

## Como citar este artigo:

TRAVAGLIA , Luiz Carlos. A gramaticalização do verbo começar. In TRAVAGLIA, Luiz Carlos et alii (org.). **Lingüística: caminhos e descaminhos em perspectiva**. Uberlândia, EDUFU, 2006. p. 514 - 529.

# A GRAMATICALIZAÇÃO DO VERBO *COMEÇAR*

Luiz Carlos Travaglia

**RESUMO:** Neste artigo trata-se da gramaticalização do verbo *começar* no Português do Brasil. Consta-se que este verbo tem hoje um uso mais gramatical do que lexical. Ele é utilizado principalmente como marcador dos aspectos começado e inceptivo (71,98 % das ocorrências), mas também é utilizado como ordenador textual. Este último uso parece estar surgindo no século XXI em textos mais formais.

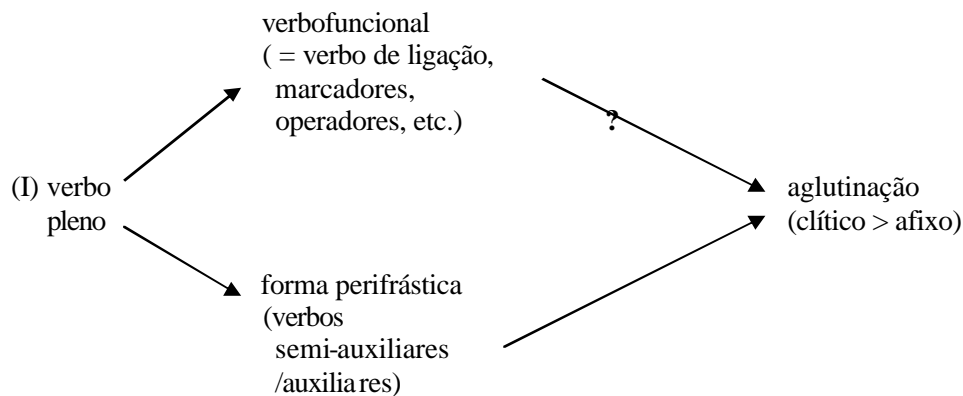
**RESUMÉ:** Dans cet article on parle de la grammaticalisation du verbe “*começar*” (commencer) en Portugais du Brésil. On constate que ce verbe a aujourd’hui une utilisation plus grammaticale que lexicale. Il est utilisé surtout comme un marqueur des aspects commencé et inceptif (71,98 % des occurrences), mais il est aussi utilisé comme un ordonnateur textuel. Il semble que cette dernière fonction est apparue au XXI<sup>ème</sup> siècle, dans des textes plus formels.

## 1. Introdução

### 1.1. Preliminares

A gramaticalização do verbo *começar*, detectando seus valores gramaticais e o seu grau de gramaticalização em cada valor gramatical, é o objeto deste estudo<sup>1</sup>.

Tomamos a gramaticalização em seu sentido estrito, entendida como a passagem de um item lexical ou gramatical a gramatical ou mais gramatical. Consideramos, conforme o que foi proposto por nós em Travaglia (2002 e 2003a), que os verbos, em sua gramaticalização, seguem geralmente uma das trajetórias em (I), em que o ponto de interrogação sugere a necessidade de se pesquisar se o verbo de ligação passa para os estágios seguintes.



Travaglia (2002: 138)

As duas cadeias de (I) podem ser melhor visualizadas como as configuramos em Travaglia (2003a: 98), nas cadeias de (II) e (III), em que os parênteses indicam estágio não obrigatório no processo de gramaticalização e o ponto de interrogação, como já dissemos, indica a necessidade de se pesquisar se o verbo de ligação passa para os estágios seguintes.

<sup>1</sup> - O estudo da gramaticalização do verbo *começar* foi realizado em conjunto com o de outros verbos, constituindo uma cadeia interligada de gramaticalização: *começar* / *passar* - *continuar* - *acabar*, *terminar* / *deixar*.

(II) verbo pleno > (forma perifrástica: verbos semi-auxiliares / auxiliares) > verbo de ligação ou verbo funcional > ? aglutinação (clítico > afixo) ?

(III) verbo pleno > forma perifrástica (verbos semi-auxiliares /auxiliares) > aglutinação (clítico > afixo)

Metodologicamente convém ressaltar que, para quantificação e correlação de fatores, usamos o programa GOLDVARB 2001 da Sociolinguística no seu aplicativo “make cell”. A seguir apresentamos outros detalhes que permitem compreender melhor a realização do estudo e seus resultados.

## 1.2. Grupos de fatores para verificar a gramaticalização

Para a realização da pesquisa foram usados dois grandes grupos de fatores, visando basicamente à verificação: a) do grau de gramaticalização; b) de se a gramaticalização tem origem ou ocorre de preferência em algum contexto linguístico ou extralinguístico; e, c) da datação dos processos de gramaticalização. Portanto, embora os grupos de fatores propostos tratem todos do processo de gramaticalização, visam observar aspectos diversos relacionados ao processo. Assim temos:

i) Fatores que verificam o **grau de gramaticalização**:

- Todos os grupos de fatores do conjunto I (Integração) de grupos de fatores;
- Os seguintes grupos de fatores do conjunto II (Gramaticalização em geral) de grupos de fatores: a) 2 . Tipo de sujeito; b) 3 Forma do sujeito; c) 4. Forma verbal; d) 5. Tempo verbal – Categoria; e) 6. Modalidades; f) 8. Aspecto – duração; g) 9. Aspecto – realização; h)10. Aspecto – desenvolvimento; i) 11. Aspecto – completamento; j) 12. Pessoa;

A variedade de tipos de sujeito (grupo 2) revela que não funcionam mais restrições de seleção do verbo pleno com relação ao sujeito, portanto o verbo estaria mais gramaticalizado. Para Heine (1993, p. 58-66) seria estágio 3 de gramaticalização do verbo. A não variedade de formas do sujeito (grupo 3) indicaria a obrigatoriedade ou a impossibilidade de ocorrência em certos contextos e quando isto acontece o verbo está mais gramaticalizado pela perda de características sintáticas. (Cf. estágio 3 de gramaticalização de Heine, 1993). Os fatores dos grupos 4, 5, 6, 8, 9, 10, 11 e 12 revelam a descategorização ou recategorização do item, pois, se muitas formas não puderem aparecer, isto revela uma neutralização de marcas morfológicas, própria do estágio 4 de gramaticalização, segundo Heine (1993).

ii) Fatores que verificam se a **gramaticalização tem origem ou ocorre de preferência em algum contexto linguístico ou extralinguístico**: a) 7. Tipo de texto fundamental; b) 13. Modalidade de língua; c) 14. Tipo de amostra; d) 16. Sexo do produtor, do conjunto II de fatores;

iii) Fatores que **“datam” os processos de gramaticalização**: 15. Época e 17. Idade do produtor.

### 1.2.1. Grupos de fatores I – Integração

Quanto mais um verbo de uma perífrase está gramaticalizado, mais está vinculado, ligado, aderido, unido ao seu verbo principal. Quanto mais forte se torna este vínculo, podemos dizer que há maior integração entre os dois verbos até o ponto de, no último estágio (cf. estágio 7, proposto por Heine, 1993: 58-66) se tornarem uma única palavra. É a ação e o resultado do princípio da coalescência. A integração então é o critério geral para verificar o grau de ligação/vinculação entre os verbos de uma perífrase. A integração pode ser

constatada por uma série de fatores. Hipotetizamos que os fatores abaixo revelariam maior ou menor integração e, portanto, maior ou menor gramaticalização, conforme as indicações para cada grupo.

