

Ronald Beline Mendes

A Gramaticalização de estar + gerúndio no Português Falado

Dissertação apresentada ao Curso de Lingüística do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas como requisito parcial para obtenção do título de Mestre em Lingüística.

Orientador: Profa. Dra. Maria Luiza Braga

UNICAMP

Instituto de Estudos da Linguagem

1999



9915112

UNIDADE: BC
Nº GRAMALA: 11111111
EX: 38311
229/99
11/08/99
Nº 000

CM-00125692-9

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA IEL - UNICAMP

M522g Mendes, Ronald Beline
A gramaticalização de estar + gerúndio no português falado./
Ronald Beline Mendes. - - Campinas, SP: [s.n.], 1999.

Orientador: Maria Luiza Braga
Dissertação (mestrado) - Universidade Estadual de Campinas.
Instituto de Estudos da Linguagem.

1. Gramática comparada e geral – gramaticalização. 2. Mudanças linguísticas. 3. Sociolinguística. 4. Língua portuguesa - verbo I. Braga, Maria Luiza.. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Profa.Dra. Maria Luiza Braga - Orientador

Profa. Dra. Margarida Salomão

Prof. Dr. Sírio Possenti

Prof. Dra. Charlotte Galves - suplente

Este exemplar é a redação final da tese
defendida por Ronald Beline Mendes

e aprovada pela Comissão Julgadora em
24 / 02 / 99.

De. Maria Luiza Braga

Ao professor e amigo

Ataliba de Castilho

AGRADECIMENTOS

Por razões óbvias, quero agradecer em primeiro lugar à CAPES, instituição que financiou meus estudos no Mestrado.

Logo em seguida, agradeço à minha orientadora, a Prof. Dra. Maria Luiza Braga - a já tão nossa Malu. Vocês não podem imaginar o que é dar uma tabela de dados quantificados nas mãos dela (ela sempre acha um defeitinho...). Orienta, sem conduzir. Sua palavra de ordem é o “talvez” (foram 37, desde a primeira versão do texto que você vai ler daqui a pouco: *talvez você pudesse...*, *talvez fosse melhor...*, e por aí vai...). Tem uma habilidade que considero importante: aliar aguda competência e vontade para o trabalho a um notável bom gosto no vestir-se e no modo de viver a vida e de tratar as pessoas. E o melhor de tudo: ela é assim mesmo, não parece ter aprendido com ninguém. Sorte de quem a tem como orientadora.

Quero agradecer também aos meus pais, que já não me perguntam a célebre “*E quando é que vai acabar, os seus estudos?*”. Agradeço-lhes imensamente por seu respeito à minha desmotivação em colocá-los a par de minha vida acadêmica, das coisas que estudo, das coisas que escrevo.

Às professoras Bernadete Abaurre, Charlotte Galves e Mary Kato, do IEL, por suas indicações bibliográficas e disponibilidade para discussões sobre assuntos relacionados à minha dissertação; à Professora Eleonora Albano, também do IEL, por seus esclarecimentos sobre o Laboratório de Fonética e sobre a Fonética Acústica; à professora Cida Torres, da USP, por suas indicações bibliográficas acerca da Teoria Gerativa; e ao professor Håj Ross, da Universidade do Texas - EUA, pelas duas leituras comentadas que chegou a fazer do texto a ser apresentado na qualificação dessa dissertação, o meu cordial obrigado.

Por fim, muito obrigado aos amigos Solange Rodrigues, Fernando Maués, Adriana Kanzepolsky e Cecília Olmos que, vez por outra, diziam a coisa certa na hora certa, sobre o trabalho, a vida, e essas coisas. E principalmente ao Guillermo, que se dignou a ler aquelas partes da dissertação que eu menos gostava, oferecendo-me a possibilidade do diálogo.

SUMÁRIO

	página
LISTA DE FIGURAS, GRÁFICOS E TABELAS.....	06
RESUMO.....	07
 Introdução.....	 08
 Parte I. Considerações Teóricas - Gramaticalização: do Léxico para a Gramática...	 10
1. Os estágios de gramaticalização.....	20
1.1. De verbo pleno para verbo auxiliar (auxiliarização).....	20
1.1.1. O modelo da “perda” (descoramento, <i>bleaching</i>).....	23
1.1.2. O modelo da implicatura.....	27
1.2. Morfologização.....	33
2. Mais, ou menos gramaticalizado?.....	41
2.1 Os “parâmetros” de Lehmann.....	42
2.2. Os “princípios” de Hopper.....	48
 Parte II. Análise dos Dados.....	 52
1. Análise Qualitativa.....	53
2. Análise Quantitativa.....	73
2.1. Os grupos de fatores.....	75
2.2. Os resultados obtidos.....	86
 Conclusão.....	 104
 ABSTRACT.....	 106
 Referências Bibliográficas.....	 107

LISTA DE FIGURAS, QUADROS E TABELAS

FIGURAS

- Figura 1:** Quadro sinóptico da gramaticalização (Lehmann 1982)
Figura 2: Do conceito-fonte para o conceito-alvo
Figura 3: Organização hierárquica dos constituintes prosódicos
Figura 4: Exemplo de representação sintática de *estar* enquanto verbo pleno
Figura 5: *Estar* em construções locativas (I)
Figura 6: *Estar* em construções locativas (II)
Figura 7: *Estar* em construções atributivas
Figura 8: *Estar* nas construções com gerúndio
Figura 9: *Estar* nas construções com locativo e gerúndio
Figura 10: *Estar* + *-ndo* através dos séculos
Figura 11: *Estar* nas construções locativas e atributivas através dos séculos
Figura 12: Cadeia de extensão metafórica do verbo *estar*
Figura 13: *Continuum* das construções de *estar*

GRÁFICOS

- Gráfico 1:** A vogal “a” em *tá*
Gráfico 2: A vogal “a” em *está*
Gráfico 3: Variação entre *estar/tá* conforme a faixa etária do falante
Gráfico 4: Uso das variantes *estar/tá* conforme o grau de atenção à fala
Gráfico 5: Uso de *tá* nas diferentes faixas etárias, de acordo com o grau de atenção à fala
Gráfico 6: Uso de *estar* nas diferentes faixas etárias, de acordo com o grau de atenção à fala.
Gráfico 7: A inserção de elementos entre o auxiliar e o gerúndio e a escolha de *estar* ou *tá*

TABELAS

- Tabela 1:** Resultados dos Grupos de fatores sociais
Tabela 2: Os falantes e a escolha de *tá* ou de *estar*
Tabela 3: O tempo do auxiliar
Tabela 4: A variação *tá/estar* de acordo com a palavra fonológica à esquerda do auxiliar
Tabela 5: Complexidade de X e a variação *tá/estar*
Tabela 6: O número de sílabas do verbo no gerúndio e a variação *tá/estar*

RESUMO

O propósito deste trabalho é verificar como se dá o processo de gramaticalização do verbo *estar* no português falado no Brasil. A questão que motivou o estudo é a seguinte: estaria a forma reduzida *tá* substituindo a forma plena *estar* nas construções perifrásticas com o gerúndio? Para encaminhar uma resposta a tal pergunta, analisei 5 das entrevistas do tipo “Diálogo entre dois informantes” do corpus compartilhado do Projeto NURC/SP. Baseado sobretudo no modelo proposto por Heine & Claudi & Hünemeyer 1991 para o estudo da gramaticalização, e tentando ver tal processo como um tipo de mudança linguística (nos termos de Labov 1972), apliquei o pacote de programas estatísticos VARBRUL aos dados qualitativamente analisados. A partir da análise quantitativa, foi possível constatar que os falantes mais velhos tendem a empregar a forma plena *estar*, ao passo que os falantes mais jovens tendem a não usar a forma plena, mas somente a reduzida - o que caracteriza uma mudança em progresso. Quanto à morfologização de *tá*, ainda que elementos interferentes (tais como locativos e adjuntos adverbiais) tendem a não ocorrer entre o auxiliar e o gerúndio quando se usa a forma reduzida, ainda não se pode afirmar que *tá* já funciona como um clítico e que vai ser aglutinado ao gerúndio.

Introdução

A idéia de desenvolver, no nível do Mestrado, um estudo sobre a gramaticalização do verbo *estar* nas perífrases de gerúndio teve sua gênese em meu trabalho de Iniciação Científica. Foi durante o desenvolvimento daquele trabalho, quando estava preocupado em descrever os usos da perífrase *estar* + *-ndo* a partir do conceito de Aspecto Verbal, que comecei a prestar atenção no fato de que a pronúncia do verbo auxiliar muitas vezes não correspondia à sua forma escrita canônica, como acontece na frase:

“mas o que que cê tá fazendo”?

Na ocasião daquele trabalho, sem me deter muito no problema, eu me perguntava sobre o que levaria o falante a usar *tá* em vez de *está*. Como estudante de linguística, não poderia incorrer na explicação simplória de que se tratava simplesmente de um “erro”, desencadeado por um “descuido” no uso da língua. Devia haver questões semânticas e sintáticas que explicassem o uso de *tá* - uma redução de *estar* enquanto palavra fonológica.

Ainda na Iniciação Científica, fazendo a pesquisa bibliográfica sobre *estar* + *gerúndio*, acabei me deparando com desenvolvimentos teóricos sobre um processo freqüente nas línguas naturais - a gramaticalização. Teoricamente, e da forma mais simplificada possível, trata-se de um processo que se dá quando um item lexical passa a desempenhar um papel gramatical. Com o avanço e o fortalecimento de tal processo, o item gramaticalizado poderia sofrer diminuição de seu tamanho, de acordo com as previsões da teoria.

Pode parecer demasiado rápido e superficial definir *gramaticalização* da forma acima, mas foi assim mesmo que comecei a pensar na redução *estar* > *tá* como um caso de gramaticalização em progresso. Foi a partir daí, então, que resolvi desenvolver este trabalho no Mestrado, dedicando-me a:

- (i) verificar como se deu a gramaticalização do verbo *estar*; e
- (ii) verificar se, na fala, a forma reduzida *tá* estaria substituindo a forma plena *estar*, nas perífrases de gerúndio, como um avanço do mesmo processo de gramaticalização.

Para tanto, procedi a uma análise qualitativa dos dados extraídos de cinco entrevistas do Projeto NURC/SP, mais especificamente cinco dos seis “Diálogos entre dois Informantes” (D2). Em seguida, fiz uma análise quantitativa de tais dados, submetendo-os ao pacote estatístico VARBRUL.

As análises qualitativa e quantitativa vêm apresentadas na Parte II desse trabalho, onde também discuto sobre a relação que haveria entre a teoria da gramaticalização e uma teoria mais geral sobre a mudança linguística - no caso, a teoria laboviana. Antes, porém, na Parte I, apresento as considerações teóricas que julguei mais relevantes para o caso da gramaticalização de um verbo.

Como já disse Castilho 1995, a respeito da gramaticalização, *“a bibliografia consultada revela de parte dos autores certa despreocupação em situar-se claramente dentro de uma 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’”*.

Tomo como minhas suas palavras, para afirmar que os estudos de gramaticalização, da forma como vêm referenciados aqui, enquadram-se numa concepção funcional de língua. Não vou me preocupar, neste trabalho, em discutir o que seja *funcionalismo*, essa que é uma *“tarefa tão difícil”*, nos dizeres de Moura Neves 1997, uma vez que *“os rótulos que se conferem aos estudos ditos funcionalistas mais representativos geralmente se ligam diretamente aos nomes dos estudiosos que os desenvolveram, não a características definidoras da corrente teórica em que eles se colocam”*.

Vou me limitar a dizer, então, a exemplo de Gebruers 1987 (apud Moura Neves 1997), que a concepção de linguagem que tomo aqui é a defendida pela gramática funcional *“porque não separa o sistema linguístico e suas peças das funções que têm de preencher”* e *“porque reconhece, na instabilidade da relação entre estrutura e função, a força dinâmica que está por trás do constante desenvolvimento da linguagem”*.

PARTE I - CONSIDERAÇÕES TEÓRICAS

Gramaticalização: do Léxico para a Gramática

Como foi dito na introdução acima, constata-se que, nos muitos trabalhos já desenvolvidos sobre gramaticalização, há sempre uma grande preocupação em definir o termo, sem contudo contextualizar tal definição dentro de um entendimento do que seja a gramática de uma língua natural.

Neste trabalho, não faço um resumo da história da utilização do termo, nem dos principais trabalhos sobre gramaticalização e seus autores. Bons históricos abrem os trabalhos de Lehmann 1982, Hopper & Traugott 1993 e Heine & Claudi & Hünemeyer 1991; um histórico mais sucinto, e também mais crítico, pode ser encontrado em Castilho 1995. Interessa dizer, aqui, que as definições de gramaticalização em geral se baseiam no fato de que as formas observadas estariam “caminhando” ao longo de uma trajetória, em cujo início estaria um item lexical e em cujo “término” estaria um item gramatical.

Ou seja, entende-se por gramaticalização, basicamente, a transição por que passa um item ao deixar de ser parte do Léxico para fazer parte do sistema gramatical. Neste sentido, a gramaticalização constitui um *processo*, sendo essa uma concepção que, de acordo com Heine & Claudi & Hünemeyer 1991, é comum à grande maioria dos trabalhos sobre o assunto.

Como decorrência dessa concepção de *processo*, aproximadamente na mesma maioria dos trabalhos considera-se que a gramaticalização:

- (i) é essencialmente um processo *diacrônico*;
- (ii) envolve sobretudo noções de morfologia; e
- (iii) é unidirecional por natureza.

Com respeito à primeira das três considerações acima, pode-se citar Traugott & König (apud Heine & Claudi & Hünemeyer 1991), para quem:

“grammaticalization (...) refers primarily¹ to the (...) historical process whereby lexical

¹ “Primarily” deve ser lido com o sentido de *sobretudo*. Numa leitura rápida da citação acima seria possível o entendimento de que Traugott & König teriam sido os primeiros a considerarem que a gramaticalização diz

items in the course of time acquire a new status as grammatical, morphosyntactic forms, and in the process come to code relations that either were not coded before or were coded differently.”

Dessa forma, a gramaticalização pode ser entendida como um caso especial de mudança lingüística, no qual as formas encontram-se alinhadas num “continuum” (Meillet 1912, apud Heine-Claudi-Hünemeyer 1991), num “cline” (Hopper & Traugott 1993), apontando, direcionando, por assim dizer, para a Gramática. Deve ficar claro, entretanto, que nem toda mudança lingüística constitui um caso de gramaticalização. Vejam-se, por exemplo, casos de mudança fonológica como /ei~/e/ em peixe - “peixe”, feira - “fera”, etc (Araújo 1998), que nada têm a ver com gramaticalização.

Por outro lado, a literatura sobre o assunto mostra que, a partir da década de 70, alguns autores passaram a tentar definir e interpretar a gramaticalização como um processo sincrônico. Heine & Claudie & Hünemeyer 1991 apontam, contudo, que em numerosos trabalhos não fica claro como o processo é concebido sincronicamente. Isto é, afirma-se que se está fazendo uma análise sincrônica, quando na verdade se está trabalhando com diacronia, a partir de dados selecionados dentro de um recorte temporal estreito, geralmente contemporâneo ao estudioso.

Aparentemente, Hopper 1987 (apud Heine & Claudie & Hünemeyer 1991), com sua proposta da “Gramática Emergente” (“*grammar is always emergent, but never specific*”), é o mais radical no uso do termo *gramaticalização* como um processo sincrônico. Em sua terminologia, ele define *gramaticização* como “*movement toward structure*”. Esse autor argumenta contra a noção que ele denomina como “*a priori grammar*”, dizendo:

“The notion of Emergent Grammar is meant to suggest that structure, or regularity, comes out of discourse and is shaped by discourse as much as it shapes discourse in an on-going process. Grammar is hence not to be understood as a prerequisite for discourse (...) Its forms are not fixed templates, but are negotiable in face-to-face interaction in ways that

respeito a um processo histórico, o que não é verdade. De fato, nenhum autor nega a possibilidade de investigar a gramaticalização diacronicamente.

reflect the individual speakers' past experience of these forms, and their assessment of the present context".

Para Hopper 1987, portanto, a rigor não existiria *gramática*, nem *gramaticalização* (nos termos definidos acima), mas sim *gramaticização*.

No entender de Heine & Claudi & Hünemeyer 1991, de todo modo, os trabalhos que começaram a ser desenvolvidos após os 70, por terem em conta que os limites entre Sincronia e Diacronia não são rigidamente incontestáveis, começaram a reconhecer na gramaticalização um eficiente parâmetro explanatório para o entendimento de uma gramática sincrônica. A respeito daqueles trabalhos, tais autores afirmam:

"Dissatisfaction with existing models of grammatical description provided a major incentive for turning to grammaticalization as a means of surmounting "static" approaches for analyzing grammar, in particular structuralism and generative transformational grammar".

Numa linha de pensamento similar, Castilho 1995:31 vê, nas postulações presentes nos trabalhos desenvolvidos após os 70, um impulso para uma Teoria do Discurso. Isto porque, sobretudo no modelo da gramática emergente proposto por Hopper, as regularidades (Gramática) seriam *"provisórias e continuamente sujeitas à negociação, à renovação e ao abandono"*, estando na enunciação o desencadeador do dito *processo de gramaticalização*. Ora, a tal rasgo de funcionalismo se opõe justamente uma concepção formal de língua, cujos fenômenos se explicam não pela enunciação, mas no enunciado.

Essa discussão teórica é importante, aqui, porque me proponho a analisar o verbo *estar* na perífrase de gerúndio dentro de um recorte temporal bastante estreito (a análise quantitativa foi feita com dados extraídos do *corpus* compartilhado do Projeto NURC/SP - Preti & Castilho (orgs.) 1986 - gravados em 1973). Dado o fato de que, neste mesmo recorte temporal, coexistem formas com diferentes graus de gramaticalização (algo que vem detalhado na Parte II), pode-se admitir que os limites entre Sincronia e Diacronia não são rigorosos.

Seguindo a lista acima, temos em (ii) o fato de que a maioria dos estudiosos concebe a gramaticalização como uma noção morfológica ou morfossintática, o que parece algo

razoável, uma vez que seus trabalhos tratam quase sempre do desenvolvimento de uma determinada palavra ou morfema a partir de outra.

Para exemplificar, pode-se citar Benveniste 1968 (apud Heine & Claudi & Hünemeyer 1991), que ao propor uma distinção entre “mutação inovadora” (transformação de “palavras principais” - *mots principaux* - em palavras acessórias - *mots accessoires*, seguindo a terminologia de Meillet 1912) e “mutação conservadora”, enfatiza o processo morfossintático envolvido na segunda. O autor argumenta que a “mutação conservadora” consiste na “substituição de uma categoria morfêmica por uma categoria perifrástica com a mesma função” - daí, “conservadora” - como no caso da substituição de flexões de caso por sintagmas preposicionais, dentre outros exemplos que o autor descreve.

Isto não quer dizer, contudo, que o processo de gramaticalização não implique correlatos com outros domínios, como o da fonologia. Conforme será visto mais adiante, o avanço no processo de gramaticalização de um verbo pressupõe sua redução fonológica, o que parece ser o caso do verbo *estar*, tema deste trabalho.

Por fim, quanto à unidirecionalidade, o aspecto do item (iii) acima, trata-se de característica inerente aos processos de gramaticalização. As definições mais comuns, explícita ou implicitamente, em geral falam em um processo evolutivo que se dá numa única direção: de “menos gramatical” para “mais gramatical”.

Kurylowicz 1965 (apud Lehmann 1982), por exemplo, propõe a seguinte definição: “*Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivative formant to an inflectional one*”.

Casos de desgramaticalização ou de regramaticalização, que poriam em dúvida a unidirecionalidade como característica intrínseca ao processo, são raros e têm, de acordo com Lehmann 1982, uma análise inadequada em sua base. Para Castilho 1996, “*a dificuldade na admissão da desgramaticalização está em que tais formas deveriam passar por um processo de ressemantização e de reforço fonológico. Exemplos claros desse processos ainda não foram encontrados.*”

Entretanto, fala-se de desgramaticalização, por exemplo, a partir do caso de *tá* enquanto marcador conversacional, como no caso abaixo.

... não faz isso não, tá?

Para o autor acima, tal fato não pode ser entendido como desgramaticalização. De acordo com ele, os linguistas que em geral assumem que ocorre desgramaticalização em casos como esse argumentam que tais itens estão desprovidos de propriedades gramaticais, como tempo e pessoa, apresentando-se sempre numa forma fixa. Seguindo Castilho 1996, essa é uma análise equivocada, pois considera as categorias do discurso como “*itens privados de propriedades gramaticais*”.

Em razão disso, o mesmo autor propõe interpretá-los como exemplos de “discursivização”, lançando a pergunta-argumento: “*Por que supor que propriedades discursivas excluem propriedades gramaticais?*” Ou seja, Castilho 1996 sugere que, em se observando um item gramatical usado pelo falante para estabelecer coesão textual e para manter a interação comunicativa, fale-se em interação da Gramática com o Discurso. Um outro exemplo disso seria o uso discursivo de itens lexicais.

Além da questão da unidirecionalidade, a definição já clássica de Kurylowicz parece merecer ainda outra aclaração. Veja-se o verbo *estar* nas construções com gerúndio abaixo:

(a) *você não está fazendo direito...*

(b) *você não tá fazendo direito...*

(c) *ele está em casa todo dia a partir das 2:00*

Considerando a forma *tá* do exemplo (b) como o resultado da redução fonológica de *está* do exemplo (a), e prevendo a perda de seu acento primário e sua conseqüente cliticização ao gerúndio, pode-se dizer que *tá*, em (b) é mais gramatical que *está*, em (a)? Ou, em outras palavras, uma construção perifrástica seria menos gramatical que um morfema flexional afixado a um semantema?

Sim. Pode-se entender dessa forma se se considera que *tá* - em (b) - desempenha um papel essencialmente gramatical (marcação de número e pessoa), enquanto que *está* - em (a) -, além de desempenhar esse papel, supostamente é mais semanticamente semelhante ao *estar* das construções locativas - em (c) -, tendo portanto “mais cara de verbo”, dito de uma maneira simples.

Entretanto, se voltamos ao título da Parte I desse trabalho - *Gramaticalização: do Léxico para a Gramática* - a sequência “*de menos para mais gramatical*” na definição já clássica de Kurylowicz pode soar estranho. Numa leitura mais descuidada, pareceria que Kurylowicz opõe léxico a gramática, ou seja, o leitor poderia entender que o léxico de uma língua não integra a gramática da mesma.

Neste caso, um clítico como *tá* seria mais parte da gramática do português que o próprio verbo *estar* conjugado em (a) e em (c). Ou então, falando de um modo mais genérico, um morfema verbal faria mais parte da gramática de uma língua que o verbo que lhe dá origem, o que é inaceitável.

Dessa forma, o título “*do Léxico para a Gramática*” - que é um dos modos mais rotineiros e informais de se referir ao processo de gramaticalização - não pode ser tomado *ipsis litteris*, justamente porque o próprio léxico integra a gramática. Castilho 1995:58 deixa claro:

“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 forte distinção entre “itens lexicais”, como elementos que compõem o Léxico de uma língua, onde imperam as idiossincrasias, e “categorias gramaticais”, como morfemas ou como palavras funcionais (verbos auxiliares, nexos, etc.) que integram a Gramática da língua, onde imperam as regularidades.”

De acordo com o mesmo autor, “*é preciso ter em conta que mesmo fora do enunciado (...) as palavras já dispõem de propriedades gramaticais (fonológicas, morfológicas e sintáticas), semânticas (referencialidade, predicatividade, dêixis, foricidade) e discursivas (topicidade, coesividade, etc)*”, de modo que é nada razoável considerar os itens lexicais como uma entidade tábua-rasa, isto é, sem propriedades linguísticas de qualquer natureza.

Feitas tais aclarações, retomo a noção de processo, que segue abaixo ilustrada pelo quadro organizado por Lehman 1982a: 13:



Figura 1 - Quadro sinótico da Gramaticalização

Este quadro proposto por Lehmann ilustra bem a idéia de “cline”, de “continuum”, bem como o “movimento” de menos gramatical para mais gramatical (levando-se em conta a ressalva feita anteriormente), permitindo visualizar essa que é, com maior ou menor rigor, a acepção que constitui a base teórica da maioria dos trabalhos até aqui desenvolvidos. Tal quadro tem sido aproveitado em estudos de gramaticalização de categorias lexicais, como no caso dos verbos, e também de categorias funcionais. O próprio Lehmann o descreve da seguinte forma:

“We assume that grammaticalization starts from a free collocation of isolating lexemes in discourse. This is converted into a syntactic construction by syntacticization, whereby some of the lexemes assume grammatical functions so that the construction may be called

analytic. Morphologization, which here means the same as agglutination, reduces the analytic construction to a synthetic one, so that grammatical formatives become agglutinating affixes. In the next phase, the unit of the word is tightened, as the morphological technique changes from agglutinative to flexional. This transition from morphology to morphophonemics will here be called demorphemicization, (...) where expression and content of the grammatical category become zero."

