

II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA  
XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA  
A PERÍFRASE CONJUNCIONAL *SÓ QUE*:  
GRAMATICALIZAÇÃO E VARIAÇÃO LINGUÍSTICA

Luíza Francisca Ferreira da Silva (UFMG)  
[luizafrancisca@yahoo.com.br](mailto:luizafrancisca@yahoo.com.br)

RESUMO

A perífrase conjuncional *só que* da língua portuguesa encontra-se num processo de variação com a conjunção adversativa *mas*. Ambas podem funcionar como marcadoras da quebra de expectativa e coocorrem nessa função. Neste trabalho, demonstramos que a perífrase em questão é fruto de um processo de gramaticalização que uniu a palavra denotativa *só* e a conjunção integrante *que*. Dentro de uma perspectiva laboviana, elaboramos uma análise quantitativa das ocorrências de *só que*, comparando-as com as de *mas*, com base em 103 gravações de fala espontânea realizadas com 162 falantes do dialeto mineiro, sobretudo moradores da cidade de Belo Horizonte. Vale lembrar que esse corpus pertence ao projeto C-oral Brasil, associado ao Laboratório de Estudos Empíricos e Experimentais da Linguagem (LEEL) da Faculdade de Letras da Universidade Federal de Minas Gerais (UFMG).

Palavras-chave: *Só que*. *Mas*. Gramaticalização. Variação linguística.

1. Introdução

As conjunções se configuram como classes gramaticais extremamente favoráveis a inovações e a mudanças linguísticas ao longo de sua existência numa determinada língua. Para Meillet (*apud* LONGHIN-THOMAZI, 2002, p. 81), "embora todas as condições pareçam favorecer a estabilidade das conjunções no curso evolutivo das línguas, elas são palavras sujeitas à constante renovação e até ao desaparecimento".

Assim, podemos supor que itens pertencentes a essa classe no português brasileiro sejam afetados por processos de mudança.

A perífrase *só que* – fruto da junção da palavra denotativa *só* e da conjunção integrante *que* – é resultado justamente de inovações operadas no português brasileiro que, recorrendo a elementos já disponíveis, formou uma expressão conjuncional de valor adversativo, em frases como:

- (1) *tava no repertório // só que nós tiramos porque já tem uma quantidade / boa de música. (BFAMCV14)*

Vale ressaltar que a perífrase em questão se configurou a partir de um fenômeno de gramaticalização, em que, grosso modo, categorias lexicais tornam-se gramaticais ou categorias gramaticais tornam-se ainda

mais gramaticais. Além disso, devemos considerar também a possibilidade de tal perífrase estar num processo de variação linguística com outra forma adversativa prototípica da língua portuguesa, a conjunção *mas*, que ocorre em sentenças como:

(2) **Mauro e Filhos é um time muito legal / eu gostaria que eles continuassem /mas eles não são veteranos. (BFAMCV 01)**

Isso porque ambas são as formas mais utilizadas pelos falantes de português brasileiro para exprimir a quebra de expectativa, tradicionalmente chamada de adversatividade.

Até 1970, a gramaticalização foi vista estritamente como ramo da linguística diacrônica que servia de instrumento para analisar as transformações sofridas por determinada língua ou mesmo para recuperar a história de uma língua ou de um grupo de línguas. Tempos depois, surgiu uma nova perspectiva de análise, criada por Givón (*apud* CESARIO & ALONSO, 2013 p. 19), que sistematizava o seguinte percurso para o processo de gramaticalização:

**Discurso → sintaxe → morfologia → morfofonêmica → zero**

Para criar esse percurso, Givón (1995) considerou alguns pressupostos iniciais, dentre os quais podemos destacar:

- i) a linguagem é uma atividade sociocultural;
- ii) as estruturas linguísticas têm funções comunicativas;
- iii) tais estruturas não são arbitrárias, mas motivadas;
- iv) mudança e variação estão sempre presentes;
- v) o significado é contextual;
- vi) as gramáticas são emergentes;
- vii) as regras gramaticais são permissivas.