Na análise optamos por considerar o verbo principal elíptico como inexistente, o que no grupo de fatores de integração faz com que os fatores 2 a 7 não se apliquem, sendo por isso marcados com o sinal “/” na codificação dos dados. Para a integração temos os seguintes grupos de fatores:

### Grupo 1- Valores do verbo *começar*:

- *Valores lexicais*: **1**<sup>2</sup> - Dar começo a; principiar; iniciar; encetar: a) dar início/ começo/ princípio a um processo, ação (que aparece como objeto): X dar início a Y (ex. 1); b) Um fato, fenômeno (que aparece como sujeito) ter início/ começo / princípio (exs. 2 e 3)<sup>3</sup>; **2- a)** Fazer a primeira tentativa ou ter sua primeira experiência em algum campo (ex, 4); **b)** Dar mostras de sua aptidão (poucos casos) (ex.5); **3-** Marca a primeira situação de uma série ou o início de uma série de situações (exs. 6 e 7)<sup>4</sup>.
- *Valores gramaticais*: **3-** Auxiliar marcando aspecto inceptivo (ex. 8)<sup>5</sup>; **4-** Ordenador textual (ex. 9 e 10)<sup>6</sup>.

- (1) O Presidente **começou** a visita pela ala infantil do hospital.
- (2) A competição **começaria** oito e meia (Tend. Rafael 14 anos)
- (3) A busca de Artemis **tinha começado** há dois anos, quando começou seu interesse por navegar na Internet. (texto 2)
- (4) Acho que o fato de eu **ter começado** como ator me fez ver o quanto outros envolvidos na produção podem contribuir para o resultado final. (Texto 83, Folha de São Paulo, homem/depoimento, narrativo)
- (5) Quem **começa** assim há de ir longe.
- (6) Meu filho é muito bonito. Sempre foi. Esse foi o **azar** da vida dele. **Começou** com o negócio de modelo, depois disse que ia ser artista de novela, que esses caras todos ganham uma fortuna só pra ficar beijando aquelas atrizes bonitas. (Texto 5)
- (7) **Começou** a leitura **pedindo** desculpas aos ouvintes. (Houaiss)
- (8) Quando eles (veículos 1.0) **começaram a rodar** por aqui, eram tão rústicos no acabamento e modestos no desempenho que receberam o apelido de pé-de-boi, lembra José Eduardo Favaretto. (texto 4)
- (9) O professor começou a conferência apresentando três hipóteses, discutiu cada uma e terminou dizendo que pelos argumentos apresentados a última era a mais plausível.
- (10) Privilégio de Amá? Como é que se diz? É, a mãe... **começô** assim: ela era novinha, aí [o]... o rapaz que ia sê padre eve relações com ela, ela ficô grávida,..... (Tend. Simone 27 anos).

**Grupo 2- Intercalação de material entre o auxiliar e o principal**, considerando as seguintes possibilidades do material de intercalação: **T** - Trecho (Frases, oração, SN com encaixada, etc.); **N** - Negação; **R** - Pronome; **S** - Sintagma; **P** - Preposição da locução; **A** - Adverbial; **0** - Nada intercalado (zero); + – Dois ou mais tipos de

<sup>2</sup> - Os números e símbolos que antecedem os fatores de cada grupo é o símbolo usado na codificação para o programa GOLDVARB-2001, utilizado para as quantificações dos dados e a observação da atuação relativa de cada fator.

<sup>3</sup> - Com este valor é comum a construção: nome de situação (construção, briga, festa, etc.) + começar + circunstância (tempo, modo, etc.). O nome de situação pode aparecer como sujeito ou objeto.

<sup>4</sup> - a) Forma: Começar + gerúndio; Começar + pôr (em?, de?) + infinitivo (Galicismo); b) Geralmente o que segue o verbo começar é uma oração subordinada adverbial modal

<sup>5</sup> - Quanto à indicação do aspecto inceptivo pelo verbo *começar* na locução “Começar + a + infinitivo, Travaglia (1981, p. 248 a 251) diz que se marca o aspecto inceptivo nos seguintes casos: a) com os presentes do indicativo e do subjuntivo (com valor de presente) e os pretéritos imperfeito do indicativo e do subjuntivo (com um presentificador) do indicativo e do subjuntivo; b) com todas as formas verbais (exceto os pretéritos perfeito e mais que perfeito), quando a locução aparece conjugada com a locução de estar + gerúndio: ESTAR + COMEÇANDO + INFINITIVO; c) para a situação narrada (Cf. Travaglia, 1981, p.71-73) que é expressa pelo infinitivo, nos futuros do presente, do pretérito e do subjuntivo, nas formas com sentido futuro do presente e do pretérito imperfeito do indicativo (raramente), e do subjuntivo (mais frequentemente) e nos pretéritos perfeito e mais que perfeito do indicativo + elemento que precise o momento de início da situação.

<sup>6</sup> - Forma: Quase sempre “começar + verbo de elocução (gerúndio ou por + infinitivo)”

material (Ex.: Prep. + Pronome objeto / Adverbial); **i** – Interjeição; **M**- Marcador conversacional; **C**- Conjunção; /- Não se aplica.

A possibilidade de intercalação de qualquer material revela menor integração entre os dois verbos do que a impossibilidade de intercalação, pois esta impossibilidade revela um vínculo maior. Quando entre os dois verbos há uma preposição e esta é responsável ou distinguidora de um valor, uso ou função não se deve considerá-la como uma intercalação, pois neste caso a preposição faria parte do “auxiliar”. A intercalação de trechos ou de dois ou mais tipos de material revelaria o menor grau de vinculação / integração e o zero (0), o maior nível, por revelar a impossibilidade de intercalação.

**Grupo 3- Status da forma nominal (ou seu equivalente): 1** - Argumento do auxiliar (sujeito, objeto, adjunto adverbial, adjunto adnominal, etc); **2**- - Não argumento do auxiliar (uma só forma); /- Não se aplica.

Se a forma nominal for argumento do verbo em gramaticalização isto revela que ele ainda tem status de verbo pleno capaz de atribuir papéis argumentais. Em caso contrário ele estará formando uma **unidade semântica** com o principal, o que é característica de verbos em estágio 3 (unidade semântica) e 4 (não poder mais ter argumentos representados por nomes) de gramaticalização .

**Grupo 4- Sujeito dos 2 verbos (auxiliar e forma nominal): M** – Mesmo sujeito; **D** - Sujeitos diferentes; /- Não se aplica.

Os dois verbos terem o mesmo sujeito indica maior grau de integração. Para Heine (1993), os verbos terem o mesmo sujeito, mas isto não ser obrigatório, é característico do estágio 2 e os dois verbos terem obrigatoriamente o mesmo sujeito é característico do estágio 3 de gramaticalização.

**Grupo 5- Tipo de subordinada que a forma nominal representa: 1**- Justaposta; **2** - Desenvolvida adverbial; **3**- Desenvolvida integrante; **4**- Desenvolvida adjetiva; **5**- Reduzida; **6**- Dessentencializada; **7**- Nominalizada; / - Não se aplica.

