação dos clíticos a Agr vai no sentido da posição de alguns autores, como Cardinaletti e Starke (1994) e Dobrovie-Sorin (1994) que os identificam como Infl. Observe-se enfim que assimilar os clíticos a Agr torna mais coerente a segunda hipótese que considera a ênclise como um caso de morfologia flexional.

Consideremos agora essa hipótese. Ela constitui uma formulação, no sistema minimalista, da afirmação de Benincà e Cinque (1993) de que, na ênclise, o verbo e o pronome clítico formam uma unidade morfológica. Além de mencionarem inicialmente o fato das línguas românicas distinguirem ortograficamente ênclise e próclise, sendo que na primeira o clítico vem colado ao verbo ou ligado por um travessão, os autores mostram como os fatos de coordenação distinguem nitidamente os dois tipos de colocações. Vejam-se os seguintes fatos do francês:

- (55) Il chantera et dansera avec nous
Ele cantará e danará conosco
- (56) * Chantera-t-il et dansera avec nous?
Cantará-ele e danará conosco
- (57) Jean le lit et relit sans cesse
Jean o lê e relê sem parar
- (58) * Lis et relis-la!
Lê e relê-a

Quando o pronome é proclítico, é possível que, numa estrutura coordenada, somente a sua primeira ocorrência seja expressa. Quando é enclítico, é impossível. Da mesma maneira, em certas línguas como o rumeno, dois clíticos podem ser coordenados quando estão em próclise, mas nunca quando estão em ênclise:

- (59) Imi si îți scrie
Me e te escreve
- (60) *Dîndu-mi si îi cartea
Dá-me e lhe um livro

18 Desse ponto de vista, os clíticos problemáticos para análises como a de Uriagereka (1995), são na realidade protótipos no sentido que eles fazem aparecer o caráter intrinsecamente não referencial desses elementos.

Benincà e Cinque notam por outro lado que a ênclise provoca em certas línguas modificações acentuais da palavra, mas que isso não acontece aparentemente nunca com a próclise.

Todos esses fatos concorrem a mostrar que *verbo + clítico* e *clítico + verbo* não constituem o mesmo objeto, o clítico e o verbo funcionando como uma unidade morfológica indissociável na ênclise mas não na próclise. Há evidentemente várias maneiras de interpretar esse fato. A interpretação proposta aqui é a mais coerente com o espírito minimalista que considera que toda a morfologia flexional já vem inscrita na palavra no início da derivação, a sua interferência com a estrutura sintática sendo tratada como o efeito do jogo da verificação. Note também que é a única maneira de distinguir morfológica e sintaticamente próclise e ênclise se admitirmos que nos dois casos, o clítico e o verbo estão na mesma categoria em sintaxe visível. A próclise é o resultado de uma adjunção do clítico ao verbo. A ênclise corresponde a uma palavra já formada no léxico¹⁹, inserida na derivação como os traços de concordância, e tendo o mesmo estatuto que eles em relação à teoria da verificação.

Ainda do ponto de vista dessa teoria, a distinção operada aqui entre ênclise e próclise pode ser apreendida em termos de interpretabilidade. O clítico independente que aparece nas construções proclíticas funciona como um pronome independente. Parece plausível pensar que esses são interpretáveis, uma vez que se comportam como os nomes quanto à concordância, podendo entrar em relações múltiplas, ou prescindindo a existência de tal relação²⁰. O clítico morfológico, por sua vez, funciona como um elemento de concordância verbal, e enquanto tal, não é interpretável. As suas diferenças de legitimação, isso é, no sistema minimalista, de verificação, decorrem dessa diferença quanto à interpretabilidade. Os traços- ϕ do clítico em próclise não precisam ser verificados porque eles são interpretáveis, os traços- ϕ do clítico em ênclise precisam ser verificados porque eles não são interpretáveis. Assim se explicam as maiores restrições se aplicando à construção enclítica. Veremos a seguir como esse contraste se articula com outras propriedades do português para produzir as distribuições da ênclise e da próclise descritas acima.