A escala que constitui a primeira linha do quadro acima - *Discurso > Sintaxe > Morfologia > Morfofonêmica > Zero* - foi proposta inicialmente por Givón 1979, como deixa claro o próprio Lehmann 1982, sendo ela uma síntese de duas postulações: "*a morfologia de hoje é a sintaxe de ontem*" (Givón 1971:413, apud Castilho 1995:30) e "*a sintaxe de hoje é o discurso pragmático de ontem*" (Givón 1979:208-9). Não vou me aprofundar nessas que são consideradas as "máximas givonianas", mas valho-me delas para lembrar que ao conceito de gramaticalização enquanto um processo soma-se a consideração de que o mesmo seja irrecorrente, interminável, no sentido de que está sempre ativo. Assim, uma construção analítica pode se gramaticalizar numa construção sintética, que pode vir a ser preterida, no uso da língua, em relação a uma outra construção analítica, que, por sua vez, pode vir a gramaticalizar-se, e assim por diante.

Lehmann 1982 também deixa claro que o quadro acima é simplificado e incompleto, pois faz parecer que a extremidade final do processo é facilmente apreensível, o que nem sempre é verdade. Além disso, o esquema proposto negligencia o ponto inicial do processo, que não pode ser prontamente identificável. Para o propósito das considerações teóricas neste trabalho, contudo, o quadro acima é bastante ilustrativo.

O autor ressalta, ainda, que não é essencial à teoria sobre gramaticalização considerar que o elemento por ela afetado deva entrar no processo pelo início e deixá-lo no fim. Desse modo, não necessariamente haverá morfologização após ocorrer a sintaticização de um termo.

Ajustando o quadro acima para a gramaticalização de verbos, dado o tema deste trabalho, podemos ainda considerar que as fases do processo sejam as seguintes:

A transformação de um verbo pleno em verbo auxiliar corresponde à fase de sintaticização. Na passagem de verbo auxiliar para clítico e em seguida para afixo temos a correspondência com a fase de morfologização, no quadro de Lehmann. Não vou aventar, neste trabalho, nenhuma hipótese acerca da desmorfemização, de modo que vou me ater apenas às duas primeiras fases de Lehmann, ou seja, estarei considerando o processo “desde” **Verbo pleno** “até” **Afixo** na escala proposta acima.

O termo indicado por Lehmann 1982 para o primeiro estágio não parece muito feliz, uma vez que se pode entender “sintaticização” como a fase em que uma unidade que não era sintática passa, então, a sê-lo - o que não é nada defensável. É verdade que problemas acerca da terminologia empregada no trabalho científico nem sempre são fáceis de resolver. No caso deste trabalho, contudo, dentro dos limites da gramaticalização de verbos, vou optar por seguir Heine 1993, usando o termo *auxiliarização*. Este me parece um termo menos problemático que *sintaticização*, ainda que a linha divisória entre um verbo pleno e o verbo auxiliar dele derivado não tenha colocação clara e indiscutivelmente definida.

Pensando em termos de verbos, entenda-se *sintaticização*, portanto, como a fase em que um verbo auxiliar evolve a partir de um verbo pleno. É o que teria ocorrido na formação dos Tempos compostos, na codificação do Modo, do Aspecto e da Voz. Em todos esses casos, um V Auxiliar e um V pleno se reuniram numa perífrase, constituindo o núcleo predicativo da sentença.

Ramat 1987 (apud Heine 1993) chama à atenção para que as fases ou estágios de gramaticalização não sejam tomados como entidades discretas, com limites claramente estabelecidos, definindo-os antes como:

“focal instances of the process between the two poles - a process which is continuously going on without breaks...”.

Este autor propõe, então, pelo menos um estágio intermediário na evolução Verbo pleno > Verbo Auxiliar, como segue no excerto abaixo:

I. *Full verbs*. At this stage, verbs have their full semantic meaning”, the object of the finite verb may be different from that of the nonfinite verb, and the clause complement may consist of a nominal instead of a nonfinite verb; that is, the nonfinite verb is not an obligatory constituent to the clause.

II. *Predicative construction*. The nonfinite verb is a compulsory constituent of the clause.

III. *Periphrastic forms*. The finite verb/auxiliary has no autonomous semantic content, it is now a marker for tense aspect, and mood.”

Castilho 1995 define os mesmos da seguinte forma : “*Verbos plenos são os que funcionam como núcleos do predicado. Os verbos funcionais transferem esse papel para os constituintes à sua direita. Os verbos auxiliares acompanham verbos nucleares na forma nominal, aos quais atribuem as categorias de número e pessoa, tempo e modo*”.

Finalmente, para falar em morfologização, podemos tomar o caso já clássico da formação do futuro românico. Após ter perdido seu sentido original, ter sofrido redução fonológica e ter sido reanalisado, um dos termos da perífrase - o auxiliar - transforma-se em morfema, integrando o sub-sistema morfológico da língua, como mostra o exemplo abaixo para o português.

amare habemus > *amarabémus > *amarémus > amaremos

Isto é, de uma forma perifrástica, analítica, desenvolveu-se uma forma sintética. Temos, entretanto, no português, uma outra forma perifrástica de codificar o futuro (*ir + infinitivo*) junto com a qual a forma sintética pode persistir ou não. Qualquer que seja o caso, o que interessa dizer para a teoria sobre gramaticalização é que o processo pode reinstaurar-se, a partir da forma mais recente:

vou amar > vô amar > ?

A hipótese para o caso do verbo *estar* nas construções com gerúndio é a de que ele está em pleno processo de morfologização. A análise qualitativa e quantitativa dos dados estão apresentadas na Parte II. Ainda em termos teóricos, faço a seguir uma discussão sobre as abordagens dos dois estágios acima apresentados. Em 2, discorro sobre duas propostas de parâmetros para a “aferição” do grau de gramaticalização de uma forma.

1. Os estágios de gramaticalização

Neste item, apresento as questões teóricas relativas à abordagem dos dois estágios do quadro de Lehmann a que me referi acima. Em 1.1 a seguir, trato do processo de “passagem” de verbo pleno para verbo auxiliar, mostrando como o fenômeno vem sendo mais comumente abordado. Em 1.2, por sua vez, trato do estágio da morfologização, apresentando o modelo teórico que serviu de base para a análise quantitativa dos dados. Sempre que necessário, ainda que no interior de uma parte dedicada a considerações teóricas, lancei mão de exemplos de uso do verbo *estar* no português brasileiro (PB).

1.1. De verbo pleno para verbo auxiliar (auxiliarização)

Vimos acima que o termo *sintaticização* não parece ser o melhor rótulo para denominar o processo de desenvolvimento de um verbo auxiliar a partir de um verbo pleno. Por essa razão, daqui por diante passo a me referir a tal estágio através do termo *auxiliarização*.

São bastante presentes nos trabalhos sobre gramaticalização - inclusive no de Heine & Claudi & Hünemeyer 1991, que primam pelas questões semânticas - observações referentes às mudanças na sintaxe das formas em processo de gramaticalização. Neste

trabalho, é na Parte II, quando da análise qualitativa dos dados, que tento descrever sintaticamente os diferentes usos do verbo *estar*.

É uma questão de ordem semântica, contudo, que parece ser privilegiada, tanto em Heine & Claudi & Hünemeyer 1991:22 - de uma forma mais genérica - quanto em Heine 1993 - de maneira mais específica, para o caso da gramaticalização de verbos: - o que o significado do auxiliar tem em comum com o do verbo do qual foi derivado?

Não me ocupo, neste trabalho, em definir o estatuto categorial dos auxiliares, questionando-me se os mesmos constituem uma categoria à parte dos verbos plenos - o que pressuporia a análise de outros verbos além de *estar*. Emprego o termo aqui, portanto, da maneira como ele é mais tradicionalmente empregado: - verbo auxiliar é aquele que entra na composição de uma perífrase verbal juntamente com outro verbo em forma nominal (infinitivo, particípio ou, como no caso em questão, gerúndio), portando flexão de tempo, número e pessoa.

A questão da diferença semântica entre um verbo dito auxiliar e o verbo pleno que lhe deu origem, entretanto, é das mais importantes nos estudos sobre gramaticalização. Willet 1988 (apud Heine & Claudi & Hünemeyer 1991), revisando a literatura das duas décadas anteriores ao seu trabalho, e focalizando as questões de “generalização semântica” envolvidas nos processos de gramaticalização, reconhece e discute as seguintes hipóteses:

- (a) a hipótese da extensão metafórica, de acordo com a qual o significado concreto de uma expressão é aplicado a um contexto mais abstrato;
- (b) a hipótese da contenção, pela qual se admite que os significados gramaticais já estariam incluídos na estrutura semântica do item lexical fonte;
- (c) a hipótese da implicatura, que assume a convencionalização de implicaturas como o mecanismo predominante na criação de significados secundários, que gradualmente tomam o lugar de significados primários.

Heine 1993, por sua vez, fala em “modelos” quando se refere à abordagem das mudanças semânticas envolvidas na gramaticalização, depreendendo os três seguintes:

- (a) o modelo da perda: a semântica lexical do item em processo de gramaticalização é “esbranquecida”, ou seja, ocorre perda de conteúdos semânticos ou conceituais; propõe-se o seguinte esquema para sintetizar o modelo:

$$ab > b$$

(b) o modelo da perda e do ganho: a perda de significados lexicais pode ser seguida pelo ganho de significados pragmáticos; esquematicamente, tal modelo é assim representado:

$$ab > bc$$

(c) o modelo da implicatura: o mecanismo predominante no desenvolvimento de significados gramaticais é a convencionalização de implicaturas ou de inferências, podendo o conceito original desaparecer em favor de um novo conceito; neste caso, o esquema proposto é

$$ab > bc > cd$$

Ao discutir esses modelos, Heine 1993 afirma que eles não são contraditórios entre si, mas focalizam o processo de perspectivas diferentes. De acordo com ele, o “modelo da perda” (*bleaching model*, em sua terminologia) enfatiza que o processo de gramaticalização envolve perdas, sobretudo de ordem semântica, ignorando outros aspectos que, indubitavelmente, também constituem o processo. Por outro lado, o “modelo da perda e do ganho” (*loss-and-gain model*) admite que as perdas ocorridas no processo são compensadas por ganhos de, por exemplo, propriedades funcionais específicas, como tempo, aspecto e modalidade (no caso de gramaticalização de verbos). O “modelo da implicatura” (*implicature model*), por fim, dá um passo à frente dos anteriores ao atentar para o fato de que um conceito original (**ab**) pode desaparecer inteiramente em favor de um novo conceito (**cd**).

Para confirmar que os modelos acima não se contradizem, Heine 1993 coloca seus esquemas explicativos lado a lado, notando que a diferença essencial entre eles é o seu grau de inclusividade, estando um modelo contido no outro, da maneira como transcrevo abaixo:

$ab > b$	modelo da perda
$ab > bc$	modelo da perda e do ganho
$ab > bc > cd$	modelo da implicatura

Comparando as considerações de Willet 1988 e Heine 1993 acima, vemos que o primeiro nomeia uma de suas hipóteses como “extensão metafórica”, que apesar de ser considerada por Heine 1993 tanto no modelo da perda e do ganho quanto no modelo da implicatura, não fica explícita em sua denominação. Vou apropriar-me, então, do termo “modelo” de Heine, discutindo a questão da perda, no sub-item 1.1.1 seguinte, e o “modelo da implicatura” em 1.1.2, onde também trato da extensão metafórica. Apesar de não detalhar os outros modelos ou hipóteses, por serem menos relevantes para o trabalho, devo dizer que não refuto seus pressupostos.

1.1.1. O modelo da perda (descoramento, *bleaching*)

Como vimos acima, o verbo auxiliar, de acordo com esse modelo, apresenta perda de conteúdo semântico em relação ao verbo pleno de que foi derivado. Tal hipótese costuma ser denominada, além de *bleaching*, por *semantic weakening*, *semantic depletion*, *desemanticization* e *fleshing out*. Por estes termos, fala-se em apagamento de traços semânticos e também em redução a um núcleo semântico abstrato. Lehmann 1982, por exemplo, refere-se ao produto final do processo que aqui chamo de auxiliarização da seguinte forma:

“(it) signifies little more than a kind of grammatical relation or an abstract image schema”.

Tal definição semântica de verbo auxiliar ensaiada por Lehmann, apesar de incidir sobre o fato de que o mesmo estabelece relações gramaticais, suscita a pergunta: - o que o auxiliar “guarda” da fonte lexical? (considerando-se “*little more*”, acima). Desse modo, está se considerando implicitamente que, relacionado ao processo de perda de parte do conteúdo

semântico lexical, haveria também um processo de preservação, podendo o auxiliar “conter” um resíduo, ainda que mínimo, de sua fonte.

Por outro lado, fala-se também em generalização do conteúdo semântico, como Bybee & Pagliuca 1985:63 (apud Heine 1993:90), através da qual o verbo em questão seria esvaziado de suas especificidades semânticas, apresentando como resultado uma distribuição mais geral, uma vez que passa a poder ser usado num maior número de contextos. De qualquer forma, tal generalização pressupõe que traços semânticos específicos são perdidos.

Para exemplificar, pode-se citar o caso do verbo *can*, no inglês, cujo uso pode ser caracterizado da seguinte forma:

I. existem condições de habilitação mental no agente;

John can use a computer

II. existem condições de habilitação no agente;

John can arrive anytime

III. existem condições de habilitação

It can be a great opportunity.

Esse é um tema que desperta muita discussão dentro da teoria sobre gramaticalização. Isto porque muitos linguistas admitem que o “desbotamento semântico” (*bleaching*) é indiscutível no decorrer do processo de gramaticalização, inclusive no estágio inicial - o da auxiliarização, no caso dos verbos.

Alguns autores, entretanto, salientam a noção de preservação do conteúdo semântico. Givón 1973 (apud Heine & Claudi & Hünemeyer 1991), por exemplo, em seu estudo sobre o desenvolvimento de categorias gramaticais de tempo, aspecto e modo a partir de verbos, sugere:

“the meaning of these categories is largely predictable because it is part of the internal semantic structure of their lexical sources” (a sublinha é minha).

Outros, a exemplo de Moura Neves 1997, questionam o uso indiscriminado do termo “perda”, lembrando que, ao contrário, pode se falar em “ganho”, alargamento, ampliação funcional do item. Neste caso, além da preservação de traços semânticos da fonte, haveria ganho de “propriedades”, um outro processo que ocorreria paralelamente ao da perda.

Outros autores, ainda, alegam que a gramaticalização não pode ser reduzida a um fenômeno de perda de traços. Para Heine & Claudi & Hünemeyer 1991, por exemplo, nos trabalhos em que se tem tal ponto de vista, a gramaticalização é vista da perspectiva do conceito-fonte, que porta o chamado “significado pleno”, sendo o “output” do processo interpretado como uma forma “empobrecida”, esvaziada, desprovida das especificidades semânticas da sua fonte.

Tais autores, ao se oporem à noção de *bleaching* como modelo para tratar da gramaticalização justificam-se da seguinte forma (as sublinhas são minhas):

*“While the bleaching model captures one important aspect of grammaticalization, it would seem that it ignores certain other characteristics of this process. In the process of grammaticalization, the source meaning may disappear completely, or it may be replaced by what - from a synchronic point of view - appear to be totally unrelated meanings or functions, such as the French negative markers *pas*, *personne*, *point*, and *rien* (which derive from nouns denoting, respectively, ‘step’, ‘person’, ‘point’, and ‘thing’), which in certain colloquial uses, where ‘ne’ is omitted, form the only expression of negation. Given the right context, grammaticalization may take directions that are difficult to reconcile with this model”*

Além disso, tais autores, ao enfatizarem que os processos de gramaticalização ocorrem como o resultado de uma interação entre aspectos cognitivos e pragmáticos, afirmam que a noção de *bleaching* acaba sendo inadequada como parâmetro descritivo e explanatório, uma vez que não se explica de fato como tais processos são desencadeados.

Heine 1993:95, por outro lado, explica o fato de que o modelo da perda é o mais popular nos trabalhos sobre gramaticalização. Em primeiro lugar, a aceitação da noção de perda como meio para descrever o processo está ligada à noção de isomorfismo na codificação lingüística. Para este autor, de acordo com o *princípio da quantidade*, em Givón 1985, o desenvolvimento de marcadores gramaticais de tempo e aspecto, a partir de verbos enquanto itens lexicais de significados concretos, está diretamente relacionado à quantidade de material morfológico e fonético necessária para a expressão de conteúdos gramaticais.

Explicando melhor, se se admite que a relação entre forma e significado seja de fato isomórfica², tem-se que a redução de substância morfológica e fonética de um item é consequência da redução da substância semântica a ser expressa. É esperado, portanto, que um verbo em processo avançado de gramaticalização sofra redução fonológica, como no caso de *estar* para *tá*. Pensando dessa forma, fica difícil não falar em perda.

Em segundo lugar, a noção de perda se encaixa muito bem na visualização do processo de gramaticalização como um todo. Pensando em termos de verbos, temos no extremo inicial do processo um verbo semanticamente “pleno”, enquanto no extremo final temos o zero, conforme quadro visto anteriormente. Parece então óbvio concluir que a evolução como um todo envolve uma redução constante do significado, de maneira que o conteúdo semântico de categorias gramaticais como tempo, aspecto e modalidade seria de fato “mais pobre” que o da fonte lexical que lhes deu origem.

Deve ter ficado claro, contudo, que “perda” é um modelo demasiado limitado para descrever tudo o que está envolvido na gramaticalização. Como já foi dito, o modelo da perda pode até explicar alguns casos de gramaticalização, visto que enfoca uma instância específica do processo, mas não constitui uma propriedade geral da gramática.

² Pelo *princípio da quantidade* de Givón, considera-se *isomorfismo* como sinônimo de *iconicidade*: para expressar uma menor quantidade de substância semântica, é necessária uma menor quantidade de substância fonética. Outros autores, entretanto, diferenciam tais características. Haiman 1983, por exemplo, questiona o isomorfismo em lingüística entendendo-o como a existência de uma única forma para a expressão de um dado significado, numa relação biunívoca. Para este autor, o isomorfismo seria uma face da iconicidade, e não seu sinônimo. A outra face seria a *motivação icônica*, de acordo com a qual as relações entre os componentes de um diagrama simbólico buscariam reproduzir as relações entre os componentes do referente a ser simbolizado.

1.1.2. O modelo da implicatura

Conforme já vimos mais acima, o modelo da implicatura pressupõe como mecanismo predominante no desenvolvimento de significados gramaticais a convencionalização de inferências. Isto faz desse modelo o mais abrangente dentre os reconhecidos, dado que os significados secundários que vão sendo estabelecidos via convencionalização não têm que estar, necessariamente, contidos na fonte lexical. Ou ainda, se estiverem, o conceito original pode desaparecer, em favor de outro.

Ao contrário do modelo descrito acima, neste não se dá importância à questão da perda de traços, mas sim à transferência de conceitos: entende-se e expressa-se uma coisa em termos de outra (Lakoff and Johnson 1980:5, apud Heine 1993), ou seja, opera-se uma transferência de conceitos de natureza metafórica. De acordo com Heine & Claudi & Hünemeyer 1991,

“(...) concrete concepts are employed in order to understand, explain, or describe less concrete phenomena. In this way, clearly delineated and/or clearly structured entities are recruited to conceptualize less clearly delineated or structured entities, and nonphysical experience is understood in terms of physical experience, time in terms of space, cause in terms of time, or abstract relations in terms of physical processes or spatial relations.”

Transferências conceituais, tais como as que foram rapidamente citadas acima, envolvem, de acordo com os autores, a abstratização do termo que constitui a fonte lexical. Em seu trabalho, eles identificam três tipos:

- abstratização generalizadora (*generalizing abstraction*): que consiste na redução do número de traços distintivos de um conceito para as suas características mais centrais ou nucleares;
- abstratização isoladora (*isolating abstraction*): através da qual se ressalta uma propriedade ou um traço particular, que não é necessariamente a “característica nuclear”;
- abstratização metafórica (*metaphoric abstraction*): que caracteriza o modo como entendemos e conceituamos o mundo ao redor de nós: relacionando conceitos mais abstratos a conceitos mais concretos; mais complexa que as anteriores, é a que interessa aqui, sendo justificada pelos autores da seguinte forma:

“Objects that are close to us are clearly structured and clearly delineated; they are ‘less’ abstract than objects that are more distant, less clearly structured and/or delineated. It also refers to referentiality or manipulability in discourse. Objects that refer, that are autonomous speech participants, are less ‘abstract’ than those that show a low-degree referentiality or manipulability”.

Com respeito aos dois primeiros tipos de abstratização, os autores também afirmam que eles podem ser levados em conta quando a gramaticalização é vista em termos de perda (*semantic bleaching*), conforme foi visto no item anterior. Em ambos os casos, o “output” do processo é parte de seu “input”, ou seja, no curso da gramaticalização ocorreria meramente uma redução de seu conteúdo intensional paralelamente a um aumento de sua extensão.

O que a abstratização metafórica traz a mais em relação ao desbotamento semântico, enquanto modelo para o estudo da gramaticalização, seria uma capacidade explanatória maior, na medida em que ela nos permite entender o que está por trás dos estágios iniciais do processo, sobretudo de uma perspectiva em que se vê a criatividade linguística correlacionada a questões cognitivas e pragmáticas.

Para exemplificar, utilizo aqui casos de *be going to* do inglês³, extraídos de Heine 1993:

(a) Suzanne *is going to* town.

(b) Suzanne *is going to* wake up in a minute

Segundo o autor, são boas as razões para postular que a transferência do uso de *be going to* de concreto/lexical (a) para abstrato/gramatical (b) seja de base metafórica:

(i) em primeiro lugar, trata-se da expressão de um “tipo de coisa” (tempo futuro - b) em termos de outro (movimentação espacial - a) ;

³ Exemplificando brevemente com o caso de *estar + gerúndio*, teríamos um auxiliar que, enquanto marcador de tempo, número e pessoa, constitui um uso mais abstrato do que aquele que pode ser identificado em construções locativas, de sentido mais concreto. Maiores detalhes são encontrados na análise qualitativa dos dados (Parte II).

(ii) em segundo, envolve a transferência de um “mundo real” - o mundo das entidades referenciais e das atividades cinéticas - para o mundo do discurso - entidades que só existem dentro do ato de fala⁴;

(iii) e finalmente, de acordo com um critério comumente empregado na definição de metáfora, pode-se considerar que o sentido de *be going to* em (a) é literal, ao contrário do sentido que assume em (b).

Fala-se, desse modo, em “conceito de origem” (*source*) e em “conceito de chegada” (*target*). Segundo Heine & Claudi & Hünemeyer 1991 e Heine 1993, tais conceitos, para o caso do exemplo acima, se encadeiam do modo simplificadaamente representado abaixo:



Figura 2 - Do conceito-fonte para o conceito-alvo

Representar a transferência conceitual dessa forma significa deixar claro que a gramaticalização não deve ser entendida como uma transição que se faz com entidades discretas, “*mas como uma extensão gradual do uso de uma entidade original*” (Moura Neves 1997:135).

Parece haver, entretanto, uma regularidade previsível na transferência de conceitos. De acordo com Heine & Claudi & Hünemeyer 1991, o processo metafórico envolvido na gramaticalização é governado por uma escala de abstratização, em que categorias conceituais se arranjam linearmente:

⁴ “*Transferência do mundo real para o mundo do discurso*” é tradução literal de Heine 1993:96. Este segundo ponto talvez fique mais claro falando-se em *transferência de um domínio de entidades concretas para um domínio de entidades abstratas*.

PESSOA > OBJETO > ESPAÇO > TEMPO > PROCESSO > QUALIDADE

Nesta escala, qualquer elemento pode conceitualizar um outro, que faça parte da categoria à sua direita. O exemplo acima - *be going to* - constitui um caso de abstratização ESPAÇO > TEMPO, no uso de verbos, o que se enquadra no processo de auxiliarização. Os autores que propuseram tal esquema, contudo, exemplificam-no também com palavras de natureza nominal, preocupados que estão em tratar o fenômeno da maneira mais genérica possível. Eles lançam mão, entre tantos outros, de uma caso na língua africana Ewe - “*megbé*”. Significando “costas” originalmente (OBJETO, parte do corpo), passa a ser empregada para os significados “atrás” (ESPAÇO), “depois” (TEMPO) e “retardado” (QUALIDADE).