Este fator é, em certa medida, correlacionado com o fator 3, já que os tipos de oração de 1 a 5 podem ser argumentos do verbo “auxiliar” e as do tipo 6 não podem ser. No caso da justaposta, não há qualquer vinculação entre as duas formas nem mesmo indireta por meio de um conectivo. No caso das orações desenvolvidas, já há um vínculo, mesmo que indireto por meio de conectivo, entre duas formas verbais. No caso da adverbial, o vínculo é menos rígido (da natureza dos adjuntos adverbiais), havendo mais possibilidade de mobilidade de uma em relação à outra; no caso das integrantes, o vínculo é maior (da natureza dos complementos) e a mobilidade é menor e, no caso das adjetivas, o vínculo é maior ainda devido à natureza de encaixada deste tipo de oração e a mobilidade é nula. No caso das reduzidas, o vínculo é direto entre as duas formas verbais e o grau de integração é alto. No caso do fator 6 (forma nominal dessentencializada), a forma nominal do verbo principal não pode ser vista mais como uma oração e não representa argumento da forma em gramaticalização, o que remete ao fator 3. A forma nominalizada não ocorreu no *corpus* analisado, mas nos parece agora, no fim da pesquisa, bastante problemática, porque se o verbo principal aparecer como argumento nominal do verbo em gramaticalização isto representaria um retorno ao verbo pleno?

No exemplo (11) pode-se perceber a diferença entre a reduzida (11a) e a dessentencializada (11b).

- (11) a- Caetano **começou** o show **cantando** uma música de João Gilberto.  
b- No show, Caetano **começou a cantar** uma música de João Gilberto.

**Grupo 6- Pausa entre auxiliar e principal:** [ - Pausa; # - Não pausa;

A possibilidade de uma pausa entre as duas formas verbais revela menor grau de integração, pois mostra que não constituem ainda uma **unidade semântica** (cf. característica do estágio 3, conforme Heine, 1993, p. 58-66) e **sintática**. A impossibilidade da pausa revela maior integração.

**Grupo 7- Mobilidade do que vem depois do verbo em gramaticalização:** **A** – Pode passar para antes; **B** – Não pode passar para antes.

A impossibilidade de movimento tem a ver com o princípio da **fixação** e revela uma construção que vai se tornando cada vez mais vinculada. Heine (1993, p. 58-66) diz que o fato de o verbo em gramaticalização começar a poder ser usado somente em uma posição em relação ao seu complemento (verbo principal da perífrase) é uma característica do estágio 5.

### 1.2.2. Grupos de fatores II – Gramaticalização em geral

Para a gramaticalização em geral temos os seguintes grupos de fatores:

**Grupo 1- Valor do verbo** (cf. grupo 1 de fatores da seção 1.2.1)

**Grupo 2- Tipo de sujeito:** **T-** texto; **H-** humano; **A-** animal ou outro ser animado; **O-** objeto; **S-** situação; **P-** período de tempo; **X-** sem sujeito

O tipo de sujeito O (Objeto) inclui objetos concretos (exemplos 12), abstratos (exemplos 13) e instituições (exemplos 14). Não foi considerado o sujeito elíptico porque sempre é possível recuperar o tipo de sujeito inclusive nos casos de sujeito indeterminado. O sujeito tipo “*texto*” incluiria coisas como a carta, o livro, o filme, a conferência, etc.

(12) a- Mas o playground **começava a se povoar** e os gritos chegavam ali em cima, ..... (Texto 123 Conto 1997)

b- O cenário de Manhattan, então, **passa a ostentar** duas chaminés. (Texto 49, Carta Capital, homem, narrativo)

(13) a- Num mundo em que a grandeza dos povos e a relevância de seus desejos e necessidades **passaram a se medir** só pelo seu poder aquisitivo, ..... (Texto 54, Carta Capital, homem, narrativo)

b- ..... Saul ergueu a taça e brindou à nossa amizade que nunca vai **terminar**. (Texto 105, 1982, escrito, culto, homem, narrativo)

(14) a- A substituição não seria para alguma autarquia **passar a produzir** itens banais, nem seria calcada em reserva de mercado..... (Texto 41, homem, dissertativo)

b- ..... só que agora, depois que a INTERBRAS **acabou**, essa coisa me incomoda tanto que eu quero mais que o circo peque fogo, ..... (Tendência, Eucy, 55 anos, narrativo)

Para Heine (1993, p. 58-66), a variedade de tipos de referentes dos SNs sujeitos é característica do estágio 3. Assim, se o verbo em gramaticalização admitir vários tipos de referente quando em locução com outro verbo ele já estaria no terceiro estágio de gramaticalização, o dos quase-auxiliares.

**Grupo 3- Forma do sujeito:** #- nome; \*- pronome; \$- sintagma nominal; &- oração; = - elipse (zero); X- sem sujeito.

Propusemos este fator para verificar se havia alguma correlação entre as formas possíveis de sujeito e um determinado valor/uso do verbo em gramaticalização. Se isto ocorre representa um obrigatoriedade ou uma impossibilidade em dados contextos, o que tem a ver com o princípio da **obrigatoriedade**.

**Grupo 4- Forma verbal**<sup>7</sup>: **a)** presente do indicativo; **b)** pretérito imperfeito do indicativo; **c)** pretérito perfeito do indicativo; **d)** pretérito mais-que-perfeito do indicativo (simples e composto); **e)** futuro do presente; **f)** futuro do pretérito; **g)** presente do subjuntivo; **h)** pretérito imperfeito do subjuntivo; **i)** futuro do subjuntivo; **j)** imperativo afirmativo; **k)** imperativo negativo; **l)** infinitivo; **m)** gerúndio; **n)** particípio; **p)** pretérito mais-que-perfeito do subjuntivo (tempo composto); **q)** infinitivo composto; **r)** futuro do pretérito composto; **s)** pretérito perfeito composto; **t)** gerúndio composto.

**Grupo 5- Tempo verbal (categoria)**<sup>8</sup>: **P-** passado; **S-** passado até o presente; **A-** presente; **R-** presente para o futuro; **F-** futuro; **T-** onitemporal; **0-** tempo não marcado

**Grupo 6- Modalidades:** **#-** certeza; **1-** possibilidade; **+** probabilidade; **2-** necessidade (inclui intenção); **3-** volição; **4-** obrigação; **5-** proibição; **6-** ordem (inclui pedido, súplica, conselho); **7-** permissão; **8-** prescrição (inclui conselho); **9-** obrigatoriedade; **@-** permissibilidade; **0-** modalidade não marcada.

**Grupo 7- Tipo de texto fundame ntal:** **D-** descritivo; **V-** dissertativo; **I –** injuntivo; **N-** narrativo

Como uma de nossas hipóteses era a de que certos processos de gramaticalização do verbo ou de suas formas e categorias poderiam estar correlacionados com determinados tipos de textos em sua constituição e/ou ocorrência, propusemos este grupo de fatores para verificar se certos valores, usos ou funções do verbo em gramaticalização apareciam apenas ou preferencialmente em determinado tipo de texto.

Para a investigação do **aspecto**<sup>9</sup>, propusemos os seguintes grupos de fatores:

**Grupo 8 – Aspectos da duração:** **[ -** durativo; **\$ -** indeterminado; **% -** iterativo; **+ -** habitual; **? –** pontual; **0 –** aspecto não atualizado.

**Grupo 9 – Aspectos da realização:** **N –** não começado; **C-** começado; **A-** acabado; **0-** aspecto não atualizado.