Desse modo, vê-se que Givón admite que os processos de transformação pelos quais passa uma determinada língua são motivados pelo uso, isto é, para ele, a gramática, a estrutura gramatical de uma língua, emerge de seu discurso, do uso que se faz dessa estrutura. Portanto,

*discourse over time gives rise to the emergence of syntactic constructions, which in turn over time become morphologized. Progressive phonological erosion and lexicalization eventually lead to the loss of overt grammatical material. (KRUG, 2000, p. 9)*

## II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA

Nota-se ainda que esse percurso criado por Givón desvincula os processos de gramaticalização exclusivamente da diacronia e evidencia a possibilidade de analisarmos fenômenos de mudança sincronicamente. Apesar disso, ainda se pode dizer que a gramaticalização carece de um lugar teórico definido. A respeito dessas questões, Lorenzo Teixeira Vitral & Jânia Martins Ramos (2006, p. 9) tecem as seguintes considerações:

na literatura recente, encontra-se muito fecunda a discussão acerca da questão de saber se a gramaticalização constitui, epistemologicamente, um modelo teórico de pleno direito ou se, ao contrário, os chamados processos de gramaticalização são, tão-somente, epifenômenos que poderiam ser deduzidos de enunciados oriundos de outros quadros teóricos. A questão é, evidentemente, complexa e envolve tomada de decisões que, seguramente, ultrapassam o campo da linguística, estritamente considerada.

Portanto, nota-se, diante da relevância e da complexidade que esses estudos têm assumido atualmente na literatura, que é de extrema importância que as novas formas gramaticalizadas surgidas na língua portuguesa, como a perífrase conjuncional *só que*, sejam analisadas.

Sobre isso, podemos dizer ainda que, conforme Lilian Sousa & Lorenzo Teixeira Vitral (2010, p. 9)

a teoria da variação e mudança, por avaliar uma possível mudança em curso, com a apresentação de fatores externos e internos que demonstram sua estabilidade ou não, apresenta uma metodologia que pode contribuir para a avaliação de um processo de gramaticalização. Se os resultados de uma análise nessa metodologia indicam mudança em progresso para uma forma que tende a tornar-se mais gramatical, trata-se de um perfil compatível com as etapas previstas por um processo de gramaticalização. Por outro lado, se há estabilidade no uso das formas ditas concorrentes, pode-se supor que tais formas caminham para funções diferentes, isto é, sofreram recategorização, o que poderia indicar também a gramaticalização de uma das formas.

Logo, a teoria da variação torna-se um instrumento de análise do fenômeno da gramaticalização. Neste artigo buscamos explicitar justamente essa interface entre os dois campos para verificarmos a gramaticalização da perífrase *só que* e sua variação com a conjunção *mas*.

### 2. Aspectos teórico-metodológicos

Pretende-se, primeiramente, demonstrar que a perífrase conjuncional *só que* do português brasileiro contemporâneo é fruto de um processo de gramaticalização que uniu a conjunção subordinativa prototípica

*que* e a forma adjetiva do português arcaico *sola*, que funcionava, esporadicamente, como palavra denotadora, como em:

- (3) *en hũa persona sola & de hũa en outra* (port. “em uma pessoa só e de uma em outra” – tradução nossa)<sup>24</sup>

Em seguida, à luz da Sociolinguística Variacionista, pretende-se comprovar que existe uma concorrência entre a forma conservadora *mas* e a inovadora *só que*, já que a segunda adquire o valor adversativo ou de *quebra de expectativa*, que é normalmente expresso pela primeira, deixando-as, pois, com o mesmo valor de verdade. Além disso, pretende-se verificar, com base em mostras de língua falada, se existe a tendência de a forma inovadora suplantará a conservadora.

Para fundamentarmos nossas suposições, vamos dissertar sobre a teoria da gramaticalização, a fim de compreendermos as características do processo no qual a perífrase em estudo se insere. Faremos também brevemente um histórico das conjunções na língua portuguesa, de modo a demonstrar a junção dos termos *só* e *que* como um processo natural e até previsível para a formação de novas perífrases conjuncionais.