Teoricamente, um significado novo pode substituir o que lhe deu origem, mas não necessariamente. Desse modo, numa análise sincrônica de dados pode-se deparar com um lexema que se refere a mais de uma das categorias acima, num mesmo contexto. É o que acontece, por exemplo, em

é tsí megbé

frase da língua Ewe (Heine & Claudi & Hünemeyer 1991:66), que pode ser entendida tanto como “ele está atrás” (categoria conceitual de ESPAÇO), quanto “ele está atrasado” (categoria conceitual de TEMPO), ou ainda “ele é retardado” (categoria conceitual de QUALIDADE)

Retomando a questão da auxiliarização, e lembrando o pressuposto de que há regularidades nos processos de transferência conceitual, é razoável pensar em conceitos concretos básicos expressos por verbos que, por extensão metafórica, passam a expressar conceitos gramaticais, mais abstratos. Em outras palavras, as chamadas “fontes conceituais”, das quais derivam conceitos gramaticais como Tempo e Aspecto, são previsíveis. Kuteva 1991 (apud Heine 1993), por exemplo, constatou que em 11 línguas de 3 famílias distintas apenas 20 verbos foram empregados para codificar 117 construções diferentes como auxiliares.

A respeito da natureza dos conceitos, Heine 1993 considera que os conceitos-fonte mais básicos são:

Rótulo Proposto	Forma Conceitual
Locação	<i>X está em/ permanece em/ vive em Y</i>
Movimento	<i>X move-se para/ de / através de Y</i>
Ação	<i>X faz Y</i>
Volição	<i>X quer Y</i>
Mudança de Estado	<i>X torna-se Y</i>
Equação	<i>X é como / é igual a Y</i>
Acompanhamento	<i>X está com Y</i>
Posse	<i>X tem Y</i>
Modo	<i>X está de modo Y</i>

As chamadas “formas conceituais” acima são, de acordo com a literatura, “esquemas de eventos”. Langacker 1978, apud Heine 1993, define-as como unidades semânticas simples que se constituem de um predicado associado a variáveis, como em *x QUER y*.

O autor lembra que os conceitos acima não esgotam a experiência humana, mas afirma serem estes os mais básicos, estando todos envolvidos na emergência de verbos auxiliares. Além disso, ao constatar que nem todos são igualmente básicos, o autor distribui tais conceitos hierarquicamente, apontando os três primeiros na lista como os mais salientes na conceptualização humana.

Ainda que sem muita profundidade, Heine 1993 preocupa-se em verificar a correlação entre os processos de extensão metafórica e as respectivas mudanças sintáticas no desenvolvimento de verbos auxiliares, seguindo a lista de conceitos básicos que ele mesmo propõe. Tampouco vou me aprofundar, aqui, nessas questões. Cito, contudo, na forma de quadro sinótico, os conceitos gramaticais mais comumente derivados dos conceitos considerados básicos:

Esquemas de Eventos	Conceitos Gramaticais Derivados
Locação	progressivo, ingressivo
Movimento	ingressivo, futuro, perfeito, passado
Ação	progressivo, contínuo, ingressivo, completivo, perfeito
Volição	ingressivo, futuro
Mudança de Estado	ingressivo, futuro
Equação	resultativo, progressivo, perfeito, futuro
Acompanhamento	progressivo
Posse	resultativo, perfeito, futuro
Modo	progressivo

Finalmente, quanto a esse modelo, um problema teórico ainda pode ser levantado. Em princípio, a metáfora envolve um “pulo” direto de um domínio para o outro, de um conceito para outro, algo que seria difícil de conciliar com a natureza gradual do processo. Lembre-se, entretanto, o esquema da figura 1 acima, representando o encadeamento de conceitos para o caso de *be going to* no inglês: há uma zona de intersecção entre os mesmos, indicando que aí eles não se distinguem categoricamente. Trata-se, justamente, da operação de transferência conceitual. O caso do verbo *estar* vem abordado no item 1 da Parte II: a análise qualitativa dos dados.

Tal problema teórico - a antítese “gradual x instantâneo” na definição de metáfora dentro do quadro da gramaticalização - é aparente. De forma coerente com o esquema proposto por Heine 1993 e transcrito na figura 1, considera-se que a gramaticalização envolve na verdade dois mecanismos: por um lado, a transferência conceitual - que é metafórica e que estabelece, sim, relação direta entre diferentes domínios cognitivos; e, por outro lado, o que se chama de “reinterpretação induzida pelo contexto”, que é metonímica, dada a contigüidade entre as partes depreendidas no processo.

Em outras palavras, portanto, justifica-se o termo implicatura: observando o processo de gramaticalização num nível macro, verifica-se a metáfora operando; num nível

micro, entretanto, opera a reinterpretação, conforme os novos significados inferidos vão sendo convencionalizados.

1.2. A Morfologização

Conforme mostrou o quadro de Lehmann (figura 1), tendo passado pelo processo de auxiliarização, é uma tendência que o verbo entre na fase de morfologização, ou seja, pode ocorrer cliticização e até mesmo afixação do auxiliar em relação ao verbo em forma nominal com que constitui uma perífrase. Em geral, este processo envolve redução fonológica do auxiliar.

Alguns autores falam em redução fonológica como um critério de auxiliaridade. De acordo com Pullum & Wilson 1977 (apud Heine 1993), por exemplo:

“(...) auxiliaries tend to reduce in phonetic substance and/or lose the ability to carry distinctive tone or stress, and become clitics of some immediately adjacent constituent”; (...) main verbs never do so, even though they may be homonyms of the auxiliaries”.

A teoria sobre gramaticalização de fato prevê, para o caso dos verbos, que a partir de seu uso como marcador gramatical, o auxiliar tende a sofrer erosão, ou seja, sua substância fonética é passível de ser reduzida, tornando-se mais dependente do verbo principal. Heine 1993:106 atenta para o fato de que a gramaticalização pode afetar apenas alguns usos de um determinado verbo, ou apenas alguns contextos em que o mesmo aparece, de modo que a erosão tende a ocorrer em tais contextos; em outros, pode perfeitamente persistir a forma fonologicamente plena.

Pode-se entender erosão, em linhas gerais, como uma simplificação da substância fonética empregada para a expressão de um dado conceito. “Simplificação” tem tanto uma dimensão temporal - no sentido de que a redução das articulações é seguida de uma diminuição no tempo que as mesmas levam para ser pronunciadas - quanto uma dimensão dita substantiva - diminuição dos gestos articulatórios. De acordo com Heine 1993, os efeitos da erosão podem ser:

- itens polissilábicos tornam-se monossilábicos;
- sons complexos, como encontros consonantais, são substituídos por sons mais simples;
- fonemas longos/geminados são substituídos por fonemas curtos/não-geminados;
- fonemas segmentais dão lugar a não-segmentais (supra-segmentais), ou são perdidos;
- vogais nasais tendem a ser substituídas pelas orais correspondentes;
- tons de contorno (e.g., low - high) tendem a ser substituídos por tons de registro (e.g., low);
- a marcação de acento tônico tende a ser perdida.

Hoje em dia, no português falado no Brasil, são fartos exemplos como os que seguem:

(a) *é justamente isso que ele **ta** fazendo, entendeu?*

(b) *eu tenho que estar atualizado (...) **ta** entendendo*

(c) *eu acho muito legal tu **ta andando** na rua e encontrar alguém conhecido*

(d) *A Tatá **taa contando** outro dia né (...) que é muito bonita a topografia da cidade*

Nestes exemplos, o verbo *estar* em sua forma fonologicamente reduzida - *tá* em (a, b, c) e *taa*, em (d) - substituiu a forma fonologicamente plena - respectivamente *está* e *estava*. Em (a), (b) e (c), não se marcou a acentuação gráfica da vogal de *ta* a fim de ilustrar que, hipoteticamente, tal monossílabo não é tônico em seu conjunto com o gerúndio, o que figuraria um forte argumento para a sua cliticização à forma nominal. Em (d), suprimiu-se a consoante “v” do auxiliar conjugado no pretérito imperfeito (de *tava* para *taa*), no intuito de grafar o mais fielmente possível o que o falante da língua teria pronunciado.

Em qualquer um desses exemplos, contudo, fica claro que a primeira sílaba do auxiliar foi “apagada”, não tendo sido pronunciada pelo falante. Dessa forma, no que diz respeito ao estágio da morfologização, a questão que os dados suscitam é a seguinte:

- já se pode afirmar que os conjuntos *tá* + *gerúndio* constituem uma única palavra fonológica, sendo *tá* um clítico em relação ao gerúndio, e provável candidato a afixo?

O encaminhamento a uma resposta a tal pergunta é desenvolvido na Parte II, quando os exemplos acima e outros são mais detalhadamente analisados.

Teoricamente, contudo, uma questão que parece mais espinhosa e que se põe aqui diz respeito à relação entre forma e significado no caso da redução fonológica. Alguns autores, a exemplo de Bybee & Perkins & Pagliuca 1922:22, apud Heine 1993, admitem uma forte correlação entre ambos, correlação esta de natureza icônica. Sendo assim, ainda que não o afirmem categoricamente, tais autores consideram que a redução fonológica está fortemente relacionada à redução do conteúdo semântico que a precede, nos moldes do “modelo da perda” e de acordo com o princípio da quantidade, já abordados acima: para a expressão de uma menor substância semântica, pode-se lançar mão de menor substância fonológica.

Por outro lado, pensando nos termos do modelo da implicatura, e portanto sem falar necessariamente em “redução do conteúdo semântico”, mas sim em extensão metafórica, como correlacionar forma e significado?

Braga 1994:17 fala numa “*relação inversa entre a dimensão dos vocábulos e sua frequência de uso*”, ou seja, quanto mais usado um vocábulo maior é a tendência em “encurtá-lo”. Esta é uma formulação feita inicialmente por Zipf 1935 (apud Heine 1993:110), de acordo com quem:

“people will shorten the linguistic expressions that are used most commonly for economy, that is to simplify their linguistic utterances”.

Ora, marcadores de conceitos gramaticais, tais como Tempo e Aspecto, tem uma frequência de uso maior que a dos verbos - enquanto itens lexicais - que lhes deram origem.

De acordo com esse mesmo ponto de vista, Lehmann 1974 (apud Heine 1993:110) afirma que, quanto maior a probabilidade de ocorrência de um “símbolo linguístico”, menor é o valor da informação por ele expressa e, quanto menor o valor de uma informação, menor é a quantidade de material fonológico necessária para expressá-la.

Heine 1993, ao referir-se a esses dois pontos de vista sobre a correlação “forma X significado” - o princípio da iconicidade e do isomorfismo, de um lado, e o princípio da economia, relacionado ao valor da informação, de outro - enfatiza que ainda não há meios para decidir qual deles é mais básico e genérico. O autor afirma:

“(...) we are still far from being able to draw on a model that would allow us to view erosion in particular and the relationship between form and meaning in general in a wider pattern”.

Todavia, no caso de *estar* + *gerúndio*, tema deste trabalho, é inevitável questionar:

- a redução fonológica do auxiliar *estar* ocorreu em consequência da mudança semântica desenvolvida no estágio inicial do processo de gramaticalização? (Independentemente do modo como tal estágio é entendido - em termos de perda ou de extensão metafórica, de acordo com os modelos delineados em 1.1.1 e 1.1.2 acima).

Ou, em contrapartida:

- estaria a redução fonológica correlacionada, antes, a fatores prosódicos?

Nos termos desta segunda pergunta, seriam fatores de natureza prosódica - como o ritmo - que viabilizariam a redução fonológica de *estar* na perífrase de gerúndio, para o caso do português falado no Brasil.

Hipoteticamente, a redução do verbo *estar* teria acontecido inicialmente naqueles casos em que o conjunto *tá+-ndo* viesse a constituir uma única palavra fonológica formada por dois troqueus, o pé básico do português. Em (a), por exemplo, “*tafazendo*” seria uma única ω (palavra fonológica), encontrando-se o acento primário sobre a sílaba “*zen*” e o acento secundário sobre *ta*.

(a) *é justamente isso que ele ta fazendo, entendeu?*

Nesse sentido, estaria seguindo a hipótese de Carvalho (1989, 1988/1992), apud Galves & Abaurre 1997, de que no português brasileiro (PB) o acento secundário segue um padrão binário, preservando-se a estrutura morfológica na pronúncia da palavra, por oposição ao português europeu (PE), em que tal acento é atribuído à primeira sílaba da palavra, mesmo quando há três sílabas antes do acento primário.

Considerar que a redução fonológica está correlacionada a fatores prosódicos é uma hipótese que se enquadra numa teoria fonológica chamada teoria dos domínios. De acordo com a proposta de Nespor & Vogel 1986, com tal teoria pode-se desviar a atenção, nas abordagens de fenômenos de ordem fonológica, dos sistemas de regras para princípios mais gerais que estariam regendo a aplicação de processos gramaticais.

Já nos primeiros parágrafos da introdução ao seu trabalho, as autoras fazem referência enfática à existência de hierarquia na organização dos constituintes do componente fonológico da gramática, no sentido de argumentar a favor dos domínios:

“According to prosodic theory, the mental representation of speech is divided into hierarchically arranged chunks. In the typically continuous flow of speech, such mental chunks, the prosodic constituents of the grammar, are signaled by different types of cues ranging from actual segmental modifications to more subtle phonetic changes”.

Com isso, as autoras querem dizer que cada constituinte prosódico funciona como o domínio de aplicação de regras fonológicas ou de processos fonéticos específicos. Por constituinte prosódico, entenda-se uma unidade lingüística complexa de natureza binária, constituída de uma cabeça (dominante) e um ou mais dominados. A teoria prevê que não há isomorfismo entre os constituintes prosódicos e os constituintes da morfologia ou da sintaxe, tendo cada um suas próprias regras e princípios:

“(...) an adequate theory of phonology must provide a way of making reference not only to the morpho-syntactic bracketings of the surface syntactic structure, but also to other syntactic as well as semantic notions. This is achieved in the model presented in this book by means of the mapping rules that group the terminal elements of a string in a way that creates units that are not necessarily in a one-to-one relation with the constituents of the morpho-syntactic hierarchy.”

Segundo Bisol 1996, apresento os constituintes prosódicos e sua organização hierárquica:

- sílaba	(σ)	
- pé	(Σ)	
- palavra fonológica	(ω)	
- grupo clítico	(C)	
- frase fonológica	(Φ)	
- frase entonacional	(I)	(do inglês <i>intonational</i>)
- enunciado	(U)	(do inglês <i>utterance</i>)

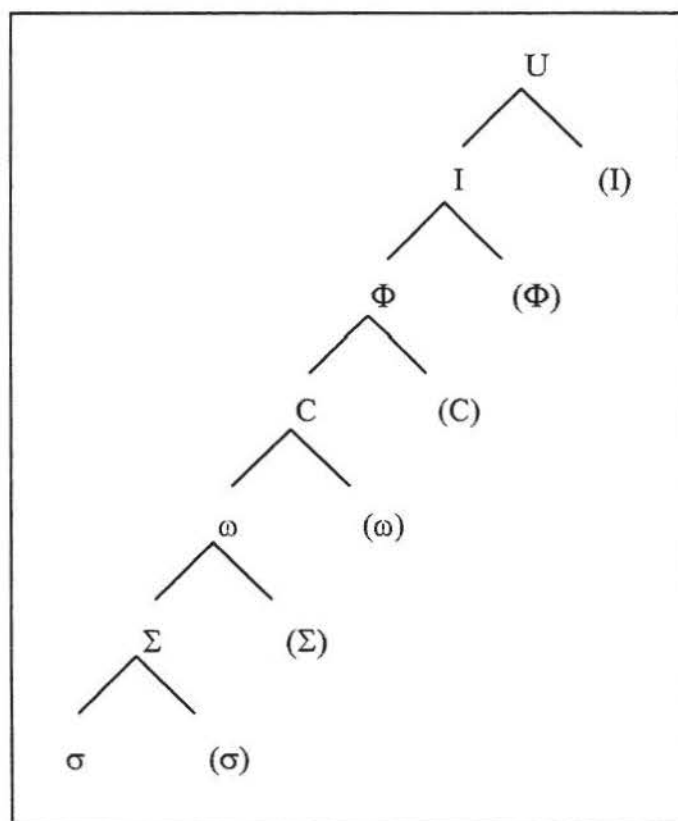


Figura 3 - Organização Hierárquica dos Constituintes Prosódicos

Para regular tal hierarquia, a teoria prevê os seguintes princípios:

1. cada unidade (X^p) da hierarquia prosódica é composta de uma ou mais unidades da categoria imediatamente mais baixa (X^{p-1});

2. cada unidade está exaustivamente contida na unidade imediatamente superior de que faz parte;
3. os constituintes são estruturas n-árias;
4. a relação de proeminência relativa, que se estabelece entre nós irmãos, é tal que a um só nó se atribui o valor forte (s - *strong*) e a todos os demais o valor fraco (w - *weak*).

A regra geral para a formação do constituinte é chamada, por Nespor & Vogel 1986, de Construção do Constituinte Prosódico, e é formulada da seguinte forma:

- *Incorpore numa ramificação n-ária X^p todos os X^{p-1} incluídos em uma cadeia delimitada pelo domínio de X^p .*

Para o caso de *tá + -ndo*, os constituintes prosódicos de maior interesse são o pé métrico e a palavra fonológica. O primeiro é a relação de dominância que se estabelece entre duas ou mais sílabas, sendo que o PB constrói basicamente pés binários de cabeça à esquerda. Quanto à segunda, transcrevo a definição dada por Bisol 1996:251 (as sublinhas são minhas):

“A palavra fonológica (...) é a categoria que domina o pé. Por exigência dos princípios que regem a hierarquia prosódica, todos os pés de uma cadeia, e nenhuma outra categoria, são agrupados em palavra fonológica. Além disso, sendo a palavra fonológica ou prosódica um constituinte n-ário, tem ela um só elemento proeminente, do que se conclui que a palavra fonológica não pode ter mais do que um acento primário. Porém, dentro do domínio da palavra fonológica, pode ocorrer reagrupamentos de sílabas e pés, sem compromisso de isomorfia com os constituintes morfológicos.”

Devo dizer que há certa polêmica acerca dos clíticos. Isto porque pode-se entender o clítico de duas maneiras: (i) é um elemento da palavra fonológica, de modo que “fala-se”, “te considero” e “o leque” constituem, cada um, uma única palavra, dada sua atonicidade;

ou (ii) é um elemento independente, sendo sujeito às regras fonológicas específicas ao domínio - como neutralização, sândi, etc.

Como vimos, a proposta de Nespor & Vogel considera o grupo clítico como um constituinte prosódico com um lugar definido na hierarquia. Outras propostas, entretanto, desconsideram tal constituinte. Com respeito a isso, Bisol 1996:25 toma a seguinte posição: *“tomar o clítico junto à palavra adjacente por locução (...) ou tomá-lo como parte da palavra fonológica é ainda uma questão em aberto”*.

Apesar dessa ser uma questão teórica das mais importantes dentro da Fonologia Prosódica, não vou ocupar-me em detalhá-la. Entendendo que *tá* pode se comportar como um clítico em relação ao gerúndio, importa aqui perceber que *tá* + *-ndo* constituem um único domínio prosódico. Vou considerar que tal domínio é o da palavra fonológica, como nos exemplos citados por Bisol 1996:252⁵:

	Uma só palavra fonológica
te considero	[te konsideru]ω
me leve	[me leve]ω
o leque	[o lequi]ω
ta fazendo	[ta fazedu]ω

Para constituir uma única palavra fonológica com o gerúndio, *ta* deve perder sua tonicidade dentro do domínio. Neste caso, como já foi dito anteriormente, a sílaba “-zen-” teria o acento primário, e a sílaba “*ta*” o acento secundário da palavra como um todo.

Tendo em mente as linhas gerais dessa teoria, e voltando agora à questão da redução fonológica, pode-se hipotetizar que tal redução teria ocorrido inicialmente naqueles casos

⁵ Os três primeiros exemplos da lista foram extraídos de Bisol 1996:252, onde também vem detalhada a questão teórica sobre os clíticos.

em que o conjunto *ta* + *-ndo* constituísse uma palavra fonológica formada pelos dois pés básicos no ritmo do PB - o troqueu - exatamente como no caso

[ta fazedo]_ω

(* .) (* .)

Essa seria uma hipótese alternativa para análise do fenômeno, se lembramos que a grande maioria dos autores fala em redução fonológica como uma consequência do processo de abstratização semântica do item em processo de gramaticalização, conforme já vimos acima. É óbvio que não se trata de absolutamente desconsiderar a correlação semântica-fonologia nos fenômenos de redução. Trata-se, ao contrário, de considerar a possibilidade de que, além da interface semântica-fonologia, também atuem regras prosódicas, que podem desencadear processos de mudança na gramática.

A análise quantitativa dos dados para verificar se fatores prosódicos estão de fato correlacionados ao uso de “*tá*” vem relatada na Parte II.

2. Mais, ou menos gramaticalizado ?

Vimos, no item 1 acima, que o processo de gramaticalização é esquematicamente representado pela disposição linear das fases que o compõem (figura 1). É bastante lógico pensar, então, que um item comprovadamente em processo de gramaticalização vai estar mais ou menos gramaticalizado conforme sua “posição” nessa escala. Simplificadamente falando: mais próximo do início da escala, menos gramaticalizado; mais próximo de seu término, mais gramaticalizado.

Nem sempre é fácil, contudo, ser tão simples na resposta à pergunta-título deste item 2. Devemos lembrar que as fases do processo de gramaticalização não são entidades discretas, com domínios indiscutivelmente definidos. Entre outros, Heine and Reh 1984 (apud Heine & Claudi & Hünemeyer 1991:16) atentam para o fato:

“Grammaticalization is an evolutionary continuum. Any attempt at segmenting it into discrete units must remain arbitrary to some extent.”

Sendo assim, é possível falar em termos de “mais ou menos gramaticalizado”, ainda que se esteja abordando, teoricamente, uma única fase do processo.

De uma maneira mais geral, ainda que se critique e se refute a abordagem da gramaticalização como fenômeno de “perda”⁶, é comum os estudiosos considerarem que: quanto mais uma unidade linguística sofre gramaticalização

- i) mais ela perde em complexidade semântica, significância funcional, e/ou valor expressivo;
- ii) mais ela perde em termos pragmáticos e mais ganha em termos de significância sintática;
- iii) mais reduzido é o número de membros pertencendo ao mesmo paradigma morfossintático;
- iv) mais sua variabilidade sintática diminui, ou seja, mais a sua posição se torna fixa dentro da frase;
- v) mais o seu uso se torna obrigatório em certos contextos e agramatical em outros;
- vi) mais ela coalesce semanticamente, morfossintaticamente e foneticamente com outras unidades;
- vii) mais ela perde em substância fonética.

Alguns autores, contudo, dedicam-se especialmente ao estabelecimento de critérios que sirvam à aferição do grau de gramaticalização em que se encontra uma dada unidade. É o que passo a referenciar.

2.1. Os parâmetros de Lehmann

Lehmann 1982a:125-179, preocupado em encontrar um meio de medir sincronicamente a “gramaticalidade” de uma forma - especificamente o signo linguístico - estabelece seis parâmetros, os quais reúno no quadro a seguir, a exemplo de Heine & Claudi & Hünemeyer 1991:

⁶ As palavras desta lista que foram sublinhada remetem, com maior ou menor evidência, à gramaticalização enquanto um processo de perda. Note-se que, apesar de criticado e, não poucas vezes, refutado, o critério da perda está presente nos estudos.

Parâmetro	Gramaticalização Fraca	Processo	Gramaticalização Forte
Integridade	Conjunto de propriedades semânticas; possivelmente polissilábico	Desgaste	Poucas propriedades semânticas; monosegmental
Paradigmaticidade (Coesão Paradigmática)	Dentro de seu paradigma, o item participa escassamente no campo semântico	Paradigmatização	Fortemente integrado no paradigma
Variabilidade Paradigmática	Livre escolha dos itens, de acordo com as intenções comunicativas	Obrigatorização	Escolha sistematicamente restrita; uso largamente obrigatório
Escopo	O item é constituinte de construções de tamanho variável	Condensação	O item modifica uma palavra ou raiz (reduz-se o tamanho do escopo)
Conexidade	O item é justaposto independentemente	Coalescência	O item é afixo ou até mesmo suporte de traço fonológico
Variabilidade Sintagmática	A posição do item nas estruturas é mais livre	Fixação	O item procura lugares gramaticais fixos

Estes parâmetros estão, na verdade, estabelecidos sobre a base de três fatores que dizem respeito justamente à autonomia de um signo linguístico, que seriam o seu “peso”, sua “coesão” e sua “variabilidade”, de acordo com o quadro abaixo:

	Parâmetros Paradigmáticos	Parâmetros Sintagmáticos
peso	integridade	escopo
coesão	paradigmaticidade	conexidade
variabilidade	var. paradigmática	var. sintagmática

Lehmann 1982 reconhece como peso paradigmático, ou *integridade* de um signo, o seu “tamanho” substancial - tanto em termos semânticos quanto em termos fonológicos -

que lhe permite manter a identidade. O autor lembra que o termo *desemantização* se aplica justamente à diminuição da integridade semântica, e que *atrição fonológica* se aplica, da mesma forma, à diminuição da integridade fonológica. Tais processos já foram comentados anteriormente, ainda que a partir de outros rótulos, respectivamente *bleaching* (dentro do “modelo da perda”) e *redução fonológica* ou *erosão*.