**Grupo 10- Aspectos do desenvolvimento:** **1-** inceptivo; **2-** cursivo; **3-** terminativo; **0-** aspecto não atualizado

**Grupo 11- Aspectos do completamento:** **P –** perfectivo; **I –** imperfectivo; **0 –** aspecto não atualizado

O aspecto analisado para os fatores 8 a 11, no caso das locuções verbais, foi sempre o da situação narrada (Cf. TRAVAGLIA, 1985, p. 71-73) representada pela forma nominal, pois tal aspecto representa a categoria gramatical para a qual o verbo em gramaticalização está se tornando um “morfema”.

**Grupo 12- Pessoa:** **1-** primeira pessoa do singular; **2-** segunda pessoa do singular; **3-** terceira pessoa do singular; **4-** primeira pessoa do plural; **5-** segunda pessoa do plural; **6-** terceira pessoa do plural; **X-** impessoal; **0-** pessoa não atualizada.

Quando o sujeito era representado pelo pronome **você(s)** a análise foi feita considerando como segunda pessoa do singular (2) e do plural (5). Quando o verbo estava em uma das formas nominais, mas havia um sujeito identificável, na análise foi colocada a pessoa do sujeito, mesmo não havendo marca dessa categoria na forma verbal.

Os grupos de fatores 4 (forma verbal), 5 (tempo verbal), 6 (modalidades), 8 (Aspecto – duração), 9 (Aspecto – Realização), 10 (Aspecto – desenvolvimento), 11 (Aspecto – Completamento) e 12 (pessoa) foram propostos para verificar se, no caso dos valores, usos e funções dos verbos em estudo, em decorrência de **descategorização ou recategorização**, ocorreu ou está ocorrendo uma neutralização de marcas morfológicas, fazendo com que o item gramaticalizado: a) passe a ter forma única ou poucas formas que são usadas com a

<sup>7</sup> - Sobre a razão de se observar os grupos de fatores 4 a 12, comentamos no segundo parágrafo após o grupo de fatores 12.

<sup>8</sup> - As distinções de tempo passíveis de expressão no Português e aqui utilizadas são as propostas por Travaglia (1991).

<sup>9</sup> - O quadro de aspectos utilizado na análise é o proposto por Travaglia (1981).

função gramatical em surgimento ou b) tenha algumas formas de uso mais freqüente para a função gramatical em surgimento, o que representa um passo para **a**, acima.

Na verdade, segundo Heine (1993, p. 58-66), já no estágio 3 de gramaticalização os verbos em gramaticalização perdem a possibilidade ou capacidade de exprimir toda a gama de distinções de tempo, aspecto e modalidade. Além disso a perda da possibilidade de exprimir certas categorias próprias do verbo é característica dos verbos em gramaticalização a partir do estágio 4, tornando-os uma espécie de verbos defectivos. Assim, se isto se verifica, temos verbos em estágios mais avançados de gramaticalização.

**Grupo 13- Modalidade de língua:** **O** – oral; **E** – escrito; **R** - oral que foi escrito<sup>10</sup>; **S** – escrito para imitar o oral (escrito que busca reproduzir a oralidade que reproduz o oral/falado)<sup>11</sup>.

Quando propusemos este grupo de fatores achamos que seria pertinente verificar também os fatores **R** e **S**. Todavia no correr do estudo percebemos que essa distinção não seria produtiva, sobretudo em função do *corpus* com que trabalhamos, e por esta razão, mantivemos na análise apenas a distinção fundamental de modalidades de língua: **O** e **E**.

Propusemos este fator para verificar se algum processo de gramaticalização está acontecendo apenas em uma modalidade de língua, preferencialmente em uma modalidade ou igualmente nas duas. O objetivo não era medir o grau de gramaticalização, mas verificar se a origem da mudança estava em uma ou outra modalidade da língua. Confere-se também a hipótese corrente de que as mudanças lingüísticas preferencialmente começam na modalidade oral.

**Grupo 14- Tipo de amostra:** **1**- culta; **2**- não culta

O processo de mudança pode ser deflagrado em qualquer variedade da língua. Trabalhamos quase na totalidade apenas com uma variedade regional (a do Rio de Janeiro). Todavia separamos as ocorrências nas variedades culta e não culta, para ver se uma ou outra é o ponto de partida do processo de gramaticalização. Confere-se também a hipótese corrente de que as mudanças lingüísticas geralmente começam na variedade menos culta da língua.

**Grupo 15- Época:** **8** – segunda metade do século XX e século XXI; **7** – primeira metade do século XX; **6** – século XIX; **5** – século XVIII; **4** – século XVII; **3** – século XVI; **2** – século XV; **1** – século XIV; **0** – século XIII

A finalidade deste fator é verificar quando a gramaticalização e/ou seus estágios tiveram início / ocorreram. Além disso, verificar se um uso que apareceu em determinada época teve um incremento em sua freqüência, o que revela também que o processo de gramaticalização está avançando, já que a freqüência não só é um fator gerador, mas um fator de consolidação da mudança lingüística representada pelo processo de gramaticalização. Aqui trabalha-se com o tempo real.

**Grupo 16- Sexo do produtor:** **M**- masculino; **F**- feminino

Este fator verifica se o processo de gramaticalização está se originando e ocorrendo ligado a falantes de um ou outro gênero. Na verdade, aqui também se observam duas variedades dialetais da língua na dimensão do sexo/gênero. A proposição deste fator é motivada pelas observações dos estudos dentro da Teoria da Variação e Mudança, que sugerem que certas variações e mudanças são motivadas e/ou incrementadas pelo comportamento lingüístico de um dado sexo/gênero.

---

<sup>10</sup> - Conferência, jornal falado.

<sup>11</sup> - Peça de teatro, reprodução de diálogo ou fala em narrativas, entrevista.



17- **Idade do produtor (só para textos orais):** 1- 7 a 14 anos; 2- 15 a 25 anos; 3- 26 a 50 anos; 4- mais de 50 anos; / - não se aplica

Este fator só foi computado para os dados do *corpus* da **variedade oral**, uma vez que era impossível saber a idade dos produtores dos textos escritos, à época em que produziram os mesmos. Se a época representa o tempo real, a idade do produtor dos textos configura um tempo aparente e permite verificar basicamente os mesmos fatos que se confere com o fator da época: se há um aumento ou diminuição da frequência de dado valor, uso e função em determinadas faixas etárias. Usamos as mesmas 04 (quatro) faixas etárias com que trabalha o Projeto de Estudo sobre Usos da Língua / UFRJ – Rio de Janeiro. A idade do produtor também verifica a relação entre o processo de gramaticalização e as variedades dialetais de idade.

### 1.3. O *corpus*

Utilizamos nesta pesquisa um *corpus* de textos orais da segunda metade do século XX e início do século XXI e de textos escritos dos séculos XIII a XXI, conforme o QUADRO 1.

Para a língua oral usamos entrevistas do Projeto de Estudo sobre Usos da Língua (PEUL) e inquéritos (D2, DID e EF) do Projeto Norma Urbana Culta (NURC), ambos da Faculdade de Letras da Universidade Federal do Rio de Janeiro. Agradecemos aos dois projetos a cessão do material e a permissão para utilização dos mesmos nesta pesquisa.