19 Benincà (1994, p. 213) cita a primeira frase de Mussafia (1886): "Os pronomes pessoais oblíquos *mi, ti, si, ecc.*, e as partículas pronominais átonas *ci, vi, ne*, ou precedem o verbo de forma finita (proclise) ou vêm depois dele, formando com ele uma única palavra (ênclise)". Benincà acrescenta: "Parece-me que não temos ainda uma maneira de dar conta de intuições desse tipo". É justamente o que a análise proposta aqui se propõe fazer.

20 Note-se que essa interpretabilidade pode estar relacionada com o fato de que os pronomes, como os nomes, são portadores de função temática.

Enfim, a terceira hipótese se baseia em um conjunto robusto de fatos:

- a ênclise interfere com processos sintáticos que são tradicionalmente analisados como pondo a categoria Comp em jogo: complementação oracional, interrogação, focalização.

- outros fenômenos específicos do PE se explicam também por uma caracterização particular de Comp: objeto nulo (Raposo 1986, Galves em progresso), infinitivo flexionado (Raposo 1987, Galves 1992a,b), posição do verbo nas orações interrogativas (Ambar 1992).

- a evolução do PE em relação ao PCI faz aparecer nitidamente que a mudança da colocação de clíticos se acompanha de uma mudança na sintaxe do verbo no que diz respeito ao chamado fenômeno V2, que por sua vez envolve uma parametrização da categoria Comp.

Essa análise afirma assim que as categorias Σ (Martins 1995), W (Rouveret 1992), ou F (Uriagereka 1992, Raposo 1995) não são senão um Comp associado a uma matriz de traços- ϕ , ou seja, a um elemento de concordância, Agr na terminologia corrente²¹. Nas orações subordinadas, esse Comp pode aparecer num contexto de topicalização encaixada (cf. 11).

A associação de Agr a Comp faz com que a projeção de Comp seja interpretada como um predicado complexo. As línguas que têm essa particularidade têm portanto um certo funcionamento de língua de tópico, uma vez que além da relação sujeito/predicado que se dá em IP, elas têm a possibilidade de instanciar uma relação de predicação a nível de CP. A realização particular dessa relação nas línguas depende da maneira como se dá a verificação de Agr, que por sua vez depende dos traços associados a esse Agr. Veremos que o PCI e o PE instanciam justamente dois tipos distintos de línguas com concordância em Comp. Na primeira, o verbo se move para Comp, e temos uma língua de tipo V2, na segunda, o verbo não se move para Comp, e temos efeitos residuais de V2, a ênclise, bem como efeitos de língua de tópico, o objeto nulo²².

Passaremos agora à análise propriamente dita.

21 A idéia de Agr em Comp vem dos estudos sobre as línguas V2. Para uma boa revisão dessa literatura, ver Ribeiro (1995) e Torres Moraes (1995).

22 Rouveret (1987) foi o primeiro a perceber esse funcionamento da oração no PE e a correlação entre ênclise e objeto nulo, com a dificuldade porém de dar conta do fato de que a distribuição do objeto nulo e a distribuição da ênclise não coincidem exatamente. Para uma discussão dessa correlação e da distribuição, ver Galves (em progresso).

3. A análise

3.1 A derivação no PE

Fenomenologicamente, o PE não é uma língua V2. Na presença de um elemento inicial não sujeito, este aparece em segunda posição e o verbo em terceira posição. Temos assim (61) e não (62), que corresponderia à ordem canônica numa língua V2 como o alemão.