Por peso sintagmático, ou *escopo*, entende-se o tamanho estrutural da construção que um item ajuda a formar. Para que fique claro, aproveito um exemplo simples fornecido por Lehmann: na língua turca, um sufixo de caso tem como escopo o NP, enquanto que, no latim, um sufixo de caso tem como escopo o próprio nome. Teoricamente, quanto mais avançada a gramaticalização, mais condensado será o seu escopo.

Em seguida, temos o fator coesão. A coesão paradigmática é entendida, de um lado, como a integração semântica de um paradigma enquanto um todo, e de outro, como a integração de uma sub-categoria dentro de seu paradigma. De acordo com Lehmann, o aspecto mais superficial e evidente da coesão paradigmática, ou *paradigmaticidade*, é o que comumente chamamos de “tamanho do paradigma”.

Quanto à coesão sintagmática de um signo, trata-se do grau de “intimidade” com que ele se conecta a outros itens numa relação sintagmática. Este grau pode variar desde a justaposição até a fusão de um item ao outro, conforme o avanço do processo de gramaticalização.

Finalmente, temos o fator variabilidade. A variabilidade paradigmática é, em linhas gerais, a liberdade com que o usuário de uma língua escolhe um item dentro de um paradigma. Seguindo Lehmann, conforme o avanço da gramaticalização de um item, o mesmo tende a tornar-se obrigatório, de modo que a liberdade de escolha dentro de seu paradigma diminui.

Já a variabilidade sintagmática consiste, nos dizeres de Lehmann, na facilidade com que um item pode se mover dentro de um contexto. Trata-se, em outras palavras, da mutabilidade posicional que um item apresenta em relação a outro, dentro da construção que formam. Tal variabilidade diminui conforme o avanço do processo de gramaticalização.

No entender de Moura Neves 1997, o que Lehmann chama de “processos” no quadro supra seriam tendências que, operando conjuntamente, caracterizariam a gramaticalização de um item.

Pensando na perífrase de *estar* + *-ndo*, parece que o *estar* se encontra justamente no processo em direção a uma forte gramaticalização, como exemplifico abaixo⁷:

- INTEGRIDADE: o verbo *estar* já passou por um processo de desgaste fonológico, como vimos anteriormente em 1.2, dada a sua redução fonológica; falar em desgaste semântico, por outro lado, é menos evidente, dada a discussão sobre as mudanças semânticas envolvidas na gramaticalização (modelo da perda X modelo da extensão metafórica acima comentadas);
- PARADIGMATICIDADE: *estar* integra, a partir de sua gramaticalização, o que podemos chamar de “paradigma dos auxiliares”, constituindo uma categoria, no sentido genérico, da qual fazem parte outros verbos, tais como *ter*, *ir*, *andar* e outros; em termos algo menos genéricos, *estar* constitui o paradigma dos auxiliares que entram em construções com a forma nominal de gerúndio, tais como *viver*, *andar*, *ficar*, *vir*, *etc.*, que pressupostamente também passaram por um processo de gramaticalização, podendo estar mais ou menos avançados neste processo em relação ao *estar*;
- VARIABILIDADE PARADIGMÁTICA: como vimos no parâmetro anterior, outros verbos podem ser usados com o gerúndio além de *estar*; nos termos desse parâmetro, não apenas *estar* pode vir a ser preferido em relação aos outros verbos, na expressão de aspectos como o imperfectivo, mas também a construção *tá* + *-ndo* pode se tornar obrigatória enquanto meio para a expressão desses aspectos.

A respeito desse parâmetro, devo lembrar, com Castilho & Moraes de Castilho 1994, que o aspecto verbal é uma categoria de natureza composicional em línguas como o

português, que não dispõe de morfologia específica para o mesmo. Dizendo a mesma coisa, nos termos de Lehmann, há uma variedade de “itens” que, na frase, desempenham a função de “aspectualizadores”, como advérbios, auxiliares, o morfema de gerúndio - {-ndo} - dentre outros. Com o avanço da gramaticalização de *estar* + *-ndo*, teoricamente a construção pode tornar-se “a forma” designada à expressão do aspecto imperfectivo, em determinados contextos. Não vou me aprofundar na questão do aspecto relacionado a *estar* + *-ndo*, para o qual remeto a Mendes 1994a e Mendes 1994b.⁸

- ESCOPO: trabalhos diacrônicos sobre os usos do verbo *estar*, tais como os de Kewitz 1994 e de Mendes 1996, mostram que no decorrer de sua história este verbo nem sempre teve como escopo gerúndio, podendo ter um PP ou um AdjP; sabemos que hoje em dia *estar* pode ter como escopo um SP com a função de locativo, como em:

(a) *eles estavam na casa do Fernando quando a gente chegou*

ou ainda um escopo mais complexo - composto por um AdvP, um PP e um gerúndio, como em:

(b) *nesse horário, eu estou sempre em casa estudando*

De acordo com esse parâmetro, teoricamente, o escopo de *estar* tende a condensar-se, com o avanço de sua gramaticalização. Assim sendo, num ponto futuro de sua diacronia, o verbo *estar* apareceria na frase acima tendo por escopo nada mais que o gerúndio:

(b') *nesse horário, eu sempre estou estudando em casa.*

⁷ Apesar de reservar uma parte específica para a análise qualitativa dos dados, optei por desenvolver tal exemplificação aqui mesmo, neste item, para fazê-lo imediatamente após a apresentação dos parâmetros de Lehmann. Maiores detalhes aparecem na Parte II

⁸ Seguindo Castilho & Moraes de Castilho 1994 em ambos os trabalhos referenciados, considero que a perífrase *estar* + *-ndo* pode desencadear processos diferentes, mas não necessariamente excludentes: o da quantificação (aspecto iterativo) e o da qualificação (aspecto imperfectivo) dos estados de coisas que elas descrevem.

- CONEXIDADE: Lehmann trata esse parâmetro sem correlacioná-lo ao que vimos anteriormente - o *escopo*, assim como o faz em relação aos outros parâmetros que propõe (ainda que deixe claro que os processos operem ao mesmo tempo). Para o caso de *estar* + *-ndo*, contudo, ESCOPO e CONEXIDADE parecem estar amplamente correlacionados; ora, entende-se que o verbo auxiliar, com o avanço de sua gramaticalização, tende a tornar-se mais dependentemente ligado à forma nominal; em outras palavras, e utilizando os mesmos exemplos acima, um maior grau de gramaticalização de *estar* pressuporia a preferência de (b') em relação a (b), por parte do falante. Outro exemplo seria o da negação:

(c) ...o cinema... eu acho que não *está evoluindo*...

(c') *...o cinema... eu acho que *está não evoluindo*...

A agramaticalidade do segundo exemplo mostra que a negação tem por escopo a perífrase como um todo, não podendo “desconectar” os termos da perífrase. A negação pode ter exclusivamente o gerúndio como escopo apenas nas predicções que aqui chamo de “reelaborativas”, como mostra o exemplo abaixo:

(c'') ...o cinema... eu acho que *está não evoluindo...está regredindo*

- VARIABILIDADE SINTAGMÁTICA: reiterando o que disse, no parágrafo anterior, sobre o fato de que os parâmetros de Lehmann podem ser entendidos como inter-relacionados uns aos outros, os casos de *estar* + *-ndo* parecem exemplificar este de maneira muito similar aos anteriores; se a posição de *estar* torna-se mais fixa conforme avance sua gramaticalização, entende-se que haverá menor possibilidade de inserção de locativos e de outros adjuntos adverbiais entre o auxiliar e o gerúndio; ou seja, a variabilidade

sintagmática de *estar*, está diretamente correlacionada ao parâmetro *escopo*, bem como ao parâmetro *conexidade*.

2.2 . Os “princípios” de Hopper

Hopper 1991 afirma que a proposta de Lehmann 1982, que acabamos de ver, constitui um bom meio para descrever aqueles processos de gramaticalização já adiantados, ou seja, aqueles casos de gramaticalização cuja ocorrência já não se questiona. Dito de outra forma, a proposta de Lehmann, segundo Hopper, aponta quais são as tendências, no nível paradigmático e sintagmático, na direção de uma gramaticalização forte.

Mais preocupado com os estágios incipientes dos processos de gramaticalização, que seriam “*menos facilmente acessíveis*” (Heine-Claudi- Hünnemeyer 1991), Hopper 1991 propõe então os seguintes cinco princípios:

- **ESTRATIFICAÇÃO:** quando novas camadas emergem dentro de um domínio funcional, camadas mais antigas não necessariamente são descartadas, podendo continuar a coexistir e a interagir com as camadas mais novas;

Trata-se, portanto, de constatar que “*uma das conseqüências da gramaticalização é a convivência de soluções gramaticais distintas, num mesmo recorte sincrônico*”, de acordo com Castilho 1996.

No que toca à perífrase *estar* + *-ndo*, podemos lembrar que tal construção não existia no Latim e, mesmo assim, aspectos como o imperfeito e o iterativo - o *domínio funcional* da perífrase - eram expressos naquela língua. Já no português, *estar* + *-ndo* não é a única perífrase de gerúndio que se presta à expressão de tais aspectos: *ficar* + *-ndo*, *andar* + *-ndo*, *vir* + *-ndo* são construções também freqüentes no PB.

Nesse sentido, temos “*soluções gramaticais distintas*” para a expressão de aspectos verbais “*num mesmo recorte sincrônico*”.

- **DIVERGÊNCIA:** quando alguma entidade lingüística sofre gramaticalização, passam a existir pares ou conjuntos maiores de formas que têm uma etimologia comum, mas que divergem funcionalmente.

Assim, se um item lexical se gramaticaliza num clítico, por exemplo, a forma lexical fonte permanece como tal, ficando sujeita a outras mudanças que lhe sejam passíveis, enquanto elemento autônomo.

O verbo *estar* ilustra esse princípio na medida em que, paralelamente ao seu uso como auxiliar, seja na forma fonologicamente plena, seja na forma fonologicamente reduzida, o mesmo é usado como cópula e como verbo intransitivo seguido de locativo, como já foi anteriormente exemplificado.

- **ESPECIALIZAÇÃO:** é o estreitamento das possibilidades de escolha que caracterizam uma construção gramatical em emergência.

De acordo com Castilho 1996, este “princípio” alude ao fato de que *“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”*.

O verbo *estar*, conforme já foi comentado no item dedicado aos parâmetros de Lehmann, não é o único que forma perífrase de gerúndio, havendo também *andar*, *viver*, *permanecer*, dentre outros. A função básica das perífrases de gerúndio é a de expressar os aspectos imperfectivo e iterativo, apesar das nuances semânticas apresentadas pelos outros auxiliares em relação a *estar*. Neste caso, ocorreria especialização se este último passar a ser o único selecionado para constituir tal perífrase.

- **PERSISTÊNCIA:** quando uma forma gramaticalizada B se desenvolve a partir de uma forma lexical A (fonte), esta forma lexical, ou ainda alguns de seus traços semânticos, podem permanecer;

Considerando que as construções de *estar* + *locativo* aparecem antes de *estar* + *gerúndio* na diacronia, a ocorrência atual de um caso como

...nesse horário, estou na minha casa estudando...

ilustra bem o princípio. Ora, neste exemplo, considera-se *estar* como um auxiliar - teoricamente um marcador de Tempo - mas suas propriedades intensionais de locativo persistem, uma vez que se aceita a “inserção” do PP *na minha casa* entre os dois membros da perífrase.

- DESCATEGORIZAÇÃO: a gramaticalização leva a um decréscimo da categorialidade da forma em questão; isso implica uma perda de marcas de categorialidade, de um lado, e de autonomia discursiva, de outro.

Em outras palavras, e para exemplificar, um verbo pleno que se gramaticalize num afixo pode perder marcas morfológicas como número e pessoa, que são marcas de verbos enquanto categorias plenas.

Esse não chega a ser o caso da forma *tá*, resultado da redução fonológica de *estar*. Apesar da forma reduzida ser largamente usada (ver análise quantitativa em II), ainda não se perderam as marcas de número e pessoa:

eu to fazendo o meu trabalho
ele tá fazendo o trabalho dele
nós tamó fazendo o nosso trabalho
** eu tá fazendo o meu trabalho*

De todo modo, pode-se falar em descategorização na passagem de verbo pleno para verbo auxiliar, pois o auxiliar não tem as mesmas propriedades semânticas e sintáticas próprias à categoria dos verbos, como por exemplo subcategorizar um argumento interno.

Tanto os princípios de Hopper, quanto os parâmetros de Lehmann, anteriormente discutidos, são diferentes tentativas para medir o grau de gramaticalização de um termo. Todavia, em cada uma das propostas, não fica muito claro quantos dentre os parâmetros / princípios propostos devem ser exemplificados por uma forma, para que a mesma seja considerada mais ou menos gramaticalizada. Ou então, para não falar em “quantidade”, tampouco os autores deixam claro se este ou aquele parâmetro/princípio deveria ser preferencialmente exemplificado por uma forma para que fosse atestada sua gramaticalização.

Mesmo no caso de Lehmann, que fala em gramaticalização forte e fraca, caracterizando aquilo que seria o intermédio do processo, nem sempre é fácil compreendê-lo em toda sua extensão, dado o fato de que cada um dos parâmetros é exemplificado com diferentes casos de gramaticalização.

Nesse sentido, para estudar um caso isolado de gramaticalização, acredito que tais parâmetros possam ser tomados antes no sentido de se descrever fatos relativos ao processo do que propriamente para medir o grau de gramaticalidade de uma determinada forma. Em outras palavras, o fato de que um reconhecido caso de gramaticalização não exemplifique exatamente algum ou alguns dos parâmetros/princípios acima, não significa necessariamente que tal caso deva ser questionado.

PARTE II - ANÁLISE DOS DADOS

Nesta parte, desenvolvo a análise da perífrase *estar* + *-ndo* dentro do quadro da gramaticalização, teoricamente abordada na parte anterior.

No item 1, a análise é de caráter qualitativo, onde abordo o processo de auxiliarização de *estar* e seu avanço no processo de gramaticalização. Para tal análise, contei com dados extraídos basicamente das seguintes fontes:

- as entrevistas do tipo D2 (Diálogo entre dois Informantes) do *corpus* compartilhado do Projeto NURC/SP (Preti & Castilho 1986);
- algumas entrevistas do Projeto PEUL/RJ;
- entrevistas gravadas por alunos do Prof. Dr. Ataliba de Castilho, por ocasião do curso de Língua Oral por ele ministrado na FFLCH/USP em 1997;
- o trabalho de Chaves et alii 1996: “Diacronia das Perífrases Verbais de Estar + gerúndio/ + infinitivo / e + particípio”;
- o trabalho de Kewitz et alii 1994: “Estudo diacrônico de ser e estar em construções não perifrásticas”¹.

Apesar de ser esta uma análise qualitativa, eventualmente insiro nela resultados quantitativos obtidos nas análises diacrônicas desenvolvidas por Chaves et alii 1996 e por Kewitz et alii 1994.

No item 2, por fim, apresento os resultados da análise quantitativa, para a qual usei apenas os dados obtidos das entrevistas do NURC/SP. Tais dados foram codificados e submetidos ao pacote estatístico VARBRUL.

Quanto às árvores sintáticas que inseri nessa parte, devo dizer que elas são representações feitas nos moldes da Teoria Gerativa, mais especificamente de acordo com Haegeman 1991. Como mencionei na introdução, este trabalho foi desenvolvido a partir de um ponto de vista funcional da língua - cujos pressupostos teóricos são notadamente diferentes daqueles preconizados a partir de um ponto de vista formal. A representação

¹ Os dados para estes dois últimos trabalhos foram extraídos do *Corpus Diacrônico do Português*, organizado por Tarallo 1991.

arbórea aqui utilizada, contudo, não compromete as bases teóricas sobre as quais se desenvolveu o trabalho.

1. Análise Qualitativa

Dados da língua escrita mostram que nem sempre *estar* foi usado com a forma nominal de gerúndio. Vejam-se os exemplos abaixo. Eles vêm agrupados em quadros, de acordo com o tipo de construção em que o verbo *estar* figura e, no final de cada frase, vem indicado o século de que a mesma faz parte. Todos os exemplos do século XIII ao XIX são de língua escrita, enquanto que todos os do século XX são de língua falada:

(01) *Maltreito de sas feridas, Galaaz non estede.* (século XIII)

(02) *Entom disse [a] Agravaim: estade até que vos diga ~uu pouco, e el esteve* (séc. XIII)

Quadro A

(03) *que gran demorança aqui u estamos* (séc. XIII)

(04) *estava hi h~uu cidadão que avia enfermidade de febre* (séc. XIV)

(05) ... *da capitania de Sto. Amaro, em cujo districto e jurisdição as ditas terras estavam*
(séc. XVII)

(06) *quando eu estava lá no Rio eu ia ... lá da ... na Praça XV até Paquetá* (séc. XX)

(07) *estamos num país em que durante alguns anos não houve prova de redação* (séc. XX)

Quadro B

(08) ...vaso que **está** cheo (séc. XIV)

(09) ... e vi **estar** uu mancebo mui fremoso, vestido de ataduras de fogo mui esprandecente... (séc. XV)

(10) todos **estais** enganados (séc. XVI)

(11) é uma cidade que tem 100 annos e **está** podre (séc. XIX)

(12) **estou** brigado com o telefone (séc. XX)

Quadro C

(13) de pos de mia morte, mia molier e meus filios e meu reino e meus uassalos e todas aquelas cousas que Deus mi deu en poder **sten** en paz e en folgãcia (séc. XIV)

(14) O Fulano **está** bem (séc. XX)

Quadro D

(15) e todo o poboo **estava chorando** ante o tabernaculo de Nostro Senhor (séc.XV)

(16) de tudo fizerão sabedor a elle dito capittão que já neste tpõ **estava ouuindo** besperas dentro da Igreja (séc. XVII)

(17) esta terra **está confessando** em altas vozes que nunca governador nenhum trouxe comitiva semelhante (séc. XVIII)

- (18) *Parabéns pelo espírito que você **está derramando** a pennas cheias* (séc. XIX)
- (19) *eu não **estou privando-os** de um prazer maior ao levá-los para um sítio* (séc. XX)
- (20) ***está havendo** aí uma grande campanha* (séc. XX)

Quadro E

- (21) *E, **estando** (ele) a hũa feestra **rogando** Nosso Senhor e **louvando-o** mui de coração, viu hũa luz viir* (séc. XIV)
- (21a) *enquanto **estava** junto da janela, rogando e louvando, viu uma luz vir*
- (22) *depois de **estarmos** aly **acabando** de espalmar as fustas* (séc. XVI)
- (23) *e tambe neste porto **estive** o draque desne vinte de junho ate dezassete de agosto **esperando** por tempo no anno de 1578* (séc. XVII)
- (24) *Maria **estava** na Tupi **trabalhando** como ::... funcionária* (séc. XX)
- (25) *agora outro dia **estive** aquí na minha casa **me visitando** uma autora teatral jovem*
(séc. XX)

Quadro F

No quadro A, os dois exemplos aí contidos datam do século XIII. Nestas frases, *estar* é reconhecido como verbo pleno, detendo as seguintes propriedades:

- (i) pode significar “ficar de pé”;
- (ii) atribui papel temático /- agentivo/ ao seu sujeito;
- (iii) não subcategoriza argumento interno, operando como um verbo intransitivo
- (iv) participa da composição do aspecto permansivo².

A representação sintática para tais construções é a seguinte:

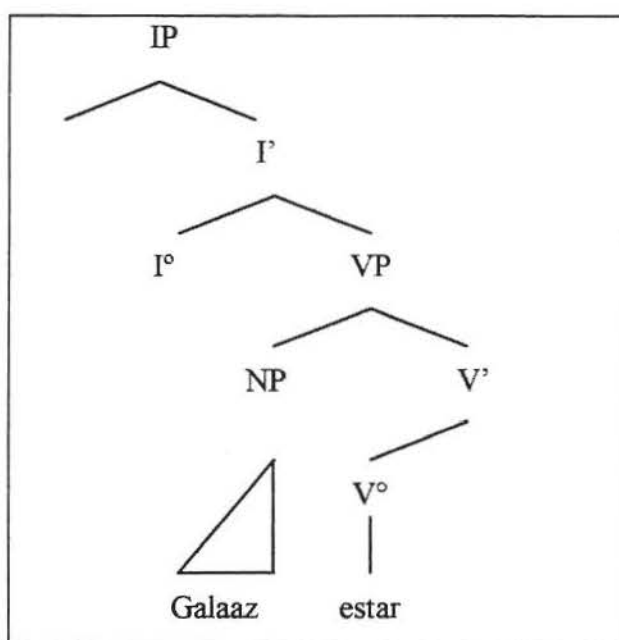


Figura 4 - Exemplo de representação sintática de *estar* enquanto V pleno

² De acordo com Castilho e Moraes de Castilho 1994, quando o estado de coisas é estático, como nas construções locativa, atributiva e equativa, o aspecto é o permansivo. Em sua proposta para o Aspecto Verbal em português, os autores incluíram o permansivo porque mesmo quando o estado de coisas é estático, tem-se a categoria do Aspecto atualizada na predicação: “O conceito de Aspecto (...) tem a autonomia que lhe é dada por sua natureza de símbolo, por sua função de representar o mundo. (...) é uma categoria nocional auto-sustentada, independente da enunciação”. Os estados de coisas que se estabelecem na predicação podem ser permansivos, operativos ou resultativos, sem que a perspectiva temporal tenha qualquer relevância.

O quadro B contém exemplos que vão desde o século XIII até o XX. Nestes, as propriedades de *estar* são as seguintes:

- (i) semanticamente, é semelhante ao *estar* pleno dos exemplos do quadro A; a diferença mais perceptível é a presença de um locativo (na forma de um sintagma preposicional ou adverbial);
- (ii) também atribui papel temático /- agentivo/ ao seu sujeito;
- (iii) em cada frase, o elemento que vem sublinhado - o locativo - pode ser entendido como argumento oblíquo;
- (iv) participa da composição, também nesses exemplos, do aspecto permansivo.

Os exemplos contidos nesse quadro são comumente rotulados de *construções locativas* (Kewitz et alii 1994), visto que sua estrutura sintagmática é, basicamente:

SN + estar + SP / dêitico

É claro que, do ponto de vista da estrutura de superfície, as sentenças desse quadro não são iguais, uma vez que o locativo se posiciona à esquerda ou à direita de *estar*. Quanto à sua estrutura profunda, a representação pode mudar conforme se considere o locativo:

(i) como argumento interno, subcategorizado pelo radical de *estar*:

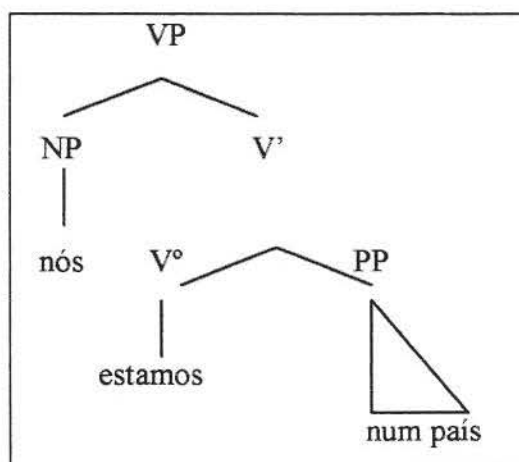


Figura 5 - *Estar* em construções locativas (i)

(ii) como núcleo do predicado³:

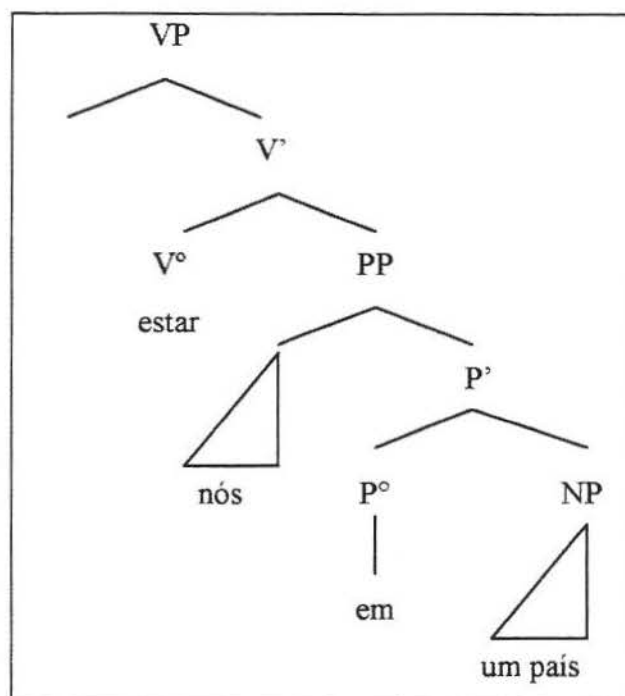


Figura 6 - *Estar* em construções locativas (ii)

Nos exemplos de (05) a (07), parece-me bastante claro que o locativo é uma sorte de argumento de *estar*, ou seja, o locativo seria subcategorizado pelo verbo. De fato, parecem agramaticais as seguintes sentenças, se comparadas a (06) e a (07) respectivamente:

(06a) * *quando eu estava eu ia da Praça XV até Paqueta.*

(07a) * *estamos*

³ Nesse caso, a representação sintagmática da representação seria:

estar [SN locativo]

da mesma forma que nas construções que tradicionalmente a gramática chama de “predicado nominal”. Tal rótulo não caberia aqui, uma vez que é um locativo que entra na construção. Neste trabalho, interessa levar em conta que tais exemplos podem ser representados em árvores de acordo com o conceito de mini-orações (*small clauses*), para cuja representação há várias propostas. Dentre elas, optei por seguir a de Stowell 1981, apud Raposo 1992.