Como se poderá observar o *corpus* tem mais material de variedades cultas da língua do que de variedades não cultas, uma vez que tomamos como cultos todos os textos escritos, inclusive os dos primeiros séculos de existência do Português, pois, com certeza, nestes, somente uma elite tinha a habilidade de escrever. Além disso os inquéritos do projeto NURC (em número de 16) são de variedade culta. Há um certo equilíbrio entre o material escrito e o oral tendo em vista o número de páginas de cada um, mas o material oral é mais extenso. O material contemporâneo é bem mais extenso que o material dos séculos XIII a XX (1ª metade). Estes, contudo, têm sempre uma extensão que vai de um mínimo de 20 páginas a aproximadamente 50 páginas. Julgamos não necessário tomar *corpus* de cada século da mesma extensão do *corpus* contemporâneo (segunda metade do século XX e século XXI), pois o objetivo não era fazer uma comparação de frequência de ocorrência entre as diferentes épocas, mas simplesmente buscar detectar desde quando determinado uso do verbo aparece na língua.

É preciso registrar ainda que um certo número de ocorrências utilizadas foi colhido em conversações espontâneas (cultas ou não, conforme o falante e a situação), novelas de televisão (cultas ou não conforme as características do personagem) e telejornais (sempre cultas, quando do repórter, e cultas ou não, conforme os falantes em entrevistas, depoimentos, etc.). Estas ocorrências não ultrapassam o número de 50 (cinquenta) em seu total.

Neste *corpus* encontramos 1961 (um mil novecentos e sessenta e uma) ocorrências dos verbos em estudo, distribuídas de acordo com o QUADRO 2. Observa-se uma predominância dos verbos de início (*começar* e *passar*) com 51,35% das ocorrências, seguidos dos verbos de término/fim (acabar, terminar, deixar) com 40,13% das ocorrências e finalmente do verbo de meio / continuidade (continuar) com apenas 8,52% das ocorrências. Entre os verbos de término/fim o verbo terminar apareceu com apenas 3,52% das ocorrências. Os valores para início e fim permitem hipotetizar que os falantes do Português tendem a marcar mais o início das situações, menos o seu término/fim e bem pouco a sua continuidade. Esta é uma hipótese a ser verificada em um *corpus* bem mais amplo e diversificado que o usado nesta pesquisa.

**QUADRO 2**  
**Número de ocorrências de cada verbo no *corpus***

VERBO	QUANTIDADE DE OCORRÊNCIAS NO <i>CORPUS</i>	PORCENTAGEM EM RELAÇÃO AO <i>CORPUS</i> ESTUDADO	PORCENTAGEM POR ÁREA
-------	--	--	----------------------

Começar	521	26,57%	51,35 %
Passar	486	24,78%	
Continuar	167	8,52%	8,52 %
Acabar	319	16,27%	40,13 %
Terminar	69	3,52%	
Deixar	399	20,34%	
TOTAL	1961	100 %	100%

## 2. A gramaticalização do verbo *começar*

### 2.1. Preliminares

Os usos gramaticais do verbo *começar* em seu conjunto são mais freqüentes (71,98%) do que os usos lexicais em seu conjunto (28,02%). Isto pode revelar que este verbo tende a se tornar um item da língua de natureza mais gramatical do que lexical. Não se pode dizer se a tendência levará ao desaparecimento dos valores lexicais. Nos usos gramaticais observa-se que o uso como ordenador textual (Valor 5) é pouco freqüente (1,15%) e o uso como marcador de aspecto começado e inceptivo (valor 4) é o que mais ocorre (70,83%). O verbo *começar* não apareceu no *corpus* com valor de verbo de ligação.

O valor lexical 1 (Dar começo a; principiar; iniciar; encetar) parece ser o que dá origem ao valor gramatical 4 que tende a ser um marcador de aspecto começado e inceptivo para o verbo na forma nominal, que representa a situação narrada, tanto que das ocorrências de valor 4 de *começar*, com aspectos de realização (grupo 9 de fatores) e de desenvolvimento (grupo de fatores 10) atualizados, apenas uma não apresenta estes aspectos. Retirando dos cálculos os (10,57% - 39/369) de ocorrências com estes aspectos não atualizados, pode-se dizer que em 99,67% dos casos este verbo está funcionando como um item gramatical marcador de aspectos inceptivo e começado.

Já o valor lexical 3 (Marca a primeira situação de uma série ou o início de uma série de situações), pode ser o que origina o valor gramatical 5 (ordenador textual), mas a pouca freqüência deste uso não nos permite fazer qualquer afirmação com mais segurança.

### 2.2. Valor 4: Marcador de aspecto começado e inceptivo

Quanto ao **grau de gramaticalização**, vejamos o que se pode dizer a partir dos fatores que o verificam, conforme relação apresentada em 1.2.

No que diz respeito aos fatores que verificam a **integração** observamos o seguinte para o verbo *começar* com o valor 4 na forma “começar + preposição + infinitivo”:

Temos intercalação (fator 2) de preposição em 86,92% dos casos, nenhum material intercalado em 2,91% e intercalação de algum outro material (ocorreram: negação, sintagma, elemento adverbial e dois ou mais tipos de material) em 10,17%. Considerando a hipótese de que a preposição é parte do “auxiliar”, sendo responsável com ele pela marcação do valor e que não representa uma real intercalação, teríamos **na verdade** 89,83% de ocorrências sem intercalação, o que revela um alto grau de integração. A intercalação de nenhum material, nem mesmo a presença da preposição, que começa a ocorrer na língua oral, reforça a proposição anterior e representa provavelmente um início de erosão da forma. Pode-se levantar esta hipótese porque os casos de não colocação da preposição (Cf. exemplos 15 e 16) ocorrem não apenas em casos em que se poderia falar de crase da preposição e da vogal inicial do verbo no infinitivo (Cf. exemplo 17).

- (15) .... tava trabalhando, NORMALMENTE, não tava fumando mais, comendo normalmente, aí de repente ele **começô passa** (hes) **vomitá, vomitá**, aí teve que í pro hospital, .... (Tend. Simone 27 anos, narrativo)
- (16) O budismo é parecido, entendeu? mas eu não sei nada de budismo. Tô... agora que eu **comecei lê**, e eu peguei esse livro justamente pra isso, pra lê alguma coisa, lê. (Tendência, Flávio, 26 anos, narrativo)
- (17) Aí ela não teve prá onde í, **começô arruma** emprego, e ninguém queria aceitá..... (Tend. Simone 27 anos).

No que diz respeito ao material intercalado é interessante observar que os sintagmas e pronomes que podem funcionar como argumento de verbo são na totalidade, argumentos do verbo na forma nominal e não do verbo em gramaticalização, ou simplesmente pronomes fossilizados<sup>12</sup>. O fato do verbo em gramaticalização não ser responsável por qualquer argumento na seqüência lingüística é um sinal de grau bastante avançado de gramaticalização em que o verbo já perdeu suas características sintáticas de verbo pleno (seria, talvez, já um verbo auxiliar, no estágio 4): o verbo *começar* não tem mais estrutura argumental.

Quanto aos demais fatores de integração observa-se que os resultados apontam para uma vinculação bastante significativa do verbo *começar* com seu principal, pois o verbo no infinitivo não representa argumento de *começar* em 100% dos casos (fator 3), o sujeito dos dois verbos é obrigatoriamente o mesmo em 100% das ocorrências (fator 4), e a forma nominal não é mais uma oração de qualquer natureza em relação ao verbo *começar*, mas uma forma dessentencializada em 100% das ocorrências (fator 5). Estes três fatores reforçam a conclusão de perda de estrutura argumental pelo verbo *começar* o que tem a ver com a perda de suas

---

<sup>12</sup> - Estamos considerando como pronome fossilizado os pronomes que são considerados parte integrante do verbo e também pronomes apassivadores e indeterminadores do sujeito.