Após essa exposição, vamos comparar, na seção *Análise dos dados*, as ocorrências de *só que* e de *mas*. Essa comparação nos dará os elementos necessários para verificarmos ou não a possível concorrência entre as duas formas da língua portuguesa. Para tanto, vamos recorrer a amostras de língua falada que evidenciem o uso dessas estruturas na fala cotidiana dos brasileiros.

## 2.1. A gramaticalização

Sobre a gramaticalização, devemos considerar que é um processo linguístico que implica a mudança de estatuto de um item lexical ou de uma construção sintática para um item ou construção essencialmente gramatical. Uma vez gramaticalizado, esse item continua a desenvolver novas funções ainda mais gramaticais (MARTELOTTA, VOTRE & CEZARIO, 1996, p. 24). Nesse processo, quanto mais o elemento avança para a gramática, mais regular e previsível ele se torna, pois abandona o nível da criatividade eventual do discurso para assumir funções restritas ao nível gramatical.

---

<sup>24</sup> Os exemplos em latim foram retirados de Longhin-Thomazi (2002).

## II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA

Como fundador dos estudos modernos sobre gramaticalização e introdutor do termo, Antoine Meillet (*apud* CEZARIO & ALONSO, 2013, p. 18) defendia que esse conceito servia para definir um processo responsável por originar novas formas linguísticas. Para ele, esse processo levaria à transformação do sistema geral das línguas, ao introduzir novas categorias para suprir expressões linguísticas até então inexistentes. Nesse sentido, Antoine Meillet postulou dois eixos principais na gramaticalização:

- i) usos gramaticais provêm de usos mais lexicais;
- ii) há um contínuo no processo que vai do concreto ao abstrato.

Quanto ao primeiro eixo, podemos entendê-lo como o fato de um termo lexical, mais flexível dentro da cláusula, perder sua expressividade pelo uso recorrente e transformar-se num termo gramatical, mais preso na sentença e mais vazio de sentido. Esse foi o caminho do item *ir* em português. Usado inicialmente como verbo, que dá ideia de deslocamento de um lugar para outro, em frases, como *ele vai ao shopping*, devido ao uso frequente e ao conseqüente desgaste dessa ideia de movimento, passa também a funcionar no português como auxiliar, em frases como: *ele vai fazer o trabalho*. Neste caso, o auxiliar esvazia seu sentido inicial de movimento e passa a ser entendido semanticamente de acordo com o significado do verbo que acompanha, nesse caso, *fazer*.

Além disso, o auxiliar torna-se mais estático dentro da sentença, precisando necessariamente estar junto ao verbo principal, o que não acontece quando *vai* funciona como verbo pleno, podendo se deslocar no interior da sentença, por exemplo: *ao shopping ele vai*, o que não ocorre com o auxiliar: *\*ele vai o trabalho fazer*.

Quanto ao segundo eixo postulado por Meillet, vale lembrar que o processo de gramaticalização envolve a abstratização, “já que o significado não-gramatical pode ser descrito como mais concreto do que o significado gramatical” (NEVES, 1997, p. 131). Nesse sentido, pode-se dizer que a gramaticalização é um fenômeno de base metafórica, na medida em que envolve a “conceptualização obtida na expressão de uma coisa por outra” (NEVES, 1997, p. 133). Falaremos mais adiante a respeito disso, mencionando a escala de abstratização de Heine et al. (1991a, 1991b)

Retomando o esvaziamento semântico a que está sujeito um item a ser gramaticalizado, pode-se dizer que esse esvaziamento está muito ligado à frequência de uso do item lexical. Bybee (*apud* CEZARIO;

ALONSO, 2013, p. 22) considera que sequências de palavras ou morfemas frequentemente usados se tornam automáticos:

a frequência leva a uma automatização das formas, que podem reduzir-se foneticamente e também costumam emancipar-se no sentido de preencher novas funções em novos contextos. Um item muito usado num determinado contexto linguístico torna-se previsível, automático e geralmente tem sua forma fonológica reduzida.