- (61) Ontem o João acabou o trabalho
(62) *Ontem acabou o João o trabalho

Teoricamente, isso significa que não é uma língua na qual o verbo se move para Comp. Propusemos acima derivar o movimento do verbo para Comp da presença de um traço-V forte associado aos traços-f de Comp. Em PE, portanto, os traços- ϕ de Comp não têm traço-V forte. Mas depois de Spell-Out, esses traços precisam entrar numa relação de verificação com outros traços- ϕ . Nas estruturas enclíticas, são os traços do clítico que efetuam essa verificação, na qual eles acham sua própria legitimação, uma vez que eles também precisam ser verificados. Essa operação mútua de verificação pode ser representada como uma co-indexação²³:

- (63) [_{CP} C_i [_{AgrP} NP V-cl_i]]

Já se mencionou acima que a presença de traços- ϕ em Comp implica que CP seja interpretado como um predicado complexo, tendo como sujeito o tópico ao qual referem os traços do clítico²⁴. Nessa concepção, os traços- ϕ e os traços operadores, QU ou Focus ou Negação, estão necessariamente em distribuição complementar em Comp. Com efeito, quando CP domina uma relação operador/variável, ele funciona como um domínio saturado, e não pode mais funcionar como predicado complexo. Nesse caso, Agr não pode estar associado a C, e a ênclise não pode ser legitimada. A estrutura da frase não implica portanto nenhuma operação de verificação de traços- ϕ

23 O recurso da co-indexação para notar a relação estabelecida entre as duas posições pelo movimento dos traços do clítico nos permite representar a idéia de que a relação entre o clítico e Comp é visível em Forma Lógica apesar da operação de verificação ter apagado os traços- ϕ em Comp.

24 O leitor atento estará se perguntando como Agr em C é verificado quando não há clíticos. Nesse caso, minha hipótese é que o PE instancia uma estrutura próxima da topicalização, em que o sujeito é gerado externamente a CP, que por sua vez é verificado pelos traços de um *pro* sujeito. Não aprofundarei aqui essa análise, nem as suas conseqüências sobre as diferenças entre frases com clíticos e frases sem clíticos. Note-se simplesmente que a idéia de que a estrutura das frases com clítico é diferente da estrutura das frases sem clíticos já está em Madeira (1992) e Manzini (1993). cf Galves (em progresso).

em Comp. Uma frase interrogativa com próclise como (15) terá assim a seguinte estrutura, na qual a relação de coindexação se dá entre o traço-QU em Comp e o especificador de Comp:

(64) [_{CP} Que livros_i Q_i [_{AgrP} A Maria [lhe [deu]]]]

A próclise acontece quando Comp participa de uma operação sobre a sentença, inscrita na estrutura antes ou depois de Spell-Out. Antes de Spell-Out, temos as estruturas em que um elemento interrogativo (15) ou focalizado (23), (25), (27), (29), (31) se deslocou para Spec/Comp. Depois de Spell-Out, temos o movimento dos traços de negação (17) e dos traços lógicos associados aos quantificadores (19), (21)²⁵. Quanto às encaixadas, elas não podem ter um Comp associado a Agr, uma vez que a relação de concordância que esse Agr implica é local. A próclise é portanto obrigatória, a não ser que seja inserido um tópico imediatamente antes da oração subordinada. Veja-se o contraste entre (11) e (12), repetidas aqui como (65) e (66), e (67) e (68), onde o tópico é gerado no início absoluto da frase, e onde se instancia a distribuição oposta da ênclise e da proclise:

- (65) Eu sei que, a Maria, o João viu-a ontem
 (66) * Eu sei que, a Maria, o João a viu ontem
 (67) * A Maria, eu sei que o João viu-a ontem
 (68) A Maria, eu sei que o João a viu ontem

Resumindo, eis as configurações lícitas e ilícitas no PE. O contraste básico é o seguinte:

- a) [_{CP} ϕ [_{AgrP} NP V-cl]]
 b) * [_{CP} X_{op} [_{AgrP} NP V-cl]]

25 Esta análise implica a ausência de movimento oculto nas orações em que o elemento normalmente desencadeador de próclise se encontra depois do verbo, e a ênclise é obrigatória, como nas frases mencionadas acima e repetidas a seguir:

- i) Contaste-lhe que mentira?
 ii) Eles leram-no a todos
 iii) Ele leu-o também

A ênclise nesses casos aponta para a presença de Agr em Comp. Isso significa que os sintagmas interrogativos e focalizados pós-verbais não são verificados via um movimento de traço para Comp na sintaxe invisível. A ausência de traço-QU em Comp em i) corresponde a uma leitura D-linked (Pesetsky 1987), ou seja referencial, do sintagma-QU, que não implica uma relação operador/variável. Quanto a ii) e iii), *também* e *a todos* não têm escopo sobre a oração inteira. Duas análises são possíveis. A primeira, baseada em May (1985) consiste em postular um movimento de alçamento de quantificador (QR) para uma posição mais baixa, por exemplo, em adjunção a VP. Na segunda, não há movimento nenhum, como em i). Esta é certamente mais fiel ao espírito do minimalismo.

Em a), os traços- ϕ de Comp são verificados pelos traços carregados pelo verbo. Em b), ao contrário, existe uma relação de verificação com algum elemento operador interna a Comp, e os traços do clítico no verbo não têm verificação possível. A derivação não converge.

c) e d) instanciam o contraste complementar:

c) [_{CP} X op [_{AgrP} NP cl V]]

d) * [_{CP} ϕ [_{AgrP} NP cl V]]

Em c), Comp não tem traços- ϕ , e a derivação converge porque o clítico, sendo interpretável, não precisa de verificação. d) demanda uma explicação suplementar. Com efeito, temos que dar conta do fato de que o clítico em próclise não é capaz de participar de uma relação de verificação com Agr em Comp. O fato dele ser interpretável não é uma razão suficiente uma vez que se a interpretabilidade não força a relação de verificação, também não constitui impedimento à essa relação. Ao contrário, permite que se dê varias vezes. A razão pela qual a próclise é excluída na estrutura d) tem que ser portanto procurada algures. Farei tentativamente a hipótese que os traços do clítico em próclise são invisíveis para o sistema de computação e que são portanto incapazes de se moverem para Comp. Essa invisibilidade é devida à posição preverbal, isto é, interna à palavra, dos traços do pronome no núcleo complexo formado pelo clítico e o verbo. Ora, o sistema de computação, por hipótese, não "vê" no interior das palavras.

3.2 A derivação no PCI

A diferença fundamental entre PE e PCI reside no fato de que o segundo, contrariamente ao primeiro, é uma língua com características V2 como mostrado por estes exemplos tirados de autores do séc. 18 (Torres Moraes 1995, pp.244-245).

- (69) A estas últimas chamam os médicos doenças contra a natureza (Oliveira)
- (70) No mesmo instante fez o governador o sinal (autor desconhecido)
- (71) Com a língua faz o arrieiro a célebre cantiga (Judeu)

A natureza exata dessa propriedade não está inteiramente clara; um dos problemas para a teoria sendo a relação que existe nessas línguas entre o movimento do verbo e o movimento do sintagma aparecendo em primeira posição. No sistema minimalista, o problema se coloca ainda mais, uma vez que os traços provocando os movimentos são a priori independentes. Dentro do quadro analítico desenvolvido aqui, propõe-se que as línguas V2 são as línguas que contêm duas propriedades: elas associam traços- ϕ a Comp,

e esses traços-f têm um traço-V forte. Deve-se admitir por outro lado que o movimento do verbo para Comp, devido ao traço-V forte de Comp, faz com que Spec/Comp seja potencialmente uma posição de sujeito saturando o elemento de concordância nuclear de CP²⁶. Ao contrário, numa língua como o PE, em que o verbo não vai para Comp, a concordância em Comp só pode ser satisfeita por um movimento de traços posterior a Spell-Out.