**QUADRO 1 - O corpus da pesquisa**

<b>Tipo de texto</b>	<b>Veículo / Fonte</b>	<b>Época</b>	<b>Quantidade</b>	<b>Modalidade de língua</b>	<b>Faixa etária (Só para oral)</b>	<b>Registro</b>
Artigos	Jornais	Séc. XXI	50	Escrita	X	Culto
Artigos	Revistas	Séc. XXI	51	Escrita	X	Culto
Contos	Livros	Séc. XX (1 <sup>a</sup> )	9	Escrita	X	Culto
	Livros	Séc. XX (2 <sup>a</sup> )	14	Escrita	X	Culto
Anúncio e Romances	Livros	Séc. XIX	7	Escrita	X	Culto
Documentos oficiais e Cartas	Livros	Séc. XVIII	8	Escrita	X	Culto
Cartas, Documentos Oficiais, Sermões religiosos	Livros	Séc. XVII	8	Escrita	X	Culto
Crônicas históricas	Livros	Séc. XVI	3	Escrita	X	Culto
Crônicas históricas	Livros	Séc. XV	4	Escrita	X	Culto
Crônicas históricas	Livros	Séc. XIV	2	Escrita	X	Culto
Crônicas históricas e Documentos oficiais	Livros	Séc. XIII	4	Escrita	X	Culto
Entrevistas	PEUL-UFRJ Tendência	Séc. XX (2 <sup>a</sup> )	4	Oral	7 a 14 anos	Não culto
Entrevistas	PEUL-UFRJ Tendência	Séc. XX (2 <sup>a</sup> )	4	Oral	15 a 25 anos	Não culto
Entrevistas	PEUL-UFRJ Tendência	Séc. XX (2 <sup>a</sup> )	4	Oral	26 a 50 anos	Não culto
Entrevistas	PEUL-UFRJ Tendência	Séc. XX (2 <sup>a</sup> )	4	Oral	Mais de 50 anos	Não culto
D2- Diálogo entre 2 informantes	NURC-RJ	Séc. XX (2 <sup>a</sup> )	1	Oral	15 a 25 anos	Culto
D2- Diálogo entre 2 informantes	NURC-RJ	Séc. XX (2 <sup>a</sup> )	3	Oral	26 a 50 anos	Culto
D2- Diálogo entre 2 informantes	NURC-RJ	Séc. XX (2 <sup>a</sup> )	4	Oral	Mais de 50 anos	Culto
DID- Diálogo entre informante e documentador	NURC-RJ	Séc. XX (2 <sup>a</sup> )	2	Oral	26 a 50 anos	Culto
DID- Diálogo entre informante e documentador	NURC-RJ	Séc. XX (2 <sup>a</sup> )	2	Oral	Mais de 50 anos	Culto
EF – Elocução formal	NURC-RJ	Séc. XX (2 <sup>a</sup> )	2	Oral	26 a 50 anos	Culto
EF – Elocução formal	NURC-RJ	Séc. XX (2 <sup>a</sup> )	2	Oral	Mais de 50 anos	Culto

propriedades sintáticas. Os poucos casos em que se tem uma pausa entre *começar* e o infinitivo (fator 6) é na língua oral por força de hesitações do falante e portanto não representa uma quebra do vínculo entre componentes da perífrase (cf. exemplo 18). Na verdade, segundo os estudos de conversação e da língua oral, a hesitação acontece quase sempre em itens gramaticais (preposições, artigos, pronomes, etc.), enquanto o falante decide o item lexical a empregar<sup>13</sup>. O verbo auxiliar neste caso estaria então sendo usado como um item gramatical no comportamento de hesitação, pois a hesitação criadora de pausa se dá após o mesmo. O que se observa até aqui é a atuação muito forte do princípio da coalescência.

- (18) a- então quando você larga o carro... e **começa a... a subir...** ..... não... é descida mesma... (D2-158 homem 3ª faixa).  
b- Ovo mexido... aí [eu]...eu pego uma tigela assim, tigelinha pequeninha, pego uma coisa de batê, boto o ovo e **começo** [a]...a <ba...> **batê** um pouquinho o ovo ..... (Tendência, Maria Carolina, 10 anos, injuntivo)

Quanto ao fator 7 (mobilidade), o que se observa é a “fixação”, pois em 100% das ocorrências a forma nominal não pode vir para antes de *começar*, ou seja, este ocupa uma posição fixa na cadeia lingüística, o que revela maior gramaticalização.

Tendo em vista o princípio da especialização, um outro elemento que pode revelar o grau de gramaticalização é a **variedade ou não de formas** com nuances semânticas diferentes dentro do paradigma. No que diz respeito a variação de formas com o verbo *começar* temos o seguinte: a) a total predominância de “começar + a + infinitivo (94,58%), seja totalmente explícita (89,43%) ou com elipse de “a + infinitivo”(5,15%); b) a ocorrência da forma “começar + de + infinitivo” em 2,44% das ocorrências. Todavia todas as ocorrências com “de” são anteriores ao século XX (uma no século XIX e as demais no século XV) (cf. exemplo 19); c) a forma “começar + em + infinitivo” não apareceu no *corpus*; d) como já vimos na intercalação, a forma “começar + infinitivo”, que representaria um princípio de erosão da forma dominante na atualidade, apareceu em 2,98% dos casos. No que diz respeito às formas com “*começar*”, podemos dizer que está ocorrendo especialização. O verbo *começar* teria vários concorrentes em sua função de marcar aspecto começado e inceptivo: passar a, dar para, destampar a, desatar a, garrar a, agarrar a, pegar a, deitar a, despejar a, cair a, disparar a, romper a, danar a, desandar a, entrar a, iniciar, botar a, pôr-se a, principiar a, desenfrear a, desembestar a, sair. O único deles que estudamos sistematicamente foi o verbo “passar” e que, embora seja responsável por 486/1961 (24,78%) das ocorrências dos seis verbos estudados na pesquisa (cf. quadro 2), tem uso sobretudo lexical (83,95% das ocorrências) e no emprego gramatical concorrente (expressão de aspecto) tem uso bem menos significativo (13,79%) e não marca o início de uma situação, mas ocorre sempre com a nuance de início de um hábito pela instauração de uma situação que não existia antes. No que diz respeito aos demais concorrentes de *começar* não fizemos um estudo sistemático para verificar se está havendo especialização em favor de *começar*, mas propomos aqui esta hipótese. Este estudo tem que ser feito posteriormente. Todavia constatamos, no *corpus*, algumas ocorrências de verbos concorrentes de *começar* com o valor 4. (Ver exemplos em (20)). Como se pode observar, a maioria não é do que consideramos contemporâneo (a partir da segunda metade do século XX).

- (19) ..... que me parecia que jazia em o inferno; ca já **começava de sentir** aqui, em esta vida presente, as penas infernaes .... (EA13, séc. 15).

(20)

- a- começava a chover você **saía cavando** vala em volta (NURC-RJ / D2-158 homem 3ª faixa)  
b- Ah! Sr. Cirino, isto de filhos, são pedaços do coração que a gente arranca do corpo e **bota a andar** por êsse mundo de Cristo. (EA27, culto, homem, séc. 19, dissertativo)  
c- A gente fica olhando para a cara do outro e **dana a rir**. (Ouvido em aula de dança, 19/09/2001, Professor, homem, 18 anos, carioca, dissertativo, estudante do 2º grau).  
d- O presidente deu a ordem e o general **desandou a prender** todo mundo (Ouvido, 15/09/2001 – Peça de teatro ambientada na década de 1950, escrita por Miguel Falabella) (Mulher – Carmem Miranda, 40 anos)

<sup>13</sup> - Sobre hesitação na língua oral, ver os estudos de MARCUSCHI, nas publicações do Projeto Gramática do Português Falado, Editora da UNICAMP / FAPESP.