Mário Eduardo Martelotta (2010, p. 147) também descreve muito bem esse processo:

um item – ou uma construção – começa a ser utilizado com muita frequência em um determinado contexto. Nesse caso, pela força do hábito, o usuário trabalha com o item ou com a construção mais automaticamente. Com isso o item – ou a construção – perde estrutura fonética. Quanto mais frequente – e consequentemente mais previsível – em um determinado contexto é a informação transmitida por um elemento linguístico, mais ele tende a ter sua estrutura fonética reduzida, até por uma questão de economia.

Outro ponto que devemos considerar é o fato de que nosso objeto de estudo, a perífrase *só que*, não é composta de um elemento lexical apenas, mas sim de dois elementos que se juntaram para formá-la. Maria Marta Cezario e Karen Sampaio Braga Alonso (2013, p. 29) reconhecem que

o processo de gramaticalização pode ser também analisado não apenas a partir de palavras e morfemas, mas também a partir de unidades morfossintáticas mais complexas, como construções de mais de uma palavra. É o caso, em português, de *apenas* (resultante da gramaticalização das unidades *a e penas*), *agora* (de *hac ora*), *embora* (de *em boa hora*) etc. (...). Ao longo do tempo, itens específicos vão sendo usados com outros para expressar novas funções comunicativas; a partir daí, padrões de uso vão se modificando pragmaticamente e se cristalizando em pareamentos de forma e sentido. Uma sequência de itens não se torna necessariamente uma única palavra. Unidades linguísticas maiores podem ser consideradas, nesse sentido, construções oriundas de gramaticalização. Do mesmo modo que formamos itens do léxico, como *planoalto*, *Maria-vai-com-as-outras*, *guarda-roupa*, também formamos itens da gramática, como *apesar de*, *embora*, *talvez*.

Quanto a isso, Heine *et al.* (*apud* MARTELOTTA, 1996, p. 106) argumentam que o caminho mais natural ao falante de uma língua é recorrer a elementos já conhecidos para, com base neles, criar novos significados para suprir necessidades comunicativas que surgem em seu cotidiano. “Essa estratégia segue o princípio de que conceitos concretos são utilizados para descrever conceitos menos concretos e mais difíceis de serem conceptualizados”.

## II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA

Considerando isso, Heine propõe a seguinte escala para demonstrar o percurso de um item lexical ao ser reutilizado para novas funções ou acepções:

**Pessoa > Objeto > Atividade > Espaço > Tempo > Qualidade**

Observando a escala, chegamos também a outro ponto fundamental da teoria da gramaticalização, que é o conceito de unidirecionalidade. Segundo esse princípio, o processo de gramaticalização se daria do concreto para o abstrato, do lexical para o gramatical, e não o contrário. Mário Eduardo Martelotta, Sebastião Josué Votre e Maria Maura da Conceição Cezario (1996, p. 27) argumentam a favor de que elementos mais concretos, justamente por serem mais fáceis de serem conceitualizados, são utilizados para expressar noções abstratas, por sua vez mais difíceis de serem sistematizadas.

Os autores ilustram a unidirecionalidade com o fato de partes do corpo serem utilizadas para expressar noções abstratas. Isso é o que ocorre, por exemplo, com a palavra *braço*, que servindo originalmente para expressar um membro do corpo, é transferida para expressar objetos, como em *braço da cadeira* e, avançando no processo, é transferida para noções metafóricas, como *braço direito*.

Nota-se que a complexidade do fenômeno da gramaticalização pode nos servir inclusive para pensarmos os processos de mudança linguística, pois seria justamente esse o nosso objetivo ao, primeiramente, comprovarmos que a perífrase conjuncional *só que* seria fruto da gramaticalização e, além disso, estaria num processo de variação com outra forma adversativa, a conjunção *mas*.