Tais sentenças não seriam agramaticais apenas se o contexto deixasse claro o “lugar” a que se estaria fazendo referência.

Os exemplos (03) e (04), mais antigos, parecem ser semanticamente mais semelhantes às construções de *estar* pleno, do Quadro A. Entretanto, também nestes casos, o locativo parece ser essencial:

(03a) * *que gran demorança aqui estamos*

(04a) ? *estava huu cidadão que avia enfermidade de febre*

A agramaticalidade de (04a) é mais discutível, se considerarmos, por exemplo, que o *local* onde o *cidadão estava* é fornecido pelo contexto maior. Em (01) e em (02), diferentemente, não há presença de locativo. Por conta disso, a interpretação de *estar* como posicionamento vertical - a exemplo do verbo *stare* no Latim - fica facilitada em (01). Já em (02), *estar* não significa, necessariamente, *ficar de pé*, podendo ser parafraseada da seguinte forma:

(02a) *Então disse a Agravaim: “ficai até que vos diga um pouco”, e ele ficou.*

Evidentemente, *estar no lugar* - “permanecer” , no sentido da sentença (02) original - não pressupõe “ficar de pé”. Sendo assim, podemos já com os exemplos dos quadros A e B vislumbrar uma mudança nos conceitos expressos por *estar*:

posicionamento	>	localização (sem locativo)	>	localização (com locativo)
exemplo (01)		exemplo (02)		exemplos (03 - 07)

Passando para o Quadro C, temos exemplos chamados de *construções atributivas*, cuja estrutura é, basicamente

SN + estar + SAdj / participípio

Nesses casos, pode-se dizer que o verbo *estar*:

- funciona como cópula, transformando um sintagma adjetival no predicado da sentença;
- também atribui papel /- agenteivo/ ao sujeito.

O aspecto, entretanto, poderá ser

- permansivo - de (08) a (11)
- permansivo ou resultativo - em (12) - dada a “ambigüidade categorial” da forma nominal (verbo ou adjetivo?)⁴.

Em geral, *estar* é chamado de verbo funcional nessas construções. Tradicionalmente, a exemplo de Bechara 1992, propõe-se o termo verbo de ligação. De qualquer modo, a estrutura pode ser assim representada:

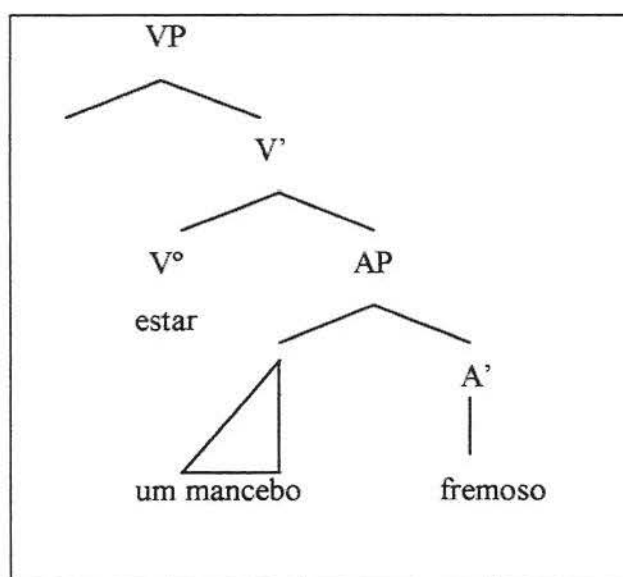


Figura 7 - *Estar* em construções atributivas

⁴ Chaves et alii 1996 propõem alguns critérios para a interpretação do participípio como adjetivo ou como verbo, nas construções em que aparece com *estar*. Um deles é o da graduação do participípio, como em:

a garganta está inflamada - a garganta está inflamadíssima

Para os autores, “nesta estrutura, o participípio teria intensificada sua face adjetiva, e sua face verbal estaria em um segundo plano”. Os autores também afirmam que, nesse caso, o Aspecto expresso não é o resultativo, ao contrário daqueles casos em que a face verbal do participípio é intensificada:

a Cantiga está feita

Os exemplos (08) e (09), entretanto, merecem mais algumas palavras. Em (08) temos, dentro corpus analisado por Kewitz et alii 1994, o exemplo mais antigo de construção atributiva com o verbo *estar*. As autoras, não exatamente preocupadas com o processo de gramaticalização, não se perguntaram se os primeiros casos de *estar* atributivo envolveriam preferencialmente um sujeito em posição vertical. Ora, em (08), *cheio* é uma qualidade atribuída a *vaso*, um objeto verticalmente posicionado.

Em outras palavras, significaria que as construções atributivas teriam sido inicialmente implementadas naqueles casos em que o verbo *estar* estabelecesse cópula entre um SN sujeito (na posição vertical) e um SAdj. A confirmação dessa hipótese constituiria um argumento a mais a favor da extensão metafórica, delineada na Parte I, nos moldes da teoria proposta por Heine & Claudi & Hünemeyer 1991.

Com relação ao exemplo (09), ele pode ser interpretado de duas formas diferentes, de acordo com as paráfrases abaixo:

(09a) *vi uu mancebo estar mui fremoso*

(09b) *vi uu mancebo mui fremoso estar*

Em (09a) temos uma construção atributiva, em que *estar* desempenha a função de cópula, cuja representação arbórea acima serviu de exemplo para os outros casos do mesmo quadro. Por outro lado, no exemplo (09b), teríamos o verbo *estar* como intransitivo, indicando a posição do sujeito, como no antigo exemplo (01).

De qualquer maneira, os dados permitem traçar o seguinte *continuum*:

posição	>	localização espacial	>	localização qualitativa
exemplo (01)		exemplos (02 - 07)		exemplos (08 - 12)
<i>estar de pé</i>		<i>estar onde? (em que lugar?)</i>		<i>estar como? (de que maneira)</i>

Justifico o termo localização qualitativa entendendo que “qualificar” - atribuir uma qualidade a algo ou alguém - é um modo de localizar, de alocar, ainda que mais abstrato. Veja-se a possibilidade de dizer a frase-sermão “*Fulano te colocou no teu lugar*”, após ouvir o tal *Fulano* ralhando com alguém (por qualquer motivo) e dizendo-lhe palavrões mais ou menos corriqueiros que, bem ou mal, qualificam.

Os dados quantitativos apresentados pelas autoras confirmam tal escala, pois indicam que as construções locativas com *estar* são mais antigas que as atributivas, as primeiras predominando em relação às segundas por três séculos. Apresento tais dados mais adiante.

Na sequência dos exemplos, os do Quadro D - (13) e (14) - também podem ser entendidos como construções atributivas. Entretanto, separei tais exemplos daqueles do Quadro C porque nestes temos, no lugar de sintagmas adjetivais, um sintagma preposicional com valor de advérbio de modo (13) e um sintagma adverbial propriamente (14).

SN + estar + SP / Sadv

Apesar disso, tais construções podem ser representadas do mesmo modo como vem ilustrado na Figura 6 acima: os exemplos (13) e (14) enquadram-se na pergunta “*estar de que maneira?*”, o que nos permite entendê-los como casos de localização qualitativa.

A partir do exemplo (15), temos casos de *estar + gerúndio*. Sobre os exemplos do Quadro E, que vão do século XV ao XX, pode-se dizer inequivocadamente:

- o verbo *estar*, por si só, não atribui papel temático a um sujeito; ele o faz em conjunto com o verbo no gerúndio;
- também não subcategoriza, por si só, argumento interno;
- ao sujeito não é atribuído, necessariamente, o traço /-agentivo/;
- participa na composição dos aspectos imperfectivo e iterativo.

Tomando o exemplo (15) como modelo, temos a seguinte representação sintática:

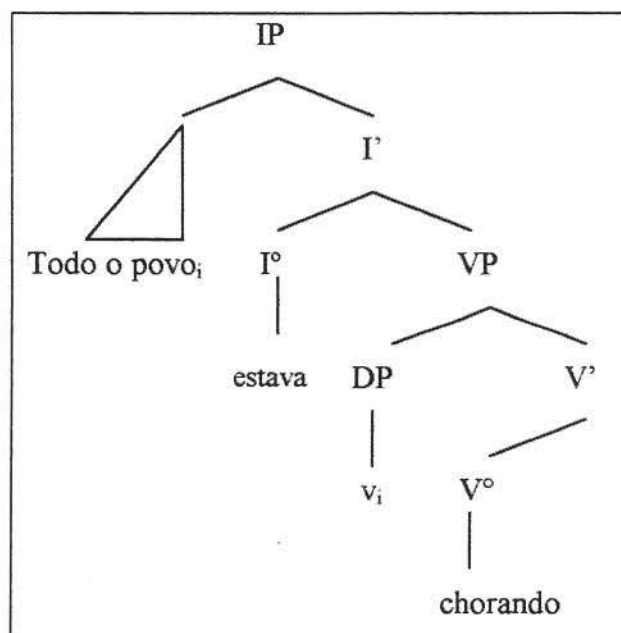


Figura 8 - *Estar* em construções com gerúndio

Dessa forma, nos exemplos do Quadro E, o verbo *estar* funciona como flexão, sendo o núcleo do IP (*inflectional phrase*). O termo *auxiliar* se aplicaria justamente a esses casos.

Os exemplos agrupados no Quadro F, por sua vez, ainda que possam ser analisados da mesma forma que os do quadro E, apresentam diferenças claras. A presença de um locativo entre o *estar* e o *gerúndio* - com exceção do exemplo (23), em que o locativo está à esquerda de *estar* - remete a uma interpretação sintática semelhante àquela representada na figura 5. Veja-se a representação do exemplo (24):

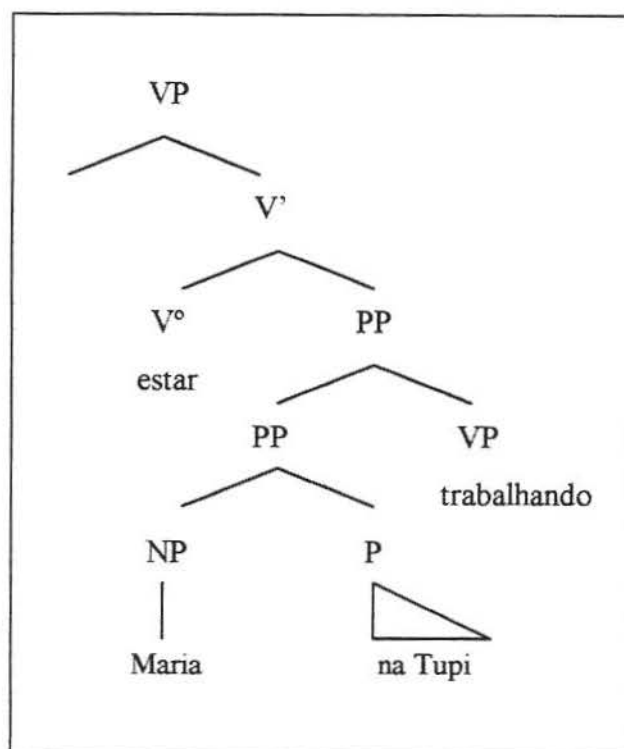


Figura 9 - *Estar* + locativo + gerúncio

Nesta representação, *trabalhando* está numa relação de adjunção com a mini-oração [Maria na Tupi]. O verbo *estar*, portanto, não ocupa nessa sentença o lugar da flexão, como no caso da figura 7. Esta relação de adjunção também é bastante perceptível em (21a), paráfrase de (21), em que se pode dizer que *rogando* e *louvando* funcionam como adjuntos adverbiais de modo.

Por outro lado, apesar da proposta de estruturação sintática acima, os exemplos do Quadro F são considerados ambíguos, do ponto de vista da gramaticalização. Todos esses exemplos têm uma estrutura semelhante, resumida no quadro abaixo:

<i>ESTAR</i>	<i>Em que lugar?</i>	<i>De que maneira?</i>
(21)	a hua feestra	rogando/louvando
(22)	aly	acabando
(24)	na Tupi	trabalhando
(25)	aqui na minha casa	visitando

O exemplo (23), que data do século XVII, não foi inserido no quadro acima por ser um pouco diferente: além do locativo *neste porto* ter sido colocado antes do verbo *estar*, há um adjunto adverbial de tempo, entre este e o gerúndio. Neste exemplo, então, teríamos, teríamos um dado a mais - além da localização espacial, um alocalização temporal:

<i>ESTAR</i>	<i>Em que lugar?</i> (<i>onde?</i>)	<i>Em que momento?</i> (<i>quando?</i>)	<i>De que maneira?</i> (<i>como?</i>)
(23)	neste porto	desne vinte de junho ate dezassete de agosto	esperando

Em todos esses casos, a ambigüidade estaria justamente na dupla possibilidade de sua interpretação:

- trata-se de construções locativas (similares às do Quadro B), o gerúndio figurando como uma sorte de “adjunto adverbial de modo”;
- ou
- trata-se de construções perifrásticas (similares às do Quadro E), sendo o locativo um adjunto.

Diante dessa dupla possibilidade de interpretação, e tendo a gramaticalização como ponto de vista teórico, consideram-se tais exemplos como a transição pela qual teria passado o verbo *estar* - das construções locativas e atributivas para as construções perifrásticas com o gerúndio:

<i>estar</i> + <u>locativo</u> exemplos (03-07)	>	<i>estar</i> + <u>locativo</u> + gerúndio exemplos (21-25)	>	<i>estar</i> + gerúndio+(locativo) exemplos (15-20)
<i>estar</i> + atributivo exemplos (08-14)				

Por conta disso, nos casos que constituem esse estágio de transição, não se pode afirmar categoricamente nem que *estar* funciona como flexão para o gerúndio, nem que *estar* é um verbo que subcategoriza um locativo como argumento interno.

A ocorrência dos exemplos (24) e (25) ainda no século XX - época em que *estar* é tido como flexão nas perífrases de gerúndio, de uso muito frequente - mostram, por outro lado, que as propriedades locativas de *estar* persistem mesmo nas construções perifrásticas. Com efeito, as construções abaixo (em que o locativo foi posicionado fora do conjunto *estar* + *-ndo*) não são agramaticais:

(24a) *Maria estava trabalhando na Tupi como funcionária*

(25a) *outro dia esteve me visitando aqui na minha casa uma autora teatral jovem*

Chaves et alii 1996 demonstram que, no *corpus diacrônico* organizado por Tarallo, a partir do século XVI ocorre um aumento significativo da produtividade de *estar* + *gerúndio*, computando tanto casos como os do Quadro E quanto do Quadro F:

Século	Número de ocorrências de <i>estar</i> + <i>-ndo</i>
XIII	01
XIV	02
XV	01
XVI	12
XVII	26
XVIII	41
XIX	20
XX	73
TOTAL	176

Figura 10 - *Estar* + *-ndo* através dos séculos

Por outro lado, Kewitz et alii 1994 demonstram que as construções atributivas e locativas são freqüentes desde o século XIV, as últimas predominando quase sempre:

Século	Construções Locativas (%)	Construções Atributivas (%)
XIII	100	0
XIV	71	29
XV	86	14
XVI	95,5	4,5
XVII	33	67
XVIII	73	27
XIX	60	40
XX	25	75

Figura 11 - *Estar* nas construções locativas e atributivas através dos séculos

Estudando também a diacronia do verbo *ser*, Kewitz et alii 1994 justificam o baixo número de ocorrências de *estar* em construções atributivas nos três primeiros séculos. Segundo as autoras, até por volta do século XV, *ser* e *estar* concorriam na formação de construções atributivas, havendo prevalência do verbo *ser*. A partir do século XVII, contudo, *estar* passou a ser mais usado neste tipo de construção, sem deixar de ser usado, contudo, como locativo.

Voltando à tabela da figura 09, vemos que a perífrase *estar* + *gerúndio* passou a ser mais amplamente usada a partir do mesmo século XVII. É razoável considerar, portanto, que a gramaticalização de *estar* começou a avançar por volta daquele século.

Dada a freqüência de uso de *estar* nas diferentes construções através dos séculos, é uma hipótese bastante defensável que o seu processo de auxiliarização tenha se dado da seguinte forma

verbo pleno > construções locativas > construções atributivas > construções perifrásticas

O fato de que *estar* + *-ndo* aparece já no século XIII não invalida o *continuum* acima proposto, dada a baixíssima frequência de tal construção naquele século.

Retomando toda a sequência de exemplos acima, podemos falar em termos de extensão metafórica e de implicatura, a partir da seguinte cadeia:



Figura 12 - Cadeia de extensão metafórica do verbo *estar*

Lembrando o que vimos nas considerações teóricas, as elipses acima não são entidades discretas. Sendo assim, os exemplos ditos “ambíguos” na verdade não podem causar espécie, pois são eles justamente que melhor mostram a gramaticalização em processo:

- exemplo (02) - localização espacial sem a presença de um locativo: de posição (exemplo 01) para localização espacial (03-07)
- exemplo (09): de posição para construção atributiva;
- o exemplo (12) e a questão do particípio nas construções atributivas: é provável que, “dentro” da própria elipse das construções atributivas haja um pequeno *continuum*:

estar + adjetivo > *estar* + particípios ambíguos > *estar* + particípio

Em outras palavras, antes de que *estar* fosse usado com qualquer verbo na forma nominal de particípio, é provável que tenha sido usado com aqueles cuja face adjetival fosse mais perceptível - como no caso de *inflamada*, acima exemplificado;

- os exemplos (13) e (14) - *estar* + adjunto adverbial de modo - podem ser entendidos como a transição entre as construções atributivas do tipo *estar* + adjetivo e as construções *estar* + *gerúndio*;

- como já vimos anteriormente, os exemplos do quadro F (*estar* + *locativo* + *gerúndio*) estariam na transição entre *estar* + *locativo* e *estar* + *atributivo*, de um lado, e *estar* + *gerúndio*, de outro.

Em termos de estruturas, em correspondência com a representação acima, teríamos então o seguinte:

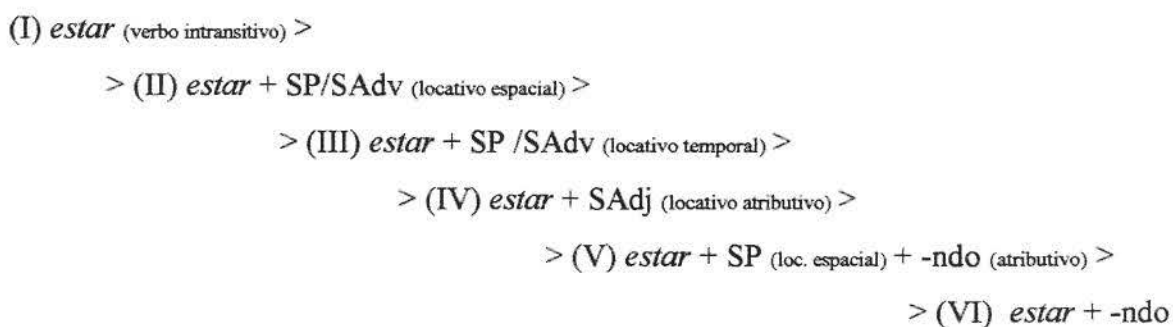


Figura 13 - *Continuum* das construções de *estar*

Retomando a escala de abstratização proposta por Heine & Claudi & Hünemeyer 1991, apreendida nas transferências conceituais de natureza metafórica, temos o seguinte:

OBJETO (I) > ESPAÇO (II) > TEMPO (III) > QUALIDADE (IV a VI)

Nesta escala, estou considerando a posição corporal “*estar de pé*” como objeto. Daí, temos o verbo *estar* localizando no espaço e depois no tempo. Finalmente, temos a categoria conceitual de “qualidade”, que abriga os três últimos estágios da figura logo 12 acima. Isto mostra, mais uma vez, que nada pode ser considerado como entidade discreta nessas representações.

Para finalizar, a análise qualitativa dos usos de *estar* não poderia ser fechada sem a inclusão dos seguintes exemplos

(26) *eu tô procurando encaminhá-la para outra coisa*

(27) *tão falando muito nisso, viu*

(28) *meu relógio tá atrapalhando nossa conversa*

(29) *eu tava pensando nisso*

(30) *a Tatá taa contando outro dia né (...) que é muito bonita a topografia da cidade*

Quadro G

Supostamente, se comparados aos exemplos do Quadro E, em que *estar* aparece em sua forma fonologicamente plena, os exemplos do quadro acima estão mais avançados no processo de gramaticalização. Isto porque, com a redução fonológica, *tá* pode tornar-se um clítico em relação ao gerúndio e, futuramente, pode vir a ser reinterpretado como “forma presa”, codificando tempo e pessoa verbais, noções gramaticais que prevaleceriam em relação ao conceito de localização (espacial, temporal ou qualitativo, conforme vimos acima).

Para que a cliticização ocorra, é necessário que *tá* perca seu acento tônico, constituindo uma única palavra fonológica com o gerúndio. Tal processo não foi estudado a fundo neste trabalho. No entanto, a fim de verificar a probabilidade de sua ocorrência, foi feita a análise espectral da vogal “a” de *tá* no exemplo (28), dentro de um contexto maior:

(28) *o meu pai fazia, ah no no primeiro colegial... e eu precisava... ter uma ah
optar por uma carreira pro — o meu relógio tá atrapalhando a nossa conversa —...
por uma carreira profissionalizante... eu achei que as coisas dali pra frente seriam
mais difíceis eu comecei o colegial...*

Foi dado a um falante nativo do português brasileiro uma folha de papel contendo o trecho acima, escrito exatamente como aqui. Numa outra folha foi escrito o mesmo trecho, mas com a forma *está*, ao contrário da anterior. O tal falante leu cada um dos trechos três vezes, alternando-os: uma vez “usando” a forma *tá*, uma vez “usando” a forma *está*. Este falante não conhecia o fenômeno em estudo e, ao ler as duas versões do trecho acima, não se deu conta de que uma das palavras era diferente.