- e- Não demorou nada e o raio do saxofone **desandou a tocar**. (Texto 101, 1969, escrito, culto, mulher, narrativo)
- f- ..... então ele (*um professor*) ria dela... fazia piada dela... a turma **dava de rir** e eu ficava séria... eu achava aquilo uma covardia... né? (NURC-RJ/DID-261, 4ª faixa, mulher, descritivo)
- g- Um dia os lucros **entraram a escassear**. (Texto 91, Machado de Assis, 1906, narrativo)
- h- Daí em diante o rio **pegava a estreitar**-se entre barrancos atacados, até cair na cachoeira. (Texto 95, 1944, escrito, culto, homem, descritivo).
- i- Às vezes eu paro e fico a pensar / e sem perceber me vejo a rezar / o meu coração **se põe a cantar** / para a Virgem de Nazaré. (Canto religioso – Pe. Zezinho, contemporâneo, homem, mais de 50 anos).
- j- Mas **puseram-se a discutir** exaustivamente os preços. (Texto 108, 1931, escrito, culto, homem, narrativo)
- l- Ana Rosa **princiou a emagrecer** visivelmente. (EA22, culto, homem, séc. 19, narrativo)
- m- Bom, **princiou-se a comer** em silêncio, ..... (Texto 25 Conto 1942)

Quanto aos grupos de fatores (2, 3, 4, 5, 6, 8, 9, 10, 11, 12) do conjunto II de grupos de fatores que verificam o grau de gramaticalização, conforme especificado em 1.2, podemos observar o seguinte:

- 1) tem-se uma grande variedade de tipos de referentes para os sujeitos (grupo 2). Só não ocorreram o referente texto, período de tempo e mais de um tipo. Pelos critérios adotados isto revela maior grau de gramaticalização, a partir do estágio 3.
- 2) quanto ao grupo 3 de fatores (forma do sujeito), só não ocorreu sujeito oracional, não havendo uma especialização significativa quanto à forma do sujeito, revelando qualquer obrigatoriedade ou impedimento de uso, o que revelaria menor grau de gramaticalização neste particular;
- 3) quanto às formas verbais (grupo 4), tempo verbal (grupo 5), modalidades (grupo 6), aspecto-duração (grupo 8), aspecto-completamento (grupo 11), pessoa (grupo 12), não se pode observar nenhuma neutralização significativa de marcas morfológicas devida à descategorização ou recategorização. Algumas frequências mais altas de ocorrência como as das formas verbais do presente do indicativo e pretérito perfeito do indicativo; as dos tempos passado e onitemporal; a da modalidade de certeza; as de não-marcado para os aspectos dos grupos 8 e 11 e as de terceira pessoa se devem não a fatos ligados à gramaticalização, mas a correlações, por exemplo, com tipos de texto (as formas verbais, o tempo, a modalidade), formas verbais (os aspectos) e referente textual (as pessoas) (cf. Travaglia-1991);
- 4) sobre a frequência dos aspectos de realização (grupo 9 de fatores) e desenvolvimento (grupo 10 de fatores), já comentamos ao falar, no início deste item, da frequência do valor gramatical de marcador de aspectos do verbo *começar*. Como vimos, o verbo *começar* marca os aspectos começado e inceptivo para o verbo no infinitivo em praticamente 100% das ocorrências o que revela um bom grau de gramaticalização, pois já não há divergência.

Quanto à **ocorrência preferencial da gramaticalização em algum contexto lingüístico ou extralingüístico**, vejamos o que pode ser verificado a partir da análise quantitativa dos grupos de fatores (7, 13, 14, e 16) do conjunto II de grupos de fatores, conforme especificado em 1.2:

- 1) quanto ao tipo de texto fundamental (grupo 7), cremos que a maior ou menor frequência verificada para os tipos (descritivo: 8,13%, dissertativo: 16,80%; injuntivo: 2,44% e narrativo: 72,63%), não pode ser atribuída ao fato de a gramaticalização em foco estar sendo levada a termo mais por um tipo de texto (como o narrativo, por exemplo) do que por outro, porque a quantidade de *corpus* de cada tipo não é, em absoluto, equivalente e então a maior ocorrência em um tipo de texto pode se dever ao fato de que se tem mais material desse tipo que de outro. O que se pode afirmar sem problema é que o verbo *começar* com o valor 4 ocorre em todos os tipos de texto observados;
- 2) a gramaticalização de *começar* com valor 4 ocorre de forma aproximadamente equivalente na língua oral (59,35%) e escrita (40,65%); na língua culta (58,81%) e não-culta (41,19%). Quanto ao sexo há uma ocorrência maior com os homens (63,41%) do que com as mulheres (36,59%), mas isto pode se dever ao fato de que o *corpus* do sexo masculino é bem maior do que o *corpus* do sexo feminino, tanto na língua oral, quanto na escrita.

Quanto à **datação do processo de gramaticalização**, vejamos o que pode ser verificado a partir da análise quantitativa dos grupos de fatores (15 e 17) do conjunto II de grupos de fatores, conforme especificado em 1.2. Observa-se que esse valor de *começar* aparece desde o século XIII e em todas as faixas de idade. O que se observa, como falamos na variedade de formas, é uma especialização em favor da forma “começar + a + infinitivo” que, atualmente, na língua oral começa a sofrer uma erosão com perda da preposição “a”.

Infelizmente, tendo em vista a constituição do *corpus*, em termos de quantidade de material de cada época (predomina o *corpus* contemporâneo) e faixa de idade, não podemos fazer afirmações sobre se esta gramaticalização está se consolidando com um aumento de frequência no correr dos séculos em que está se processando ou se está sendo incrementada por alguma faixa de idade. Todavia pela sua frequência de ocorrência de 70, 83%, na época 8 (contemporânea) pode-se afirmar que é um valor e função bastante consolidados.

De tudo o que dissemos, pode-se propor que o verbo *começar* com o valor 4 (marcador de aspectos começado e inceptivo), está muito provavelmente no **estágio 4** de gramaticalização, se não plenamente, já francamente com diversas características deste estágio.

### 2.3. Valor 5: Ordenador textual

Quanto ao **grau de gramaticalização** vejamos o que se pode dizer a partir dos fatores que o verificam, conforme relação apresentada em 1.2 item (i).

No que diz respeito à **integração** (Conjunto I de grupos de fatores), apesar da forma ser de perífrase, não se pôde verificar nada tendo em vista o pequeno número de ocorrências do verbo *começar* com este valor (apenas seis ocorrências). O mesmo se pode dizer quanto aos grupos de fatores (2, 3, 4, 5, 6, 8, 9, 10, 11, 12) do conjunto II de grupos de fatores que verificam o grau de gramaticalização, conforme especificado em 1.2. A única observação que nos parece pertinente é o fato de este uso ter sérias restrições contextuais à sua ocorrência, que podemos assim especificar:

- a) em primeiro lugar seu sujeito (exemplos 21) ou objeto (exemplos 22 e 23), sejam eles explícitos ou elípticos, mas recuperáveis de alguma forma têm de ser um SN com referente do tipo “texto” ;
- b) em segundo lugar, quando em construção perifrástica, sempre ocorre com verbo de elocução (exemplo 22c e 23).