### **2.2. As conjunções e locuções conjuntivas no português brasileiro**

Como dito, precisamos dissertar também sobre a teoria que permeia as conjunções e locuções conjuntivas, sobretudo dentro da coordenação, já que nosso objeto de estudo se insere nessa categoria. Segundo a tradição gramatical, as conjunções são as “expressões que ligam orações ou, dentro da mesma oração, palavras que tenham o mesmo valor ou função”. (BECHARA, 1973, p. 159)

Para Antoine Meillet (*apud* LONGHIN-THOMAZI, 2003, p. 81), a expressividade vai comandar o surgimento de novas conjunções. Se-

gundo ele, “o valor expressivo das palavras é sempre transitório” e, nesse contexto, dois fatores são atuantes para a perda de expressividade de uma palavra e a consequente necessidade de sua substituição por outra mais carregada de conteúdo. O primeiro fator diz respeito à perda de conteúdo fonético. Nesse caso, segue-se a tendência natural de as línguas perderem substância fonética devido ao uso pelos falantes. Estes tendem, pelo princípio da economia linguística, a “simplificar” as palavras em sua fala cotidiana de modo a dar mais dinamismo ao processo de comunicação. Essa redução, entretanto, leva inevitavelmente a uma perda semântica, já que a palavra perde também partes significativas em seu processo de redução.

O segundo fator determinante é a frequência de emprego do item. Antoine Meillet (*apud* LONGHIN-THOMAZI, 2002, p. 10) diz que, quanto mais um termo é empregado, mais desgastado ele fica. Como dito, esse desgaste faz com que o item aos poucos tenha o seu significado básico esquecido, sendo necessária a criação de um novo termo que o substitua nesse significado básico. Quando isso ocorre, é bem provável que esses itens que receberam novas funções caiam também em desgaste, percam expressividade e sejam substituídos por novas formas, que, por sua vez, também assumirão novas funções. Desse modo, vê-se que o ciclo de renovação de uma língua é infinito.

Nesse sentido, dois aspectos fundamentais ao desenvolvimento das conjunções são apontados (MEILLET *apud* LONGHIN-THOMAZI, 2002, p. 82-83):

- i) as conjunções constituem uma classe de palavras dominada por uma necessidade contínua de transformação;
- ii) palavras de classes diferentes podem ser recrutadas para assumir o papel de conjunção, por meio de mudanças semânticas pressionadas pelo contexto de uso.

Com relação a isso, Longhin-Thomazzi (2002, p. 106) acrescenta dois pontos complementares:

- iii) o português habilitou palavras de natureza diversa - adverbial, preposicional, pronominal e nominal - ao papel de conjunção.
- iv) o português generalizou o processo - iniciado no latim vulgar - que consiste em combinar a partícula subordinativa *que* com palavras de diferentes categorias, para a formação de perífrases conjuncionais.

## II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA

No caso de *porém*, por exemplo, trata-se de uma forma filiada ao advérbio latino *per inde*, que dava o sentido de causa/ explicação às orações que ligava, como em:

- (4) “*E vtry-se tam toste que o nom poderedes já oje acalçar. E porende vos louvaria de ficardes*<sup>25</sup>. (port. “E vai tão cedo que já não podes alcançar. Por isso acho bom que fiques”).

Para Manuel Said Ali (1966, p. 187-188),

a língua usual privilegiou o uso do termo mais curto, condenando *porende* ao desaparecimento. Com o tempo, o emprego de *porém* em contextos específicos de contrajunção (assinalados principalmente pela presença da negativa) levou-o a uma transformação semântica: em lugar de significar “por isso”, “por essa razão”, passa a significar “apesar disso”, “mas”, “contudo”. *E ainda que quando o levaram diante del-rei desmaiou, não defalleceu porem em sua firmeza, mas foi hum natural pejo*. Assim, enquanto anafórico, *porém* era advérbio (“por isso”), mas a partir do momento em que co-ocorre com a negação, esse primitivo advérbio começa a funcionar como conjunção adversativa. De seu elemento fonte (o advérbio *per inde*), a conjunção *porém* só conservou a mobilidade característica dos advérbios, que lhe garante emprego no início, intercalada ou no final de sentenças.

Vemos, portanto, em *porém*, um caso em que a necessidade de expressão comanda o surgimento de uma nova conjunção, a partir de uma outra classe gramatical, nesse caso, um advérbio.