O fato é que, em uma das três repetições, a vogal de *tá* foi pronunciada sem tonicidade, se comparada ao “a” de *está*. Os dois gráficos abaixo, fornecidos pelo aparelho *Computerized Speech Lab*, mostram a natureza do a de *tá* e de *estar* respectivamente, para o caso do exemplo (28):

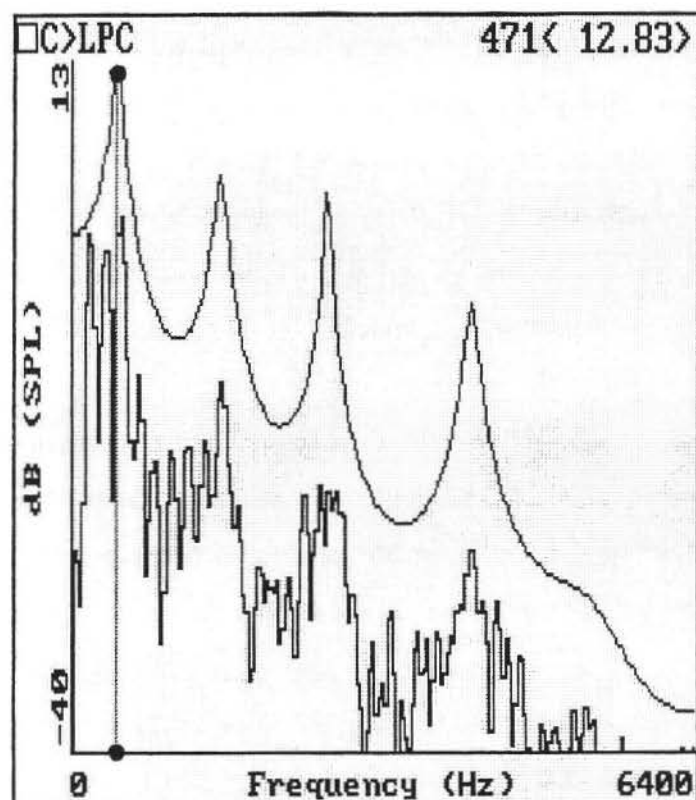


Gráfico 1 - a vogal “a” em *tá*

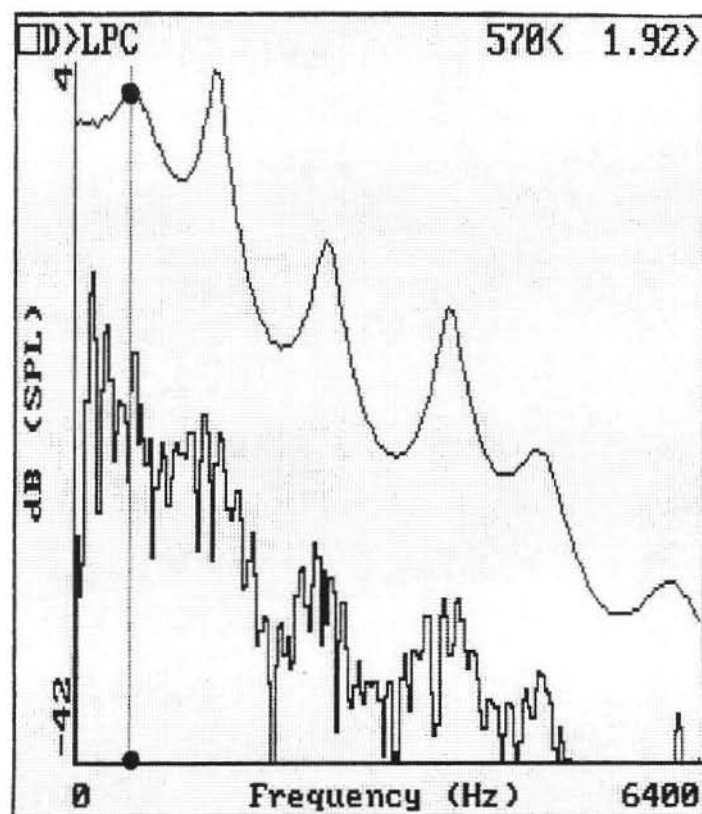


Gráfico 2 - a vogal "a" em *está*

A sigla LPC nos gráficos significa *Linear Prediction Code*. Os números impressos no canto superior direito de cada gráfico - 471 no gráfico 1 e 570 no gráfico 2, correspondem aos pontos marcados nas respectivas curvas - são valores diretamente proporcionais à abertura da mandíbula no momento da pronúncia do "a". No gráfico 1, referente à forma *tá*, o valor é menor que no gráfico 2, referente a *estar*. Isto significa que, no primeiro caso, *tá* não é tônico, de modo que pode ser entendido como um clítico em relação ao gerúndio.

É interessante dizer, ainda, que numa das repetições de leitura do trecho, o falante nem se deu conta de que vinha escrito *está* em sua folha de papel, pronunciando o auxiliar na forma reduzida.

2. Análise quantitativa dos dados

Desde a introdução a este trabalho, passando pelas considerações de natureza teórica sobre a gramaticalização e pela análise qualitativa anteriormente desenvolvidas, venho fazendo referência à análise de ordem quantitativa. Com efeito, esta é uma parte do trabalho que deve causar uma certa expectativa, o que se justifica de duas formas: primeiramente, porque desenvolver uma análise quantitativa a partir de uma amostra de dados deve permitir a confirmação (ou não) das intuições acerca do fenômeno lingüístico, delineadas sob a forma de análise qualitativa; e também porque, na grande maioria dos trabalhos sobre gramaticalização, inclusive naqueles aqui referenciados, não se realiza quantificação de dados, de modo que aqui tal análise tem o caráter de novidade.

Neste ponto, temos que voltar a algo que apenas rapidamente já foi mencionado no corpo do trabalho: a relação entre gramaticalização e mudança lingüística. Já vimos que todo processo de gramaticalização constitui um tipo de mudança, mas nem toda mudança lingüística consiste em gramaticalização. Ora, quando se fala em quantificação de dados, toca-se numa questão muito cara à Teoria da Mudança, na ótica de William Labov (Labov 1972 e 1975, por exemplo):

- em sentido amplo, considera-se que existem fatores de natureza lingüística e extralingüística que se correlacionam com a implementação da mudança;
- em sentido mais estrito, considera-se que se devem submeter os dados a análises estatísticas que permitam verificar em que medida tais fatores influenciam no desenvolvimento do processo.

Como vimos na primeira parte deste trabalho, a Gramaticalização vem sendo mais recentemente abordada a partir de um ponto de vista cognitivo, como mostram as considerações acerca das transferências conceituais de natureza metafórica envolvidas no processo. Fala-se também em inferências pragmáticas, em reinterpretações induzidas pelo contexto (conforme vimos, em linhas gerais, na abordagem do modelo da implicatura, no item 1.1.2.), o que nos permite falar em Gramaticalização como uma teoria co-textual.

Até aqui, parece coerente pensar no processo de gramaticalização como um subtipo de mudança. Na prática, entretanto, os trabalhos sobre gramaticalização não vão mais longe

na sua inserção dentro de uma teoria mais geral da mudança lingüística. A Teoria da Mudança, conforme preconizam Labov e seus seguidores, defendem que a heterogeneidade é uma característica intrínseca às línguas naturais e o ponto de partida para falar em variação e mudança lingüísticas. Todavia, essa heterogeneidade não se manifesta apenas internamente nas línguas - sob a forma de variantes em “competição”. É também notável a consideração de variáveis extralingüísticas, tais como sexo, idade, escolaridade e contexto situacional - fatores que dizem respeito aos usuários da língua e que também se correlacionam com os processos de mudança.

Nestes termos, a proposta laboviana se harmoniza com a chamada “Sociolingüística Variacionista” ou quantitativa, como vem discutido em Mollica 1992 (org.): a Teoria da Mudança tem uma notável dimensão social, visto que o pressuposto básico da heterogeneidade lingüística é diretamente relacionada aos usuários da língua e às comunidades de que os mesmos fazem parte. A Teoria sobre Gramaticalização, por outro lado, não pode ser tomada como uma teoria social. Com base no que vimos anteriormente, ela deve ser tomada antes como uma teoria associal, pois não prevê qualquer correlação entre fatores sociais e o desenvolvimento de processos de gramaticalização.

Ora, nesta parte do trabalho que chamo de “Análise Quantitativa de Dados”, diferentemente daquilo que em geral se faz nos trabalhos sobre gramaticalização, verifico justamente como a gramaticalização do verbo *estar* na perífrase de gerúndio se correlaciona não só a fatores lingüísticos, mas também a fatores sociais, extralingüísticos. Nos termos da Teoria da Variação e da Mudança, *estar* e *tá* seriam variantes de uma mesma variável, depreendendo-se uma mudança em progresso se for observada a tendência de substituição da primeira pela segunda.

É para verificar isso que se lança mão de um pacote de programas - o VARBRUL (Sankoff 1988, apud Scherre 1992) - usado no tratamento estatístico de dados lingüísticos. Devo lembrar que as primeiras aplicações desse software foram feitas a fenômenos de variação de natureza fonológica e morfológica, em que há relativamente maior segurança quanto à “mesmidade funcional” das variantes de uma variável. Quando os estudiosos se propuseram a aplicar a metodologia da Teoria da Variação a fenômenos de variação sintática, contudo, surgiu uma grande polêmica referente à validade dos resultados, polêmica

essa mais reconhecidamente representada pelas discussões entre Beatriz Lavandera e William Labov. Isto tem um certo interesse aqui pois, como veremos adiante, no estabelecimento dos grupos de fatores linguísticos, há questões sintáticas importantes no uso das variantes *estar* / *tá*. Apesar disso, não vou me aprofundar nessa questão, para o que referencio Bentivoglio 1987, que sintetiza tal polêmica.

Para tomar a gramaticalização de *estar* como um fenômeno de mudança nos termos labovianos, faz-se necessário, portanto, retomar os princípios básicos da Teoria da Mudança, repassando os problemas dos quais a mesma deveria dar conta - é o que faço, brevemente, no sub-item 2.1 a seguir, onde também estabeleço os grupos de fatores que considere na análise quantitativa que desenvolvi. Em 2.2, finalmente, apresento os resultados obtidos com tal análise.

2.1. Os grupos de fatores

De acordo com Weinreich & Labov & Herzog 1968, uma Teoria da Mudança deve ter como pressupostos básicos os seguintes pontos:

- (i) A mudança linguística não deve ser identificada como uma deriva aleatória a partir de variação inerente ao discurso. A mudança começa quando a generalização de uma alternância particular num dado subgrupo da comunidade de fala toma uma determinada direção e adquire o caráter de diferenciação ordenada;
- (ii) A associação entre estrutura e homogeneidade é uma ilusão. A estrutura linguística inclui a diferenciação ordenada de falantes e de estilos através das regras que governam a variação na comunidade de fala; a competência do falante nativo de uma língua inclui o controle de tais estruturas heterogêneas;
- (iii) Nem toda variabilidade e heterogeneidade na estrutura da língua envolve mudança; mas toda mudança envolve variabilidade e heterogeneidade;

(iv) A generalização da mudança lingüística através da estrutura da língua não é nem uniforme nem instantânea; ela envolve a co-variação de mudanças associadas no decorrer de substanciais períodos de tempo, e é refletida na difusão de isoglossas no espaço geográfico.

(v) As gramáticas nas quais a mudança lingüística ocorre são as gramáticas das comunidades de fala. Pelo fato de que as estruturas variáveis contidas na linguagem são determinadas por funções sociais, idioletos não podem fornecer a base para gramáticas internamente consistentes.

(vi) A mudança lingüística é transmitida dentro de uma comunidade como um todo; ela não é confinada a passos discretos dentro da família. Quaisquer que sejam as descontinuidades encontradas na mudança lingüística, elas serão produtos das descontinuidades dentro da comunidade, ao invés de produtos inevitáveis das lacunas geracionais entre pais e filhos.

(vii) Os fatores sociais e lingüísticos estão intimamente interrelacionados no desenvolvimento da mudança lingüística. As explicações que são confinadas a um ou outro aspecto, não importando quão bem estejam desenvolvidas, vão falhar ao tentarem dar conta das regularidades que podem ser empiricamente observadas no comportamento lingüístico.

De acordo com os mesmos autores, com base nesse pressupostos, uma forte Teoria da Mudança deveria dar conta dos seguintes problemas:

- o problema das restrições: partindo do pressuposto básico de que a heterogeneidade é inerente às línguas naturais, e que da variação lingüística pode-se passar à mudança, uma teoria forte deveria prever que mudanças vão ocorrer, que curso tais mudanças vão tomar e quais seriam os condicionamentos mais importantes para a sua implementação;
- o problema da transição: a mudança lingüística é um processo diacrônico, gradual, que geralmente mostra uma distribuição contínua através de sucessivas faixas etárias da

comunidade de falantes da língua em questão; dessa forma, entre quaisquer dois estágios A e B de uma mudança em progresso, normalmente encontra-se um estágio intermediário que define o caminho pelo qual uma estrutura A evolui para uma estrutura B. Em linhas gerais, a mudança ocorre, então, conforme o falante de uma determinada língua aprende uma forma alternante, dentre pelo menos duas que coexistam na sua época, e quando uma das formas se torna obsoleta. Isto pode ocorrer entre grupos de faixas etárias muito próximas: de acordo com os autores, as crianças, por exemplo, tendem a não preservar as características do dialeto de seus pais, mas sim as do grupo com o qual mais conviveu nos seus anos pré-adolescentes;

- o problema do encaixamento: a mudança lingüística raramente é a transição de um sistema inteiro para um outro; ao contrário, é a partir de um conjunto limitado de variantes que uma mudança pode vir a se implementar. As variantes de uma variável encontram-se conjuntamente disponíveis para o uso por parte de uma comunidade de fala; a escolha de uma variante, entretanto, vai depender:
 - de seu encaixamento na estrutura lingüística: ou seja, o uso desta ou daquela forma variante se correlaciona a aspectos internos da própria língua; considerando *estar* e *tá* como formas variantes, por exemplo, de acordo com o encaixamento na estrutura lingüística, a escolha de uma ou de outra forma pode estar correlacionada ao “tipo de elemento” com que ocorre (a forma nominal de gerúndio, um SP com função de locativo, um SAdv com função de adjunto adverbial de tempo, etc.), ou ainda à pessoa (primeira do plural, segunda do singular, etc) e/ou tempo verbais (pretérito perfeito do indicativo, presente do subjuntivo, e outros);
 - e de seu encaixamento na estrutura social: neste caso, a escolha de uma variante se correlaciona a fatores que podem ser chamados de “externos à língua”, tais como a classe social, a escolaridade, a idade e outros. O importante para a Teoria

da Mudança é considerar que as variações sociais e geográficas são parte integrante do sistema de uso lingüístico⁵.

- o problema da avaliação: a Teoria da Mudança deve estabelecer empiricamente os correlatos subjetivos das diferentes variáveis numa estrutura heterogênea. Tais correlatos subjetivos de avaliações não podem ser deduzidos do lugar das variáveis dentro da estrutura lingüística. Mais além, o nível de consciência social é uma propriedade maior da mudança lingüística que deve ser diretamente determinada. Os correlatos subjetivos da mudança são mais categoriais em natureza do que os padrões de comportamento em mudança: a sua investigação aprofunda nosso entendimento dos caminhos através dos quais a categorização discreta é imposta sobre o contínuo processo de mudança;
- o problema da atuação: do ponto de vista laboviano, atuação quer dizer o mesmo que implementação da mudança; dar conta de tal problema não é uma tarefa fácil, pois:
 - se tomarmos com rigor a proposição de que a mudança lingüística está correlacionada à mudança no comportamento social, as hipóteses preditivas não poderão ser prontamente aventadas.
 - considera-se que a mudança lingüística começa quando uma variante se difunde através de um específico subgrupo da comunidade de fala; tal variante assume, então, certa significância social - simbolizando os valores sociais associados ao grupo que a utiliza; pelo fato de que a mudança está encaixada na estrutura lingüística, ela será gradualmente generalizada para outros elementos do sistema. O mecanismo é, portanto, cíclico, de modo que é empiricamente difícil determinar

⁵ Na explanação da mudança lingüística, Weinreich & Labov & Herzog afirmam que os fatores sociais abarcam o sistema lingüístico como um todo; entretanto, deixam claro que a significância dos fatores sociais não é igualmente distribuída ao longo de todos os elementos do sistema, assim como nem todos os aspectos do sistema são igualmente marcados por variações regionais. Os autores afirmam, por exemplo, que pode haver uma mínima correlação com fatores sociais nos estágios mais incipientes e mais adiantados da mudança.

com rigor o peso real com que dado fator atua na implementação de uma mudança.

A teoria sobre gramaticalização, por sua vez, é quase que totalmente cega para os cinco problemas acima listados. De certa forma, ela dá conta do problema das restrições ao tentar, por exemplo, arrolar as propriedades de uma forma que funciona como fonte para a transferência de conceitos⁶, nos processos que chamamos anteriormente de extensão metafórica. Ou ainda, quando os teóricos afirmam que a gramaticalização de uma forma está diretamente relacionada à frequência de uso da mesma, também está se tratando do problema da restrição. Um outro exemplo é dado pela abordagem da gramaticalização como um *continuum* e um processo gradual, conforme ilustra o quadro de Lehmann 1982, citado e transcrito no início das considerações teóricas (Parte I - item 1). Em qualquer um desses casos, a teoria sobre gramaticalização pretende delinear que “mudanças” vão ocorrer e de que forma elas vão se processar, algo que se enquadra no chamado “problema das restrições” acima.

O segundo problema - o da transição - é parcialmente elaborado pela teoria da gramaticalização, quando ela enfatiza a gradatividade com que se dá o processo. Entretanto, quando Weinreich & Labov & Herzog 1968 falam em transição estão pensando mais exatamente em termos de distribuição da mudança através das diferentes faixas etárias. Ora, como sabemos, o conceito de gradualismo em gramaticalização é tomado no interior do processo linguístico: os autores não se preocupam com a transição em termos de “distribuição contínua através de sucessivas faixas etárias”.

Com respeito ao encaixamento da mudança na estrutura linguística, trata-se de um problema que a gramaticalização prevê, na medida em que a teoria admite que o processo possa ser desencadeado em contextos linguísticos muitos específicos e, portanto, reduzidos em quantidade. É com o aumento do grau de gramaticalização de uma forma que, teoricamente, ela passa a ser usada numa variedade maior de contextos. Além disso, pode-se falar em encaixamento na estrutura linguística quando se considera que o avanço dos

⁶ Lembrem-se, por exemplo, os conceitos básicos listados por Heine 1993 e referenciados no item 1.1.2.

processos de gramaticalização envolve reinterpretações que são induzidas pelos contextos em que as formas figuram.

Entretanto, Gramaticalização e Mudança Lingüística parecem distanciar-se quando o assunto é a correlação com fatores sociais. De fato, a teoria sobre gramaticalização não considera a ação de fatores sociais e, conseqüentemente, não cogita que os diferentes fatores - lingüísticos e extralingüísticos - atuam com pesos diferentes na implementação do processo. Os trabalhos sobre gramaticalização raramente incluem, portanto, uma análise estatística dos dados sobre os quais se está pesquisando.

O que me propus a fazer neste trabalho foi justamente tomar o caso da gramaticalização do verbo *estar* na perífrase de gerúndio como um caso de mudança lingüística, considerando a possibilidade de correlação do mesmo com fatores de ordem social. Procedi então à aplicação do pacote estatístico VARBRUL - um conjunto de programas desenvolvidos com base em modelos matemáticos adequados para tratar dados lingüísticos estatisticamente (Sankoff 1988, apud Scherre 1992) - a dados de *estar* + *gerúndio* e de *estar* + *X* (onde *X* = SP, SAdj ou SAdv).

Os exemplos foram retirados de cinco das seis entrevistas publicadas no *corpus* compartilhado do Projeto NURC/SP (Preti & Castilho 1986), sendo duas entrevistas com informantes de 3ª faixa etária (mais de 60 anos, em 1976), duas com informantes de 2ª faixa etária entre 30 e 40 anos, em 1976), e apenas uma entrevista com informantes de 1ª faixa etária (até 26 anos, em 1976). Optei por trabalhar com apenas uma entrevista de 1ª faixa porque o número de dados era substancialmente maior que nas outras entrevistas, além do que ambas as entrevistas de 1ª faixa apresentam um comportamento muito semelhante, com respeito aos usos do verbo *estar*.

Na verdade, desenvolvi duas análises quantitativas - de um lado, tomei como variável dependente *tá* + *-ndo* (versus) *estar* + *-ndo*, e de outro, *tá* + *x* (versus) *estar* + *X*:

VARIÁVEL DEPENDENTE	
Perífrase de Gerúndio	Outras Construções
<i>tá</i> + <i>-ndo</i>	<i>tá</i> + <i>X</i>
<i>estar</i> + <i>-ndo</i>	<i>estar</i> + <i>X</i>

A partir da comparação das duas análises, a intenção seria responder basicamente as seguintes perguntas:

- a redução fonológica de *estar* - pré-requisito para sua cliticização à forma de gerúndio, aqui entendida como um avanço no seu processo de gramaticalização - ocorre com mais frequência nas perífrases de gerúndio, ou ocorre indistintamente tanto nas perífrases quanto nas outras construções?
- *estar* + *-ndo* e *tá* + *-ndo* são variantes de uma mesma variável dependente ou há realmente uma mudança em progresso?
- os fatores extralingüísticos se correlacionam, de fato, a tal caso de gramaticalização?
- o conjunto *tá* + *-ndo* é mais coeso do que o conjunto *estar* + *-ndo*, permitindo-nos falar em redução fonológica como aumento do grau de gramaticalização de *estar*?

Com o intuito de responder a essas perguntas e inspirado nas considerações de Weinreich & Labov & Herzog 1968 sobre a Teoria da Mudança, estabeleci os seguintes grupos de fatores sociais:

GRUPOS DE FATORES (Variáveis Independentes)	
FATORES SOCIAIS	
1. <u>Falante</u> (indicado pelo número da entrevista)	2. <u>Sexo</u>
1 L1-396	m masculino
2 L2-396	f feminino
3 L1-333	3. <u>Idade</u>
4 L2-333	1 primeira faixa etária (até 26 anos, em 1976)
5 L1-255	2 segunda faixa etária (entre 30 e 40 anos, em 76)
6 L2-255	3 terceira faixa etária (mais de 60 anos, em 76)
7 L1-360	4. <u>Grau de atenção à fala</u>
8 L2-360	+ maior atenção
9 L1-343	- menor atenção
0 L2-343	

O primeiro grupo de fatores, através do qual identifico os falantes (ou informantes) nas entrevistas, justifica-se no sentido de que era necessário verificar se, dentro daquele *corpus*, haveria comportamento idiossincrático da parte de um deles, ou seja, preferêncica pela forma *estar* ou pela forma *tá*, nos conjuntos com gerúndio e com X.

Quanto ao segundo grupo de fatores sociais - o do sexo - pode parecer estranho que o uso de *estar* ou de *tá* esteja correlacionado ao fato de que o falante seja homem ou mulher. Como afirma Chambers 1995:105, há pouquíssimas diferenças biológicas entre homens e mulheres que se refletem na língua por eles falada (segundo este autor, documentos lingüísticos mostram que as mulheres são menos passíveis de sofrer de gagueira e de afasia). Desse modo, tal grupo de fatores não pode ser tomado a partir de um ponto de vista biológico, mas sim social. Segundo Labov 1972:243, *“na fala cuidada, as mulheres usam menos formas estigmatizadas do que os homens, e são mais sensíveis às formas prestigiadas do que eles*. Labov 1990 (apud Chambers 1995:103) também afirma que *“dentro de uma estratificação sociolingüística estável, os homens usam formas não-padrão mais freqüentemente que as mulheres”*. De acordo com Chambers 1995:112, este é um dos padrões mais comumente aceitos e citados na literatura sociolingüística.

Passando para o terceiro grupo de fatores, temos o dado social da idade. Tal grupo de fatores pôde ser estabelecido graças ao fato de que as entrevistas do Projeto NURC foram gravadas com falantes de diferentes faixas etárias, conforme indicado na lista de grupos de fatores acima. Sendo que a primeira faixa etária vai até os 26 anos e que a terceira faixa etária começa a partir dos 60, tem-se um considerável intervalo temporal para a análise dos dados, ainda que todas as entrevistas tenham sido gravadas em 1976.

De acordo com uma posição considerada “clássica”, segundo Naro 1992, *“o processo de aquisição da linguagem se encerra mais ou menos no início da puberdade, e (...) a partir desse momento a língua do indivíduo fica essencialmente estável.”* Nesse sentido, um falante de 60 anos em 1976 fala basicamente do mesmo modo com que falava em 1931 (admitindo que a puberdade começa aos 15 anos).

De tal ponto de vista, com um falante de 60 anos em 1976 tem-se acesso à língua que, de certa forma, falava-se há 45 anos atrás. Dessa forma, as mudanças observadas a partir da comparação entre a fala de tal indivíduo com a fala de outros, mais novos, permite falar em *mudança em tempo aparente*. Naro 1992 lembra, por outro lado, que tal hipótese clássica se encontra cada vez mais enfraquecida, enfatizando que “*cada caso de variação em tempo aparente terá que ser investigado numa tentativa de determinar seu status histórico verdadeiro*”.

De qualquer modo, justifica-se a proposta do grupo de fatores “idade” para o caso da variação *estar* / *tá* porque, graças à noção de mudança em tempo aparente, a análise dos dados a partir da idade dos falantes deverá permitir verificar se *tá* e *estar* encontram-se em variação estável ou se realmente constituem uma mudança em progresso.