Articulando essa análise com a teoria da ênclise proposta acima, é agora possível entender porque se o sujeito ou um outro sintagma da oração ocupa Spec/CP, a ênclise é ilícita. É porque, nesse caso, Agr é verificado pelo especificador de Comp, e não há verificação possível do clítico. Numa língua V2, os traços- ϕ de Comp podem ser verificados pelo sujeito ou por um outro sintagma da oração. Em ambos os casos, a ênclise é excluída, uma vez que o clítico não tem nenhum elemento disponível para sua verificação. A próclise é portanto obrigatória. Se nenhum sintagma se move para a posição de especificador de Comp, os traços- ϕ de Comp só podem ser verificados pelo clítico, a ênclise é obrigatória. Deriva-se assim a complementariedade das estruturas enclíticas com as estruturas em que Spec/CP está preenchido, representada de novo a seguir (com a co-indexação marcando as relações de verificação):

- a) [_{CP} XP_i [_{Cl} cl V] ...]
- b) XP [_{CP} [_{Cl} V-cl] ...]
- c) * [_{CP} XP_i [_{Cl} V-cl] ...]

De fato, especificador e clítico competem para a mesma relação de verificação envolvendo os traços-f de Comp. Nessas condições, um sintagma preverbal, sujeito ou não, não pode estar no domínio de verificação desses traços numa construção enclítica. Isso significa que deve estar fora da fronteira de oração, como representado acima.

A hipótese do traço-V forte de Agr em Comp, associada à teoria dos ênclise e da próclise proposta acima permite derivar a análise do PCI proposta por Salvi (1990) e Benincà (1995). Afastamo-nos porém da análise desses autores no que diz respeito à mudança.

26 Lembre-se que o movimento de uma projeção máxima para uma posição de especificador não pode ser desencadeado pelos traços de concordância, por não serem traços fortes, mas somente por um traço-D. Ora parece que um traço-D só pode ser legitimamente associado a Agr em Comp quando o verbo se move para lá. Uma interpretação possível desse fato é a de que o movimento de V torna Comp uma projeção estendida do verbo, e que isso é uma condição para que um traço-D seja associado a seus traços- ϕ .

3.3 A mudança

Tanto Salvi quanto Benincà afirmam que no PE, o sujeito referencial está sempre na posição externa a CP, o que explica a ênclise obrigatória. Para eles, a mudança do PCI para o PE consiste portanto simplesmente na perda da disponibilidade de Spec/CP para os não operadores. Das duas posições possíveis para o sintagma preverbal tópico em PCI, só sobra uma no PE.

A análise proposta aqui, ao contrário, considera que a posição do sujeito com ênclise no PE é diferente da posição do sujeito com ênclise no PCI, como representado pelas duas estruturas abaixo:

PCI: NP [_{CP} V-cl [_{AgrP}]]

PE: [_{CP} [_{AgrP} NP V-cl]]

Como Salvi, atribui-se aqui a origem da mudança à perda do fenômeno V2, ou seja, à perda do movimento do verbo para Comp. Por hipótese, Agr em Comp só pode então ser verificado por um movimento de traços posterior a Spell-Out. A posição de especificador de Comp deixa de ser disponível para os não operadores. Mas isso não significa que todo sintagma preverbal tenha que ser externo a CP. Para o sujeito, existe uma posição interna a CP disponível, o especificador de AgrP²⁷.

A pergunta fundamental à qual uma teoria da mudança se deve responder é como e porquê acontece a passagem de um sistema para o outro. Qual é a razão da perda de V2 em português? Esse assunto foi bastante discutido na literatura recente a respeito da história do francês (Adams 1987, Roberts 1993, entre outros). Como observado por Torres Moraes (1995), as condições nas quais se deu a mudança são muito diferentes em português. Em particular, não está ligada a um aumento das estruturas V3, que aparecem mais como uma consequência do que como uma razão da mudança. O aumento das construções com ênclise, que se dá ao longo do século 18, com um salto qualitativo no início do sec. 19, pode, ao contrário, ter desempenhado um papel crucial. Galves e Galves (1995) arguem que esse aumento está relacionado com a implementação de uma mudança prosódica que favorece o uso da ênclise sobre a próclise. A mudança gramatical só se dá porém quando uma geração de crianças reanalisa as estruturas com ênclise como tendo um sujeito não mais externo mas interno a CP²⁸. Nesse