Todavia esta afirmação só tem status de hipótese a ser verificada num estudo com um número de ocorrências mais representativo. Estas restrições estão ligadas ao princípio da obrigatoriedade e revelam um certo grau de gramaticalização bem significativo.

- (21) a- Quando o *filme/livro começa*, o irmão mais velho..... (Texto 20, homem, escrito)  
b- Privilégio de Amá? Como é que se diz? É, a mãe...? **começô** assim: ela era novinha, aí [o]... o rapaz que ia sê padre eve relações com ela, ela ficô grávida,..... (Tendência, Simone, 27 anos). (O sujeito elíptico é “a novela” cujo nome ela citou: Privilégio de Amá)
- (22) a- podemos **começar** ? pelo jogo.... (NURC-RJ / D2-374 mulher 35 anos) (o objeto elíptico é “a conversa”).  
b- aí eu **estou começando** ? com uma descrição que não tem nada de descrição..... (NURC-RJ / D2-374, mulher, 4ª faixa) (o objeto elíptico é “minha fala = meu texto”)  
c- ..... o professor **começa**... ? **dando** um conceito do que seja signo lingüístico.. por exemplo... (NURC-RJ/EF-356, 3ª faixa, mulher, dissertativo) (o objeto oculto é “a aula”) (“dar um conceito” pode ser visto aqui como um verbo de elocução, pois tem a ver com dizer)
- (23) .....tenho que **começar** ? **dizendo**... a primeira vez que nós fomos à Europa... você se lembra bem o clima como era? (NURC-RJ/D2-369, 4ª faixa, mulher, injuntivo) (o objeto oculto é “minha fala”)

Quanto à **ocorrência preferencial da gramaticalização em algum contexto lingüístico ou extralingüístico**, vejamos o que pode ser verificado a partir da análise quantitativa dos grupos de fatores (7, 13, 14, e 16) do conjunto II de grupos de fatores, conforme especificado em 1.2. Aqui também o pequeno número de ocorrências impede qualquer afirmação segura sobre tipo de texto (grupo 7), modalidade de língua (grupo 13), tipo de amostra (grupo 14) e sexo/gênero do produtor do texto (grupo 16), todavia parece que podemos hipotetizar que é um uso mais culto, não só com base na frequência observada (83,33%), mas também na percepção de analista e na projeção de que este tipo de organização textual aconteceria mais em situações e textos mais formalizados. É interessante observar também que, como a maioria das ocorrências (5) é da língua oral e como nesta modalidade de língua os *corpora* estão equilibrados em extensão, pode-se pensar que as mulheres usam mais o verbo com esta função, ou este recurso de ordenação textual e poderiam ser as responsáveis por este processo de gramaticalização. Na verdade com apenas seis ocorrências estas duas últimas observações são meras hipóteses a serem verificadas com um *corpus* maior. Pode-se também observar outros recursos de ordenação textual<sup>14</sup>, e se não está havendo uma especialização desfavorável ao verbo *começar* com o valor 5, daí a pouca frequência. Alguns recursos alternativos são o verbo “iniciar”, que não estudamos (inicia o texto, o filme, a aula, a conferência, etc.) e seus cognatos como o nome (no início do texto o autor fala de.....) e o advérbio (Inicialmente expusemos os problemas, vamos agora propor possíveis soluções).

Quanto à **datação do processo de gramaticalização**, a análise quantitativa dos grupos de fatores 15 (época) e 17 (idade) do conjunto II de grupos de fatores sugere que este é um uso recente na língua (100% de ocorrências a partir de 1951). Se for este o caso, isto explicaria sua pouca frequência? Todas as ocorrências apareceram em falantes das 3<sup>a</sup> e 4<sup>a</sup> faixas etárias, o que sugere o uso por falantes já com maior elaboração lingüística, o que corrobora a hipótese da norma culta acima.

Propusemos hipóteses para a baixa frequência deste uso. Mais uma pode ser acrescentada: a de que explicitações sobre a ordem dos elementos no texto não são muito comuns, por não serem necessárias a não ser em casos especiais a serem determinados. Neste caso, uma frequência de 1,15% em 521 ocorrências do verbo “começar”, seria significativa em termos de uso como ordenador textual. Todas as hipóteses só podem se confirmar com um estudo mais amplo em que se consiga um número mais significativo de ocorrências.

### 3. Considerações finais

Como se pode observar, o verbo *começar* tem hoje um uso mais gramatical do que lexical sobretudo como marcador de aspecto. Não deixa de ser significativo o fato de que dois de seus valores lexicais parecem ser estágios na gramaticalização para dois de seus usos gramaticais. Parece muito interessante observar que o uso como ordenador textual é um uso relativamente novo, o que não é comum detectar-se nos estudos de gramaticalização. Alguns fatos verificados se apresentam mais como hipóteses a serem confirmadas por estudos posteriores com a observação de aspectos aqui apontados, tais como: ampliação do *corpus* em certos aspectos e verificação de ocorrência de recursos concorrentes. Cremos todavia que o estudo avança aspectos interessantes sobre o uso do verbo “começar” no Português.

### 4. Referências bibliográficas

BYBEE, Joan, PERKINS, Revere e PAGLIUCA, William. **The evolution of grammar: tense, aspect and modality in the languages of the world**. Chigago / London: The University of Chigago Press, 1994.

HEINE, Bernd. **Auxiliares: cognitive forces and grammaticalization**. New York / Oxford: Oxford University Press, 1993.

---

<sup>14</sup> - Sobre ordenação textual, seus princípios e recursos (marcadores de ordenação textual), veja Travaglia (1991: cap. 5, item 5.4).



TRAVAGLIA, Luiz Carlos (1981). **O aspecto verbal no português**; a categoria e sua expressão. Uberlândia: Editora da Universidade Federal de Uberlândia, 1996 (3<sup>a</sup> ed.). 552 pp.

\_\_\_\_\_. **Um estudo textual-discursivo do verbo no Português do Brasil**. Campinas: Tese de doutorado, IEL/ UNICAMP, 1991. 330 + 124 pp.

\_\_\_\_\_. A gramaticalização dos verbos começar / passar – continuar – acabar, terminar / deixar. In: TRAVAGLIA, Luiz Carlos. **Gramaticalização de verbos** – Relatório de pesquisa. Rio de Janeiro: Faculdade de Letras / UFRJ, Relatório de Pós-Doutorado em Lingüística, 2002. (131 pp.)

\_\_\_\_\_. Verbos gramaticais – verbos em processo de gramaticalização. In: TRAVAGLIA, Luiz Carlos. **Gramaticalização de verbos** – Relatório de pesquisa. Rio de Janeiro: Faculdade de Letras / UFRJ, Relatório de Pós-Doutorado em Lingüística, 2002a. (56 pp.)

\_\_\_\_\_. A gramaticalização de verbos. In HENRIQUES, Cláudio Cezar (org.). **Linguagem, conhecimento e aplicação** – Estudos de língua e lingüística. Rio de Janeiro: Editora Europa, 2003: 306-321.

\_\_\_\_\_. Verbos gramaticais – Verbos em processo de gramaticalização. In: FIGUEIREDO, Célia Assunção; MARTINS, Evandro Silva, TRAVAGLIA, Luiz Carlos e MORAES FILHO, Waldenor Barros. (orgs.). **Língua(gem): reflexões e perspectivas**. Uberlândia: EDUFU, 2003a: 97-157.