O grupo chamado de grau de atenção à fala, finalmente, vem incluído entre os fatores sociais porque diz respeito à monitoração que o falante faria em relação à própria fala, no momento em que a entrevista estivesse sendo gravada. Supostamente, quanto maior fosse tal monitoração, menor seria a tendência em usar a forma reduzida *tá*. Ao contrário, se o grau de atenção fosse maior, o falante tenderia a usar a forma não-reduzida *estar*.

Assim como os outros fatores sociais, este também não é comumente considerado nos trabalhos sobre gramaticalização. Acredita-se, entretanto, que numa “conversa mais solta”, o uso das formas linguísticas seria mais espontâneo, favorecendo o uso das formas não-padrão, aquelas mesmas formas que um falante de alto grau de escolaridade julgaria “erradas” (o que seria o caso da forma reduzida *tá*, por exemplo).

Ainda que os diálogos entre dois informantes do NURC/SP sejam considerados “informais” (se comparados aos outros dois tipos de entrevistas - “Diálogos entre Informante e Documentador” e “Elocuções Formais”), na verdade, os falantes não são igualmente “informais” no uso de sua própria língua, assim como não mantêm o mesmo grau de informalidade (entenda-se *espontaneidade*) durante todo o tempo da entrevista.

Dessa forma, na codificação dos dados, nem sempre foi fácil “medir” o grau de atenção que o falante dava à sua própria fala. Seguindo Labov 1972:86-94, considerei o grau de atenção à fala maior ou menor de acordo com a presença das seguintes

características (no trecho da entrevista em que aparecessem as variantes das variáveis dependentes aqui estabelecidas):

Características da fala	Grau de atenção				
	Maior		Menor		
apagamento de concordância verbal e nominal	-	-	-	-	+
interrupção momentânea da entrevista	-	-	-	+	+
presença de comentários à parte	-	-	+	+	+
apagamento da consoante /r/ no final dos verbos no infinitivo	-	+	+	+	+

A ordem das características da fala foi estabelecida conforme a importância que se deu a elas. Considerando que todos os informantes das entrevistas que constituíram o *corpus* eram de nível universitário, filhos de pais universitários, tomei o apagamento da concordância verbal ou nominal como o indicador mais importante de que houve, digamos, “descuido” com a fala. Neste sentido, como mostra a primeira linha da tabela acima, caso ocorresse “apagamento” da concordância verbal ou nominal, considerei mínimo o grau de atenção prestado à fala.

Em seguida, aditi como segundo aspecto mais importante a interrupção momentânea da entrevista. Seguindo Labov 1972, considerei eventuais interrupções no curso da entrevista como desencadeadores de maior espontaneidade, e portanto de menor monitoração da fala. Por “interrupção” eventual entenda-se:

- a entrada, no ambiente onde estivesse sendo gravada a entrevista, de uma pessoa que não estivesse participando da mesma;
- alguém que entrasse para servir alguma bebida ou petiscos;
- um telefone que tocasse e quebrasse o curso da entrevista;
- um barulho externo que chamasse à atenção, etc

O terceiro aspecto da lista é o da presença de comentários à parte, desenvolvidos pelo(s) próprio(s) informante(s), que fizessem com que a conversa se desviasse do assunto inicialmente proposto pelo documentador, ainda que momentaneamente. Ouvindo os diálogos entre dois informantes, percebi que, na maioria das vezes, comentários que fugiam

ao tópico corrente da entrevista caracterizavam-se por maior espontaneidade, desencadeando uma maior informalidade na sequência da conversa, caso o tom anterior ao comentário já não fosse informal.

Por fim, considere também o apagamento da consoante /-r/ no final dos verbos no infinitivo. Trata-se de fenômeno muito freqüente no PB falado, de modo que considere a pronúncia do /r/ (*falar*, em vez de *falá*; *fazer*, em vez de *fazê*, etc) como indicador de “cuidado” (eu diria até “esmero”) no uso da língua.

Conforme pode ser visto na tabela acima, os casos extremos aparecem em sombreado. A atenção à fala seria indiscutivelmente maior quando nenhuma das características listadas fossem notadas no diálogo. Ao contrário, quando todas fossem realizadas, a atenção seria considerada mínima.

Estabelecidos os fatores sociais, passamos então aos fatores lingüísticos. A pergunta aqui é basicamente formulada da mesma forma que no caso anterior: que fatores de natureza lingüística estariam correlacionados ao uso da forma fonologicamente reduzida *ta* ou da fonologicamente plena *estar*?

FATORES LINGÜÍSTICOS⁷

5. Tempo e modo do verbo *estar*

- p presente do indicativo
- y imperfeito do indicativo
- f infinitivo
- b imperfeito do subjuntivo

6. Pessoa Verbal

- 1 primeira pessoa do singular
- 2 segunda e terceira pessoas do singular
- 3 primeira pessoa do plural
- 4 segunda e terceira do plural
- 5 impessoal

7. Tipo de palavra fonológica à esquerda de *estar*

- z pausa
- n monossílabo átono (*eu*, etc)
- m monossílabo tônico (*eu*, *cê*, *el*)
- o oxitona
- a paroxitona (*inclusive êli*, *éla*)

8. Complexidade Morfofonológica de X

- baixa (uma palavra, com qualquer acento)
- * média (duas palavras)
- # alta (locuções ou sintagmas extensos)

9. Número de sílabas do gerúndio

- 2 duas sílabas
- 3 três sílabas
- 4 quatro sílabas
- 5 mais de sílabas

10. Tipo de material interferente entre *estar* e *-ndo*

- sem material interferente
- c clítico
- * tudo/ ainda/ realmente/ também/ mais, etc
- z pausa

11. Tipo semântico do verbo no gerúndio

- r realização [+din] [+con] [+tel]
- a atividade [+din] [+con] [-tel]
- m mudança [+din] [-con] [+tel]
- d dinamismo [+din] [-con] [-tel]

2.2. Os resultados obtidos

Com a quantificação dos dados, obtiveram-se os seguintes resultados, no que diz respeito aos fatores sociais (lembre-se que X = SP, SAdj ou SAdv):

⁷ Os grupos de fatores 5, 6 e 7 dizem respeito tanto à variação *estar* + X versus *tá* + X quanto à variação *estar* + *-ndo* versus *tá* + *-ndo*. O grupo 8, entretanto, refere-se apenas ao primeiro caso, enquanto que os grupos 9, 10 e 11 se referem apenas ao último.

Grupo 1 Falante	Estar + Gerúndio			Estar + X		
	nº de casos	tá (%)	estar (%)	nº de casos	tá (%)	estar (%)
1	7	86	14	16	50	50
2	2	50	50	9	33	67
3	34	6	94	13	8	92
4	8	38	62	5	0	100
5	14	7	93	7	0	100
6	9	44	56	9	78	22
7	6	50	50	7	29	71
8	19	84	16	21	62	38
9	37	97	3	18	78	22
0	48	98	2	12	100	0
Grupo 2 Sexo	Nº de casos	tá (%)	estar (%)	Nº de casos	tá (%)	estar (%)
masc	66	70	30	50	58	42
fem	118	62	38	67	46	54
Grupo 3 Idade	Nº de casos	tá (%)	estar (%)	Nº de casos	tá (%)	estar (%)
1ª faixa	85	98	02	30	87	13
2ª faixa	48	50	50	44	50	50
3ª faixa	51	24	76	43	28	72
Grupo 4 Atenção	Nº de casos	tá (%)	estar (%)	Nº de casos	tá (%)	estar (%)
Maior	62	24	76	70	12	88
Menor	122	85	15	47	73	27
Totais	184	69	31	117	51	49

Tabela 1 - Resultados dos Grupos de Fatores Sociais

Os resultados referentes ao Grupo 1 mostram paralelismo entre os usos da forma reduzida *tá* nas construções com *gerúndio* e e nas construções com X. Em princípio, isto poderia ser um contra-argumento à morfologização de *estar*, na medida em que não seria o gerúndio uma das explicações para a redução fonológica, ou seja, não seria o fato de *estar* figurar numa construção com um verbo em forma nominal, servindo de flexão número-pessoal para o mesmo, o que estaria correlacionado à sua redução fonológica.

Há, entretanto, dois argumentos a favor:

- o primeiro é o de que não é necessariamente o gerúndio que deve explicar a redução fonológica de *estar*, uma vez que, conforme vimos nas considerações teóricas, podem ser fatores prosódicos aqueles que a motivam; ora, tais fatores prosódicos não têm que atuar, necessariamente, apenas quando o verbo *estar* venha acompanhado de gerúndio;
- por outro lado, se observamos a última linha da tabela, é com o gerúndio que ocorre uma maior porcentagem de *tá* (69% dos casos), ao passo que há um equilíbrio nos usos das variantes *tá* e *estar* (51% versus 49%, respectivamente) nas construções com *X*; isto permite dizer, então, que é provável que os falantes usem *tá* nas construções com *X* por analogia com as construções de *tá* + *-ndo*.

Pensando em cada falante isoladamente, a tabela acima mostra que a maioria deles segue essa tendência geral em usar a forma reduzida nas construções com gerúndio. Há apenas três exceções: os falantes 3, 6, e 0, destacados nas linhas sombreadas da tabela. Para que isso fique mais claro, reorganizei os falantes na tabela abaixo:

Falante	Faixa Etária	Forma Reduzida TÁ		Forma Plena ESTAR	
		+ gerúndio (%)	+ X (%)	+ gerúndio (%)	+ X (%)
1	3ª	86	50	14	50
2	3ª	50	33	50	67
4	3ª	38	00	62	100
5	2ª	07	00	93	100
7	2ª	50	29	50	71
8	2ª	84	62	16	38
9	1ª	97	78	3	22
3	3ª	06	08	94	92
6	2ª	44	78	56	22
0	1ª	98	100	02	00

Tabela 2 - Os falantes e a escolha de *tá/estar*

Para os sete primeiros falantes dessa tabela, a porcentagem de uso de *tá* nas construções com *gerúndio* é maior do que nas construções com um *X* qualquer;

inversamente, portanto, a porcentagem de *estar* nas construções com gerúndio é menor do que nas construções com *X*.

No entanto, o falante nº3, membro da terceira faixa etária (mais de 60 anos) utiliza quase que categoricamente *estar*, qualquer que seja o tipo de construção. Além disso, as porcentagens de uso de *tá* nas construções com *gerúndio* e nas construções com *X* são praticamente iguais (6% e 8%, respectivamente).

Já o falante nº6 (segunda faixa etária) utiliza mais a forma *tá* nas construções com *X* do que propriamente nas construções com gerúndio. Por fim, quanto ao falante nº0 (primeira faixa etária - até 26 anos), ele utiliza quase que apenas a forma *tá*.

Dentro do conjunto dos falantes, portanto, os três acima mostram um comportamento idiossincrático, na medida em que não seguem a tendência da maioria. Os resultados referentes a eles, todavia, são interessantes: o falante de terceira faixa etária - mais velho - utiliza mais a forma *estar*, enquanto que o falante da primeira faixa - mais novo - utiliza mais a forma *tá*. De certa forma, pode-se já aqui vislumbrar mudança em progresso em tempo aparente. Maiores detalhes serão dados quando da apresentação dos resultados sobre o grupo de fatores da faixa etária.

Quanto ao segundo grupo de fatores, há uma pequena diferença na distribuição da forma reduzida *tá*: enquanto que os homens as utilizam em 70% dos casos, as mulheres o fazem em 62% dos casos. Tal diferença quantitativa é mínima, mas mostra-se consistente, comprovando a hipótese sociolingüística da variação correlacionada ao sexo. Ora, dentro de um grupo com características mais ou menos estáveis - mesmo grau de escolaridade, numa mesma época, em entrevistas documentadas da mesma forma - as mulheres mostraram relativamente mais resistência ao uso da forma inovadora.

Na seqüência, temos o grupo das faixas etárias, que se mostrou de grande poder explanatório. Os gráficos abaixo mostram isso mais claramente:

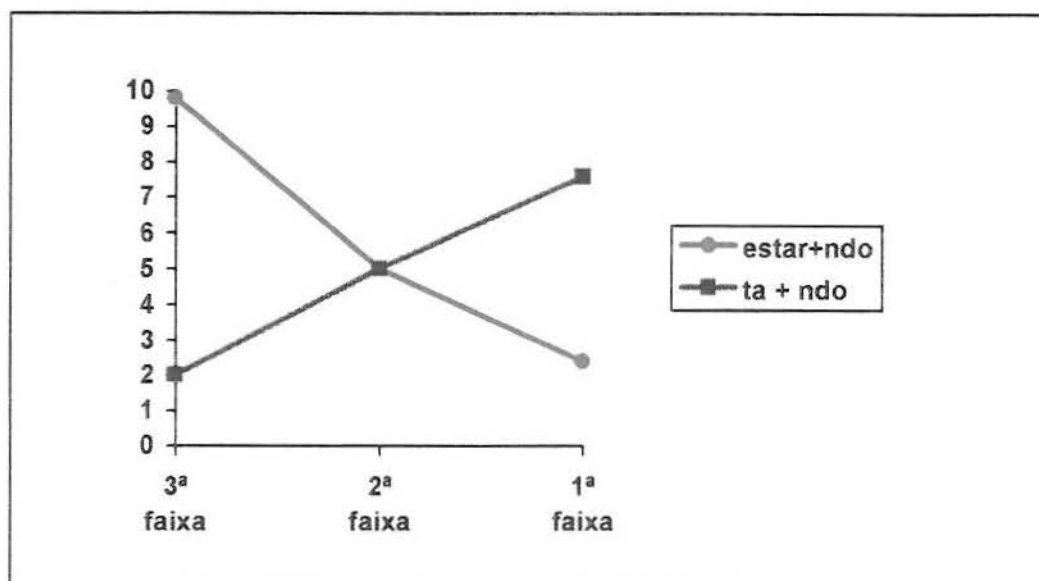


Gráfico 3

Variação *estar/tá* de acordo com a faixa etária do falante

De acordo com o gráfico, os falantes de terceira faixa etária (mais de 60 anos) usam menos a forma *tá*, enquanto que na primeira faixa etária (menos de 26 anos) há mais casos de *tá* e menos de *estar*. Entre as duas, a segunda faixa etária apresenta equilíbrio no uso das variantes. Esses resultados mostram que, não obstante as reservas feitas por Naro 1992 em relação ao conceito de mudança em tempo aparente, conforme vimos acima, é cada vez maior a preferência pela forma *tá*, com o passar do tempo.

A estes resultados, devo acrescentar que ouvi também outras três entrevistas, parte da “Amostra Censo da Variação Lingüística no Estado do Rio de Janeiro”, dentro do Programa de Estudos sobre o Uso da Língua (PEUL-UFRJ). Um dos informantes era do sexo masculino e outro do feminino, estando os dois na faixa dos trinta anos na época das gravações - fins da década de 80. A terceira informante era uma menina de 10 anos de idade.

Nos três casos, não apareceu nenhuma vez *estar* em sua forma fonologicamente plena. É verdade que, nas três entrevistas, as conversas foram bastante descontraídas (o que permite dizer que ocorreu pouco monitoramento da fala por parte dos falantes), além do que o português falado não era o culto, como no caso do NURC/SP, mas sim popular. Mesmo

assim, parece que na passagem dos 70 para os 80 houve um aumento nos usos da forma *tá* no português falado.

No intuito de comprovar a hipótese, ouvi outras três gravações de diálogos entre dois informantes realizadas em 1995. Numa delas, os informantes eram estudantes universitários do sexo feminino, na faixa dos 20 anos. Na outra, os informantes eram também estudantes universitários, na faixa dos 20 anos, mas do sexo masculino. Em ambas as entrevistas o português falado era o culto, e a variante *estar* não foi usada nenhuma vez.

Finalmente, na terceira entrevista, um diálogo entre dois informantes do sexo masculino, sendo um paulista e um paraense, ambos na mesma faixa etária acima, o único caso de *estar*, em detrimento de *tá*, foi o seguinte, articulado pelo paraense:

(31) *o que eu estava falando pra ti é o seguinte...*

Interessante notar que foi o mesmo informante que articulou o exemplo seguinte, já referenciado anteriormente:

(32) *eu acho muito legal tu ta andando na rua*

Ora, na mesma entrevista, esse falante obedece a conjugação do “tu” em qualquer outra situação que não a de *estar* como auxiliar - *tu vistes, tu fizestes, tu te lembrastes, etc.* Isso pode significar o início do processo de morfologização da forma *tá*, ainda que seja uma tendência cancelar a marcação morfológica de 2ª pessoa do singular, de acordo com Paredes da Silva.

Finalmente, com respeito ao último grupo de fatores sociais - o do grau de atenção à fala - os resultados também foram bastante interessantes:

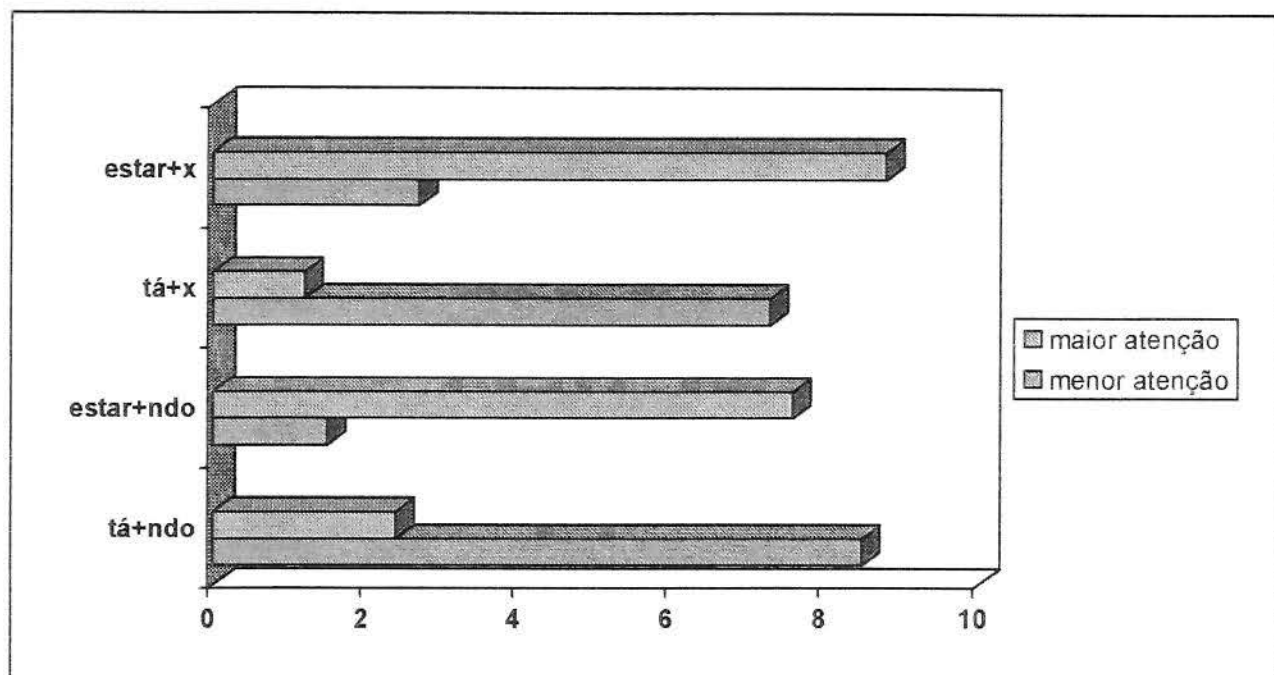


Gráfico 4

Uso das variantes *estar/tá* conforme o grau de atenção à fala

De acordo com o gráfico acima, a forma reduzida foi preferida, tanto nas construções com gerúndio quanto nas construções com X, nos momentos em que os falantes prestavam menor atenção à sua fala (nos termos que defini anteriormente em 2.1). Ao contrário, naqueles momentos em que os falantes poderiam monitorar sua própria fala, a forma canônica - *estar* - foi a preferida.

Dado que estes dois últimos grupos de fatores - o do grau de atenção à fala e o da faixa etária - mostraram-se tão elucidativos, efetuei o cruzamento entre os dados fornecidos por cada um deles, obtendo os seguintes resultados:

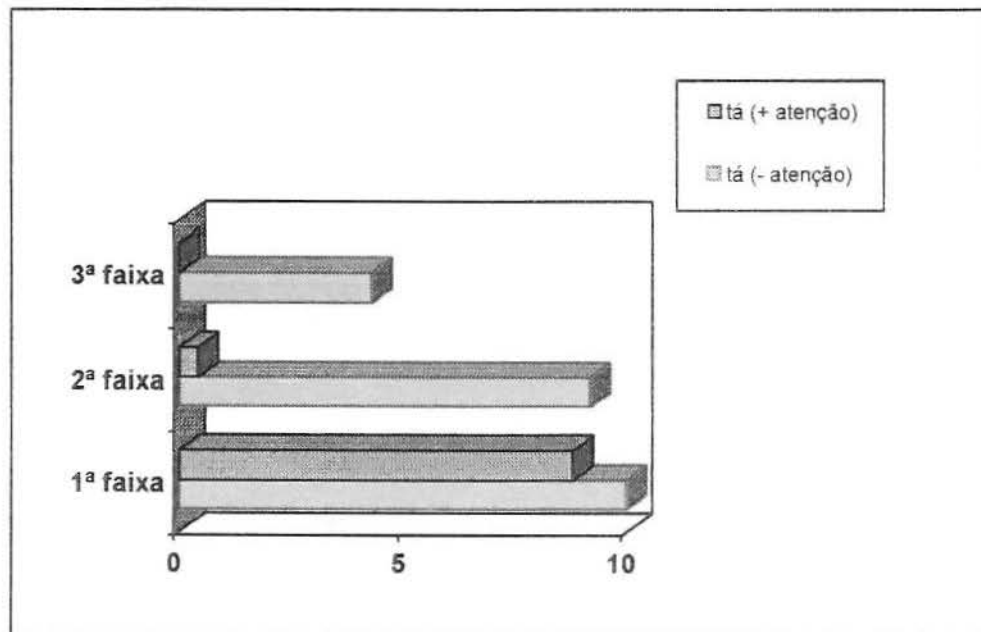


Gráfico 5

Uso de *tá* nas diferentes faixas etárias
de acordo com o grau de atenção à fala

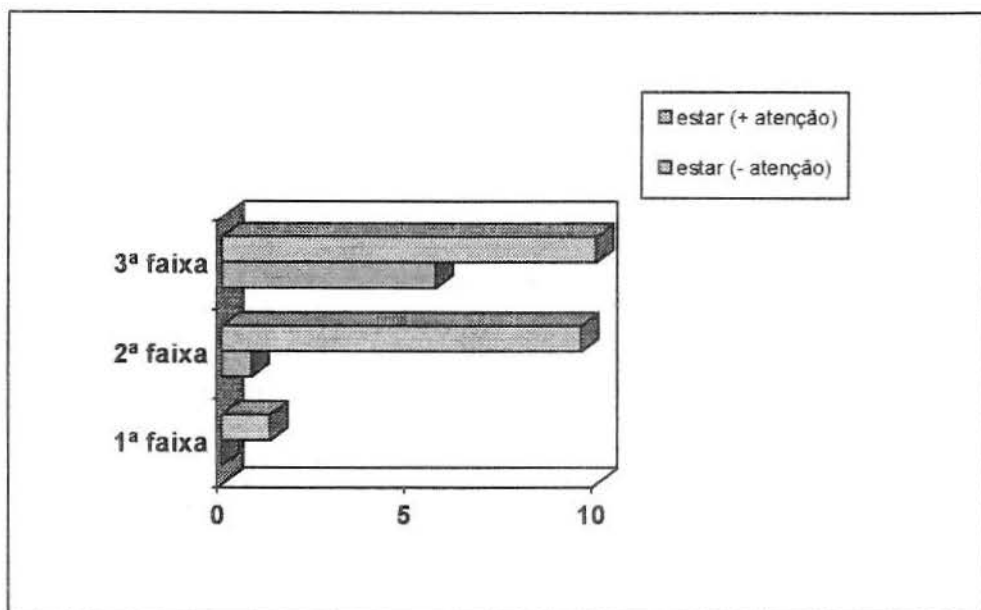


Gráfico 6

Uso de *estar* nas diferentes faixas etárias
de acordo com o grau de atenção à fala

Estes gráficos mostram que, na primeira faixa etária (falantes mais novos), nem mesmo o monitoramento da fala leva o falante a optar pela forma fonologicamente plena *estar*. No gráfico B, temos que o mínimo de uso de *estar* é feito em condições de maior monitoramento da fala. Em outras palavras, o uso da forma *tá* pelos falantes mais jovens já não depende da atenção que o mesmo presta à sua fala. Na segunda faixa etária, mais uma vez, observa-se equilíbrio: comparando o gráfico A com o B, vemos que mesmo quando o falante monitora sua própria fala, ele tende a empregar a variante *estar*; por outro lado, quando não há tal monitoramento, a tendência é de empregar *tá*. Na terceira faixa etária (falantes mais velhos), por fim, e ao contrário do que ocorreu na primeira, mesmo quando não se presta atenção ao uso da própria língua, é a forma *estar* que predomina.