27 Para os não sujeitos, evidentemente, essa posição não é lícita.

28 Para Galves e Galves (1995) a atribuição da estrutura às sequências morfológicas às quais as crianças são expostas, e em particular, a definição das fronteiras da oração, estão intimamente relacionadas à percepção pelas crianças do padrão prosódico da língua que estão aprendendo.

momento, dado que o preenchimento da posição Spec/CP é incompatível com a ênclise, a única posição lícita para o sujeito é Spec/AgrP. A posição do verbo é automaticamente afeta por essa reanálise. Ele não pode mais ser localizado em Comp, uma vez que ele segue um sujeito mais baixo que Comp. A consequência paramétrica da nova estrutura associada pelas crianças às construções Sujeito Verbo clítico é portanto a perda de V2, isto é a perda do traço-V forte de Comp.

3.5 Observações finais

A análise da ênclise proposta aqui apresenta várias vantagens em relação às demais:

- Ela representa formalmente a unidade morfológica mais forte presente na ênclise, mesmo no âmbito de uma concepção da cliticização na qual o verbo e o clítico estão sempre no mesmo núcleo.

- Para derivar por movimento estruturas enclíticas em que verbo e clítico estão no mesmo núcleo, duas análises são possíveis. Ou o clítico se adjunge à direita do verbo, como proposto por Uriagereka (1992, 1995), ou o verbo se adjunge à esquerda do clítico (Madeira 1992, Manzini 1992, 1994). A primeira solução viola o princípio de adjunção à esquerda de Kayne (1991) e levanta a questão delicada de saber em que contextos esse princípio deve ser relaxado. A segunda solução levanta a questão de saber porque o clítico atrai o verbo. Manzini (1994) reconhece que a natureza das checagens envolvidas não está clara: " Under the general framework of Chomsky (1992), however, movement of V+I to Cl takes place in order to check some feature of V+I against Cl; while movement of Cl to V+I is to check some feature of Cl against V+I. Possibly the same feature is involved in both instances of movement, say Case; but as far as I can see, nothing hinges on a precise implementation of my proposal in this respect." (op. cit., p.303)

- A distribuição complementar entre ênclise e presença de traços operadores em Comp não implica considerar o clítico como operador como faz Manzini (1992, 1994). Além de ser intuitivamente indesejável, dada a natureza referencial dos clíticos, há uma falha no paralelismo: a presença de clítico em Comp (que corresponde para ela à ênclise) bloqueia a checagem dos outros traços, mas os outros traços não bloqueiam a checagem do clítico quando ele está em próclise (cf op. cit. p.310).

- O modelo de Chomsky (1995), com o papel preponderante que dá à verificação dos traços, e sua concepção do movimento invisível como movimento de traços para um núcleo, traz grandes vantagens para a análise da distribuição da ênclise e da próclise no PE. Com efeito, a principal dificulda-

de consistia até agora em atribuir a elementos que têm realizações superficiais bastante diversas nos enunciados um estatuto idêntico, de maneira a propor uma análise generalizável a todos esses elementos. Pode-se agora incluir na mesma análise contextos em que se vê o movimento, como é o caso das interrogativas, e contextos em que não parece haver movimento, como é o caso da negação e dos sujeitos quantificados, sem forçar o movimento das categorias para posições de especificadores antes ou depois de Spell-Out. Todo movimento invisível sendo um simples movimento de traços, a representação em Forma Lógica fica consideravelmente simplificada.

Résumé

Cet article décrit et analyse le changement survenu dans le placement des clitiques en portugais européen au 19^e siècle. Dans les propositions affirmatives racines à sujet référentiel, aussi bien l'enclise que la proclise étaient possibles en portugais classique (PCI). Dans le même contexte, seule l'enclise est permise en portugais moderne (PE). On propose une analyse dans le cadre du Programme Minimaliste de Chomsky, basée sur l'idée que les formes enclitiques correspondent à des mots formés dans le lexique tandis que les formes proclitiques sont dérivées en syntaxe. Les formes enclitiques sont alors sujettes à un processus de vérification de traits qui implique l'existence de traits correspondants dans un autre endroit de la structure. Aussi bien en PCI qu'en PE, c'est la catégorie Comp qui peut contenir des traits d'accord capables de vérifier les traits du clitique en enclise. La différence entre les deux grammaires est que le verbe se déplace en Comp dans la première, mais pas dans la seconde. Cette affirmation est corroborée par le caractère "V2" du portugais classique.

References

- ADAMS, M. (1987) **Old French, null subjects and verb second phenomena**. Tese de doutorado, UCLA.
- AMBAR, M. (1992) **Para uma sintaxe da inversão sujeito/verbo em português**, Lisboa: Edições Colibri.
- BARBOSA, P. (1991) Clitic Placement in EP. Mimeo, MIT.
- BARBOSA, P. (1993) Clitic placement in Old Romance and European Portuguese and the null subject parameter. Mimeo, MIT.
- BELLETTI, A. (1990) **Generalized Verb Movement: Aspects of Verb Syntax**, Turin: Rosenberg and Sellier.
- BENINCA, P. (1994) **La variazione sintattica (Studi di dialettologia romanza)**, Bologna: Il Mulino.
- BENINCA, P. (1995) Complement clitics in Medieval Portuguese and the Tobler-Mussafia Law. In A. Battye and I. Roberts, (eds) **Language Change and Verbal Systems**.
- BENINCA, P., CINQUE, G. (1993). Su alcune differenze fra enclisi e proclisi, in **Omaggio a Gianfranco Folena**, Padova: Editoriale Programma.
- CARDINALETTI, A., STARKE, M. (1993-1994). The typology of structural deficiency, on the three grammatical classes. Mimeo, Università di Venezia/ Université de Genève.

- CARDINALETTI, A., ROBERTS, I. (1991) Clause structure and X-second, a sair em W. Chao e G. Horrocks (eds) **Levels of Representation**, Berlin: Mouton.
- CHOMSKY, N. (1989) Notes on economy of derivation. In I. Laka and A. Mahajan (eds) **Functional heads and clause structure**, **MIT Working Papers in Linguistics**, vol.10.
- CHOMSKY, N. (1993) A Minimalist Program for Linguistic Theory. In K. Hale and S. J. Keyser (eds) **The view from Building 20**. Cambridge, MA: MIT Press.
- CHOMSKY, N. (1995a) Bare Phrase-Structure. In G. Webelhuth (ed.) **Government and binding theory and the minimalist program**. Oxford, UK: Blackwell .
- CHOMSKY, N. (1995b) **The Minimalist Program**. Cambridge, MA: MIT Press.
- CORVER N., DELFITTO D. (1993) Feature assymetry and the nature of pronoun movement. Mimeo, Tilburg University /University of Utrecht.
- DOBROVIE-SORIN, C. (1994) The typology of pronouns and the distinction between syntax and morphophonology. Mimeo, Université Paris 7.
- DUARTE, I., MATOS, G. (1995) Romance Clitics and the Minimalist Program, comunicação apresentada no **5º Colóquio Internacional de Gramática Gerativa**, La Coruña.
- GALVES, A., GALVES, C. (1995) A case study of prosody driven language change. From CIP to EP. Mimeo, Unicamp/USP.
- GALVES, C. (1992a) Enclise, Infinitif fléchi et extraction du sujet : évidence pour deux accords en portugais européen, **comunicação no 3º Colóquio Internacional de Gramática Gerativa**, Vitoria (Espanha).
- GALVES, C. (1992b) Clitic placement in European Portuguese. Evidence for a non-homogeneous theory of enclisis. **Atas do Workshop sobre o Português**, Associação Portuguesa de Linguística, Lisboa.
- GALVES, C. (1994) Clitic-placement and Parametric Changes in Portuguese, comunicação apresentada no **24th Linguistic Symposium Romance Languages**, UCLA/USC A sair em M. Saltarelli and M.L. Zubizarreta (eds) **Selected Papers from the 24th Linguistic Symposium Romance Languages**. Georgetown University Press.
- GALVES, C. (em progresso) **Gramáticas minimalistas do português**.
- KAYNE, R. (1991) Romance Clitics, Verb Movement, and PRO. **Linguistic Inquiry** 22, 647-686.
- LOBO, T. (1992) **A colocação do clíticos em português. Duas sincronias em confronto**. Dissertação de mestrado, Ufba.
- MADEIRA, A. M. (1992) On clitic placement in European portuguese. In H. van de Koot (ed.) **UCL Working Papers in Linguistics** 4. University College London.
- MANZINI, M. R. (1992) Second position dependencies. comunicação apresentada no **8º Workshop on Germanic Syntax**, Tromso.
- MANZINI, M. R. (1994) Triggers for verb-second: Germanic and Romance. **The Linguistic Review** 11, 299-314.
- MARTINS, A.M. (1992) Quantifiers and clitics in European Portuguese. Mimeo, Universidade de Lisboa.
- MARTINS, A.M. (1993) Clitic-placement from old to modern portuguese. Comunicação apresentada na **11th International Conference on Historical Linguistics**, Los Angeles.
- MARTINS, A.M. (1995) **Clíticos na história do português**. Tese de doutorado, Lisboa.
- MAY, R. (1985) **Logical form**. Cambridge, MA: MIT Press.
- MUSSAFIA, A., 1886. Una particolarità sintattica della lingua italiana dei primi secoli. In **Miscellanea di filologia e linguistica**, Firenze.
- PESETSKY, D. (1987) Wh-in-situ: movement and unselective binding. In E. Reuland and A. ter Meulen (eds) **The representation of (in)definiteness**. Cambridge, MA: MIT Press.