Veja-se ainda que, quando a atenção dada à fala é mínima, os falantes de primeira faixa etária não usam *estar*. De maneira exatamente oposta, e portanto complementar, quando a atenção dada à fala é máxima, os falantes de terceira faixa etária não empregam a forma *tá*. No meio termo, estão os falantes da segunda faixa etária, os únicos que usam ambas as variantes, qualquer que seja o grau de atenção à fala.

Passando agora aos grupos de fatores de natureza lingüística, trato da questão do tempo em que está conjugado o auxiliar. Vou me deter no presente e no pretérito imperfeito do indicativo, apesar de ter levado em conta outros tempos verbais que apareceram nas entrevistas, por representarem estes a grande maioria dos casos. No caso de *estar* + *gerúndio*, de acordo com o esperado, em ambos os tempos houve predominância da forma reduzida (*tá* e *tava*). O que causou espécie foi a distribuição das formas no caso de *estar* + *X*:

Estar + -ndo			
Grupo	Nº de casos	tá (%)	estar (%)
pres.	154	65	35
imperf.	24	83	17
Estar + X			
Grupo	Nº de casos	tá (%)	estar (%)
pres.	80	60	40
imperf.	24	33	67

Tabela 3 - O tempo do auxiliar

Nestas construções, de acordo com os valores sombreados, o pretérito imperfeito inibiria o uso da forma reduzida *tá* nas construções com X, como nos exemplos abaixo:

(33) *nesse dia, eu estava aqui na minha sala e sintonizei no canal do programa*

(34) *estava muito interessante esta sua crônica*

As duas tabelas acima vêm a confirmar, então, que a redução de *estar* para *tá* é mais freqüente nas perífrases de gerúndio. Com efeito, não foram poucos os casos como

(35) *eu **taa** pensando... será que isso é... sempre... desde que começou a haver máquina*

(36) *A Tatá **taa** contando outro dia né (...) que é muito bonita a topografia da cidade*

em que a articulação da lábio-dental sonora [v] do morfema de pretérito imperfeito foi substituída ou por uma aspirada gutural [r] ou por um simples alongamento da vogal temática [a], em construções com gerúndio.

Uma última consideração a fazer em relação ao tempo do auxiliar é a de que aqueles tempos que são os menos utilizados, dentro das entrevistas do NURC/SP, inibem o uso da forma reduzida, confirmando a importância da freqüência de uso para a gramaticalização:

(37) *e preciso que ninguém mais esteja indo entende*

(38) *ela ta estará cumprindo um papel que hoje e ate desconhecido*

(39) *eu estaria fazendo uma auto-critica*

(40) *hoje eu estive vendo um livro...editado pelo...Instituto Roberto*

O exemplo (38) é particularmente interessante, pois o falante ensaiou o uso da forma *tá*, no presente, mas logo em seguida “mudou de opinião”, usando a forma plena no futuro. Nas entrevistas que foram gravadas em 1994, contudo, apareceram os seguintes casos:

(41) *ele acaba usando o que ele tiver usando numa lingua especifica*

(42) *se eu pudesse eu taria viajando*

Esses exemplos, mais recentes que os anteriores, não só reforçam a hipótese da morfologização em progresso de *tá*, mas também permitem inferir que a redução fonológica em outros tempos ocorre por analogia ao que se passa no presente e no imperfeito, muito mais freqüentes. Não há dúvida de que, no presente e no imperfeito, a forma reduzida de *estar* é muito mais passível de perder seu acento, se comparada a formas como *taria* e *tiver* - este último confundindo-se, aliás, com o futuro do subjuntivo de *ter*.

Quanto ao sexto grupo de fatores, parece não haver correlação entre a pessoa do discurso codificada pelo auxiliar e a seleção das formas alternantes *tá* e *estar*. Vale dizer, contudo, que tanto nas construções com gerúndio quanto nas construções com *X*, a forma plena foi a mais usada na primeira pessoa do plural (mais de 75% dos casos), como nos exemplos

(43)... *mas por enquanto nós não estamos desenvolvendo nada , nós estamos só ...
assimilando Know-how*

(44) L1: *como é que estamos no tempo?*

L2: *tá ... tá ótimo*

A segunda pessoa do plural é, justamente, a menos utilizada, o que empiricamente corrobora o ponto de vista teórico de que a gramaticalização de uma forma está diretamente relacionada à sua frequência de uso. Na época das gravações do NURC/SP, portanto, pode-se dizer que os falantes ainda pareciam resistir um pouco em articular frases como a seguinte:

(45) *então nós **tamo pensando...** **tamo pensando** que não; oficialmente já foi*

O grupo de fatores de número 7 - relativo ao tipo de palavra fonológica que vem à esquerda de *estar* - foi proposto com o intuito de se verificar se a mesma estava correlacionada ao uso das variantes *tá* / *estar*. A tabela abaixo contém os valores obtidos com a quantificação dos dados:

Caráter da palavra fonológica à esquerda de <i>estar</i>	Construções com -ndo			Construções com X		
	Nº de casos	tá (%)	estar (%)	Nº de casos	tá (%)	estar (%)
pausa	25	60	40	16	56	44
monossílabos tônicos	68	65	35	34	47	53
monossílabos átonos	20	75	25	14	57	43
oxítonas	20	60	40	11	45	55
paroxítonas	51	65	35	42	60	40

Tabela 4 - A variação *tá* / *estar* de acordo com a palavra fonológica à esquerda do auxiliar

Para ilustrar cada um dos casos, nas construções com gerúndio, temos os exemplos:

(46) PAUSA⁸

acho que o grande problema das pesquisas é esse ... realmente:: ... tá existindo vamos dizer que:: que no país hoje em dia está ... entrando assim muito muitos bens de capital ...

(47) MONOSSÍLABOS TÔNICOS

acho que ela não está cumprindo aquele serviço

(48) MONOSSÍLABOS ÁTONOS⁹

eu estou dizendo que... seria fácil pifar a solução deles

(49) OXÍTONAS

então tá acumulando

e é desses doutores que nós estamos precisando ...

(51) PAROXÍTONAS

o italiano tá morando aqui

Relativamente às construções com *X*, o caráter fonológico da palavra à esquerda do auxiliar não teve poder explanatório sobre os usos das variáveis dependentes: como mostra a tabela, há equilíbrio no uso das mesmas.

No caso das construções com gerúndio, os resultados foram um pouco diferentes, mas tampouco foram explanatórios. Em qualquer um dos casos, a forma *tá* foi privilegiada, sobretudo nos casos assinalados.

⁸ Nas normas para transcrição de dados do NURC/SP, a pausa é simbolizada pelas reticências [...] (Preti & Castilho orgs 1986 e 1987)

⁹ Os pronomes de primeira pessoa do singular foram considerados palavras fonológicas sem acento tônico naqueles casos em que eles vieram isolados antes do auxiliar (ou seja, sem nenhuma palavra anterior a ele na mesma frase fonológica), sendo pronunciado sem ênfase pelo falante.

Com os grupos de fatores 8 (Complexidade Morfofonológica de X) e 9 (número de sílabas do gerúndio) pretendeu-se verificar se a redução fonológica de *estar* apresentaria alguma correlação com aspectos prosódicos. No primeiro caso, os resultados mostraram que não há correlação entre a escolha da variante e a complexidade morfofonológica de X:

Estar + X			
Complexidade de X	Nº de casos	tá (%)	estar (%)
baixa (uma palavra)	55	51	49
média (duas palavras)	29	52	48
alta (locuções ou sintagmas extensos)	33	52	48

Tabela 5

A tabela mostra que há praticamente total equilíbrio no uso das formas variantes, seja X um elemento simples, um elemento de média complexidade morfofonológica ou um elemento de maior complexidade. Seguem abaixo os respectivos exemplos:

(52) ... e *está aqui* o remédio

(53) a Marília Medalha me disse que era... que *tava muito interessante* esse seu essa sua crônica

(54) Dinah *está com boas esperanças* de levar agora A Muralha

Para o caso de *estar* nas construções com *gerúndio*, como já foi dito, o grupo de fatores em questão é o do número de sílabas do verbo principal. De acordo com o que vimos no item 1.2, a hipótese é de que a forma *tá* teria sido usada inicialmente naqueles casos em que a redução se implementasse no sentido de otimizar a alternância rítmica na pronúncia da perífrase. Para tanto, a redução seria desencadeada nos casos em que o

gerúndio tivesse número ímpar de sílabas, de modo que seria então pronunciada uma sequência de dois troqueus básicos, como em

(55) *ela ta falando que não.*

(56) *pronto sarou era isso que tava inteferindo... era um espirito não sei das quantas...*

A quantificação dos dados confirmou a hipótese, pois:

Número de sílabas do gerúndio	Total de Casos	tá (%)	estar(%)
2	22	59	41
3	81	77	23
4	46	55	45
5	30	82	18

Tabela 6

Nesta tabela, fica clara a predominância de *tá* nos casos em que o gerúndio apresentava número ímpar de sílabas. Percebe-se equilíbrio no uso das formas variantes quando o gerúndio tem 2 e 4 sílabas.

Isso leva a crer que os aspectos rítmicos do PB também são fatores que se correlacionam ao avanço no processo de gramaticalização de *estar*. Retomando a questão da complexidade de *X*, o fato de que há equilíbrio na distribuição das variantes *tá* e *estar* permite dizer que a variação é livre para o caso daquelas construções (algo que é coerente com os valores totais apresentados na primeira tabela: 51% de casos de *tá + X* contra 49% de *estar + X*). Nesse sentido, também é razoável afirmar, mais uma vez, que o uso da forma reduzida naquelas construções foi desencadeada por um processo de analogia com os casos de *tá + gerúndio*, o provável contexto onde a forma reduzida teria sido primeiramente empregada.

Para a gramaticalização enquanto teoria, é interessante verificar a atuação de fatores de natureza prosódica. Como já comentamos, a maioria dos trabalhos correlaciona o avanço dos processos de gramaticalização com aspectos estruturais ou semânticos, não se atendo à possibilidade de correlação com aspectos fonológicos. No caso de *tá* + *-ndo*, a importância de tais aspectos está no fato de que eles facilitam e fortalecem a continuidade da reinterpretação “de verbo auxiliar para clítico”.

Finalmente, propus o décimo grupo de fatores envolvido na variação dos usos de *estar* na perífrase de gerúndio, pensando na questão dos “elementos interferentes”, ou seja, aqueles elementos sintáticos que figurariam entre o auxiliar e o gerúndio.

Lembrando os princípios de Lehmann 1982a, vistos na primeira parte desse trabalho, na passagem de uma gramaticalização fraca para uma gramaticalização forte, um dos processos que se observa é o da coalescência. Nesse sentido, a ocorrência de elementos entre o auxiliar e a forma nominal seria um empecilho à morfologização do primeiro. Esperava-se, portanto, que nos casos de *tá* + *-ndo*, haveria proporcionalmente menos casos elementos interferentes:

(57) *ele ainda não tá... trabalhando bem né?* (PAUSA INTERFERINDO)

(58) *o país tá se construindo aos poucos* (CLÍTICO INTERFERINDO)

(59) *eu tô sempre correndo, sempre falando, tudo depressa* (ADVÉRBIO INTERFERINDO)

Dentro do *corpus*, foram poucos casos como os exemplificados acima. Por essa razão, coloquei-os todos dentro de uma mesma categoria - “presença de material interferente” - em oposição a “ausência de material interferente”, como mostra o gráfico abaixo:

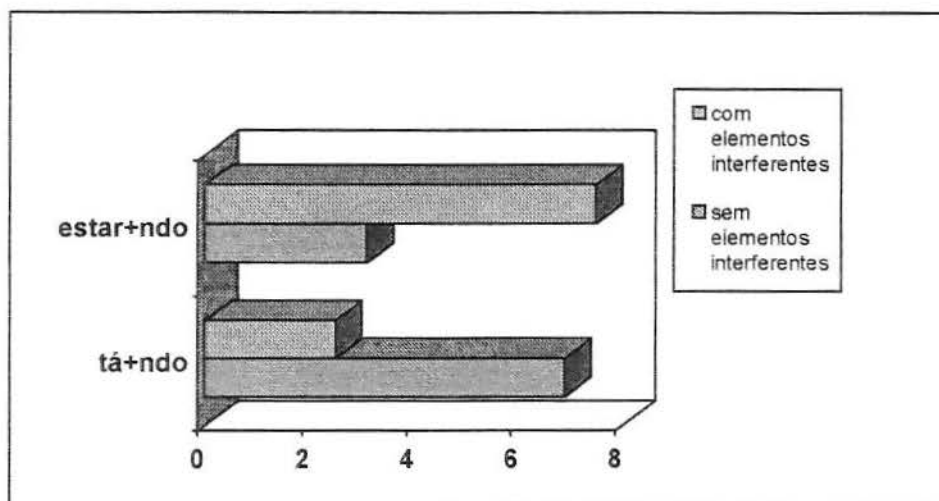


Gráfico 7 - A inserção de elementos interferentes entre o auxiliar e o gerúndio e a escolha de *tá/estar*

Como se vê, o uso da forma *tá* se correlaciona com a ausência de elementos interferentes entre o auxiliar e o gerúndio. A inserção de tais elementos é, portanto, uma medida da coalescência entre *tá* e *-ndo*: se não houvesse nenhum caso de interveniência entre *tá* e a forma nominal em questão, o grau de dependência sintagmática entre ambos seria máximo. Conforme o gráfico, isso parece ser uma tendência, dado que ocorrem menos elementos interferentes entre *tá* + *-ndo* do que entre *estar* + *-ndo*.

Além disso, observou-se que nos casos em que o auxiliar e o gerúndio fossem separados por um locativo, nas entrevistas do NURC/SP que foram examinadas, a variante de *estar* empregada era a não-reduzida:

(60) *agora outro dia esteve aqui na minha casa me visitando uma autora*

(61) *eu estava na tupi trabalhando como funcionária*

(62) *nós estamos aqui dando um depoimento sobre esse aspecto da linguagem*

Exemplos também interessantes são aqueles em que o verbo *estar* figura sem nenhum elemento expresso à sua direita. Nesses, não houve nenhuma ocorrência da forma reduzida:

(63) *oficialmente não tá encerrado, mas de fato está*

(64) *ai eu procurei, escolhi...foi escolhida a que eles estão*

(65) *eles estão... eles têm uma esperança*

(66) *daqui há uns vinte anos... como a gente estaria?*

O último grupo de fatores estabelecido foi o do tipo semântico do verbo no gerúndio. Considerando que o verbo *estar* foi inicialmente empregado com o sentido de *posição corporal* e em construções locativas (conforme vimos na análise qualitativa), a hipótese relacionada a tal grupo de fatores é a de que haveria algum tipo semântico de verbo que se correlacionasse ao uso de uma ou de outra forma variante.

Para isso, utilizei a classificação semântica de verbos proposta por Dik 1978 (ver grupo de fatores nº11, na lista de grupos de fatores acima). Os resultados obtidos, contudo, mostram quase que exato equilíbrio na escolha das formas variantes, para os diferentes tipos semânticos de verbos no gerúndio, de modo que tal grupo de fatores não se mostrou minimamente explanatório.

Conclusão

A análise quantitativa dos dados permitiu verificar que os falantes do PB tendem a usar a forma reduzida *tá* nas construções com gerúndio. Ao contrário do que ocorreu nas construções com tal forma nominal, houve equilíbrio no uso de *estar* e de *tá* nas outras construções.

Apesar das restrições relativas à concepção de mudança em tempo aparente, conforme vimos, podemos considerar que a substituição de *estar* por *tá*, nas perífrases de gerúndio, é uma mudança em progresso. Os dados colhidos das entrevistas gravadas mais recentemente mostram que determinados fatores - como o grau de atenção à fala - já não operam: o falante utiliza quase que categoricamente a forma reduzida.

Quanto à morfologização de *tá* - cliticização seguida de afixação ao gerúndio - ainda é cedo para fazer qualquer afirmação. Os dados mostram, contudo, que é uma tendência considerável a de que tal processo continue. Uma evidência importante disso é o fato de que o uso da forma reduzida comprovadamente inibe a inserção de elementos, como locativos e adjuntos adverbiais, entre o auxiliar e o gerúndio. Ou seja, comparada à forma plena *estar*, a forma reduzida é mais coalescente ao gerúndio.

Em se tratando das mudanças de natureza semântica e da redução do tamanho do item em processo de gramaticalização - uma das questões teóricas mais importantes na literatura sobre o assunto - este trabalho mostrou que não se pode estabelecer relação direta de causa e consequência entre as mesmas. Não nego que a redução *estar* > *tá* esteja correlacionada ao processo de extensão metafórica pela qual o verbo *estar* teria passado em seu processo de auxiliarização, como vimos na análise qualitativa. No entanto, quero enfatizar a ação de aspectos prosódicos, que também se correlacionam com a redução.

E por falar em “fatores”, este trabalho mostrou a importância da consideração de aspectos sociais no processo de gramaticalização. Tratando-se de um caso de mudança, tal processo se instaura não apenas num contexto lingüístico definido, mas também num contexto situacional específico, como puderam mostrar os grupos de fatores sociais aqui propostos. Teoricamente, portanto, não convém negligenciar a dimensão social do processo

de gramaticalização, se se quer um máximo de rigor possível ao descrever como o mesmo se implementa.

Um último comentário a fazer diz respeito à natureza metafórica do processo de gramaticalização. O caso do verbo *estar* na perífrase de gerúndio, conforme vimos na análise qualitativa, é uma confirmação empírica de que a gramaticalização não deve ser explicada meramente como um processo de perda de propriedades intensionais. Não significa que não se possa falar em termos de “perda”. Entretanto, deve-se fazê-lo com a consciência de que *perda* é um termo que se pode usar na descrição de uma instância específica dos processos de gramaticalização; o termo não explica, entretanto, o processo em toda sua complexidade.

ABSTRACT

The aim of this study is to check how the grammaticalization of the verb *estar* gets started and goes on in spoken Brazilian Portuguese. The question that has motivated this study is: would the reduced form *tá* replace the original one - *estar* - in periphrastic constructions with gerund? In order to try to answer that question, I have analyzed 5 interviews (two persons talking about a given subject) provided by NURC/SP (Preti & Castilho 1986 and 1987). Based mainly on the framework proposed by Heine & Claudi & Hünemeyer 1991, and trying to look at grammaticalization processes as a kind of linguistic change (according to Labov's theory) I have submitted the data to the statistical software VARBRUL. From the quantitative analysis, it has been possible to see that older native speakers employ rather the plain form *estar*, whereas the younger ones don't even happen to use it. This fact characterizes a linguistic change in progress. As for the morphologization of *tá*, although interfering elements (such as locatives and adverbial adjuncts) are not likely to be inserted between the auxiliary and the gerund, we're not yet allowed to state that the reduced form *tá* has been cliticized or agglutinated to the gerund.

Referências Bibliográficas

- ARAÚJO, M. F. R. - 1998 - *A alternância de /ey/~e/ no Português Falado na cidade de Caxias (MA)*, versão preliminar da dissertação de Mestrado a ser apresentada na UNICAMP em 1999.
- BECHARA, E. - 1992 - *Moderna Gramática Portuguesa*. São Paulo, Companhia Editora Nacional.
- BENTIVOGLIO, P. - 1987 - A variação nos estudos sintáticos. In: *Estudos Lingüísticos*. XIV Anais do GEL, Campinas.
- BISOL, L. - 1996 - Constituintes Prosódicos, in *Introdução a Estudos de Fonologia do Português Brasileiro*. Porto Alegre, EDIPUCRS.
- GALVES, C.C. & ABAURRE, M. B. - 1997 - *Rhythmic differences between Brazilian and European Portuguese: evidence for rhythm implementation at the PF interface*
- BRAGA, M. L. - 1994 - “A dimensão dos constituintes no português do Brasil”, *Tempo Brasileiro* 117:17-26.
- CASTILHO, A. T. de 1995 - “A Gramaticalização” . *Estudos Linguísticos e Literários* 19:25-64, UFBa.
- CASTILHO, A. T. de & MORAES DE CASTILHO, C. M. - 1994 - “O Aspecto Verbal no Português”
- CHAVES, T. R. A. P. et alii - 1996 - Diacronia das perífrases verbais de estar + -ndo / -r / -do. Comunicação apresentada no XLIII Seminário do GEL, em *Estudos Lingüísticos*

- DIETRICH, W. - 1973 - *El Aspecto Verbal Perifrastico en las Lenguas Romanicas*. Trad. Espanhola. Madrid, Gredos, 1983.
- DIK, S. C. - 1978 - *Functional Grammar*. Dordrecht - Holland / Cinnaminson - EUA, Foris Publications.
- GIVÓN, T. - 1979 - *On understanding Grammar*. New York, San Francisco, London, Academic Press.
- HAEGEMAN, L. - 1991 - *Introduction to Government and Binding Theory*. Oxford, Basil Blackwell
- HOPPER, P. J. 1991. On some principles of grammaticization, em E.C.TRAUGOTT and B.HEINE Eds. 1991, pp. 17-36.
- HOPPER, P.J. & TRAUGOTT, E.C. - 1993 - *Grammaticalization*. Cambrige, CUP.
- HEINE, C., CLAUDI, U. & HÜNNEMEYER, F. - 1991- *Grammaticalization - a conceptual framework*. The University of Chicago Press.
- KEWITZ, V. et alii - 1994 - Estudo diacronico de ser e estar em construções não perifrásticas. Trabalho apresentado ao professor Atliba de Castilho, no curso de Sintaxe Diacrônica do Portugues. USP.
- LABOV, W. - 1972 - *Sociolinguistic Patterns*. Philadelphia, University of Pennsylvania Press.
- LABOV, W. - 1972 - The Isolation of Linguistic Styles. In: *Sociolinguistic Patterns*. Philadelphia, University of Pennsylvania Press.

- _____ - 1975 - *Language in the inner city*. Philadelphia, University of Pennsylvania Press.
- LEHMANN, C. 1982a. Thoughts on Grammaticalization. A programmatic sketch. Köln, Arbeiten des Kölner Universalien-Projekts, vol. I.
- _____ 1982b. Grammaticalization: synchronic variation and diachronic change, *Lingua e Stile* 20: 303-318.
- LOBATO, L.M.P. - 1975 - Os verbos auxiliares em português contemporâneo. Critérios de Auxiliaridade, in *Análises Lingüísticas*, Petrópolis, Vozes.
- MENDES, R. B. - 1994a - Quantificação e Aspecto Verbal. O caso da perífrase estar + gerúndio. Comunicação apresentada no XLII Seminário do Grupo de Estudos Lingüísticos do Estado de São Paulo, em *Estudos Lingüísticos* 34.
- _____ - 1994b - Estar + -ndo e a expressão do Aspecto no português falado. Comunicação apresentada na I Jornada de Iniciação Científica na XLVI Reunião Anual da SBPC.
- MOURA NEVES, M. H. de - 1997 - *A Gramática Funcional*. São Paulo, Martins Fontes
- MOLLIKA, M.C. (org.) - 1992 - *Introdução à Sociolingüística Variacionista*. Cadernos Didáticos UFRJ.
- NESPOR, M. & VOGEL, I. - 1986 - *Prosodic Phonology*. Dordrecht - Holland/Riverton - USA. Foris Publications.

- PRETI, D. & CASTILHO, A. T. de (orgs.) 1986 e 1987 - A Linguagem Falada Culta na Cidade de São Paulo. Volumes II. São Paulo, TAQueiroz.
- SCHERRE, M. M. P. - 1992 - *Introdução ao pacote VARBRUL para microcomputadores*. Universidade Federal do Rio de Janeiro
- SWEETSER, E. - 1990 - *From Etymology to Pragmatics - Metaphorical and cultural aspects of semantic structure*. Cambridge, Cambridge University Press.
- RAPOSO, E. - 1992 - *Teoria da Gramática: a faculdade da linguagem*. Lisboa, Caminho.
- TARALLO, F. - 1991 - *Corpus Diacrônico do Português*, 8 vols., mimeo.
- TRAUGOTT, E.K. and HEINE, B. Eds. - 1991 - *Approaches to Grammaticalization*. Amsterdam/Philadelphia, 2 vols.
- WEINREICH, U. & LABOV, W. & HERZOG M. L. - 1968 - Empirical Foundations for a Theory of Language Change . In: LEHMANN, W. P. & YAKOV, M. (Eds.) *Directions for Historical Linguistics*. Austin and London, University of Texas Press.