- POLLOCK, J.Y. (1989) Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20, 365-424.
- RAPOSO, E. (1986) On the null object in European Portuguese. In O. Jaeggli e C. Silva-Corvalan (orgs) **Studies in Romance Linguistics**. Dordrecht: Foris.
- RAPOSO, E. (1987) Case Theory and Infl-to-COMP - The inflected infinitive in European Portuguese. *Linguistic Inquiry* 18, 85-109 .
- RAPOSO, E. (1995) Clitic position and Verb Movement in European Portuguese, conferência no 5º **Colóquio Internacional de Gramática Gerativa**, La Coruña.
- RIBEIRO, I. M. (1995) **A sintaxe da ordem no português arcaico: o efeito V2**. Tese de doutorado, Unicamp.
- RIZZI, L. (1990a) **Relativized minimality**. Cambridge, MA: MIT Press.
- RIZZI, L. (1990b) Speculations on Verb-Second. In Nespov, M. and al. (eds) **Grammar in progress: a festschrift for H. Van Riemsdijk**. Dordrecht: Foris.
- ROBERTS, I. (1993) **Verbs and Diachronic Syntax**. Dordrecht: Reidel.
- ROUVERET, A. (1987) **Syntaxe des dépendances lexicales. Identité et identification dans la théorie syntaxique**. Tese de doutorado, Paris VII.
- ROUVERET, A. (1992) Clitic placement, focus and the Wackernagel Position in European Portuguese. Comunicação apresenta no **ESF Workshop on Clitics**, Donostia.
- ROUVERET, A (1994-95) **Seminário**, Université Paris 8.
- SALVI, G. (1990) La sopravvivenza della legge di Wackernagel nei dialettioccidentali della penisola iberica. **Medioevo Romano** 15,177-210.
- TORRES MORAES, M. A. (1995) **Do português clássico ao português moderno: um estudo da cliticização e do movimento do verbo**. Tese de doutorado, UNICAMP.
- URIAGEREKA, J. (1992) A Focus Position in Western Romance. comunicação apresentada no 15º **colóquio do GLOW**, Lisboa.
- URIAGEREKA, J. (1995) Syntax of clitic placement in Western Romance, *Linguistic Inquiry* 26, 79-123.