O latim clássico conservou a antiga partícula enclítica *que*, que servia para unir palavras ou orações que apresentavam estreitas relações de sentido entre si, como *domi bellique* (port. “em casa e na guerra”) e *senatus populusque* (port. “o senado e o povo”); ou ainda para dar a um termo da oração a ideia expressa em outro termo como em: *ille numquam i/la dicet jacta fortuito naturalemque rationem omnium reddet* (port. “ele nunca dirá aquelas coisas feitas por acaso e chegará à razão natural de tudo”). (LONGHIN-THOMAZI, 2002, p. 84)

A estreita relação que o *que* estabelecia entre os termos os quais encadeava era tamanha que essa partícula se colocava em posição enclítica em relação a esses termos. Essa dependência dos itens a *que* fez com que essa partícula se configurasse como uma referência no quadro das conjunções latinas. Tanto que ocorria em vários casos clássicos, como em:

- (5) *atque haec urbs atque imperium* (port. “esta cidade e até este império”)

---

<sup>25</sup> Os exemplos em latim foram retirados de Longhin-Thomazi (2002).

(6) *quoque qua de causa Heluetii quoque* (port. "por essa razão os Helvécios também")

No latim vulgar, entretanto, as diversas conjunções da língua clássica não foram adotadas. Nesse sentido, as línguas românicas trataram de criar novas construções. Para a subordinação, a língua vulgar dispunha de poucas conjunções e “passou a usar por toda parte um *quod* (que posteriormente sofreu concorrência de *quia* e de *quid*) para assinalar as diversas formas de subordinação (LONGHIN-THOMAZI, 2002, p. 95). Essas formas se tornaram, então, a conjunção subordinativa por excelência do latim vulgar. “Quando necessitava de uma nova conjunção, a língua tendia a formar perífrases ou locuções conjuncionais a partir da combinação de uma base preposicional/ adverbial com *quod*” (LONGHIN-THOMAZI, 2002, p. 103). Esse fenômeno teve ainda mais produtividade nas línguas românicas.

Câmara Jr. (*apud* LONGHIN-THOMAZI, 2002, p. 105) descreve a prevalência de *que* a partir de *quid* da seguinte maneira:

o advento da conjunção subordinativa *que* resultou primordialmente de um esvaziamento da significação pronominal da forma neutra *quid* do pronome indefinido-interrogativo e sua coalescência com a outra forma neutra *quod*, reservada ao pronome relativo. Secundariamente, houve a convergência da evolução fonética da partícula de conexão comparativa *quam* e da conjunção causal *quod*. De tudo isso, resultou uma partícula multifuncional *que* para os mais variados padrões frasais.

Considerando, então, esse caráter “multifuncional” de *que*, pode-se concluir, então, que as conjunções como as conhecemos hoje, resumidamente, são fruto de um processo duplo: houve perda do latim clássico ao vulgar no quadro de conjunções, e a entrada de *que* assumindo proeminência. Depois dessa perda, novas perífrases com base em *que* vão sendo formadas para suprir a carência deixada pela passagem do latim clássico ao vulgar e posteriormente às línguas românicas.

Descoberta, pois, a origem do *que* que compõe nossa perífrase, falta-nos analisar a partícula *só*, também constituinte da locução em estudo. Tradicionalmente, o item *só* pertenceria à classe das palavras denotativas, classe esta que englobaria os termos que não se puderam encaixar nas categorias tradicionais, como a do substantivo e a do adjetivo, por exemplo. Desse modo, classificava-se como palavra denotativa todo item que não se conseguia conceptualizar em outra classe.

## II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA

Dentro das palavras denotativas, o item *só* que nos interessa encaixar-se-ia nas de exclusão, restrição ou limitação, no sentido de “exceção”, “apenas”, como em:

(7) <eu queria ver a comunidade dele> lá / ver que que es falam / entre si / assim / só pra eles mesmos. (EVN: [29])

Etimologicamente, devemos observar que, primeiramente, havia no latim a forma adjetiva *solus*, que possuía o feminino *sola* e o neutro *solum*. Estas duas frequentemente fundiam-se com outras estruturas, formando novas construções: *non solum ... sed etiam* (port. “não só... mas também”) ainda no latim, e *tã solamente por fartar ao apetito* (port. “tão somente satisfazer ao apetite”), encontrada já no português arcaico do século XIV.

Como não encontramos *solus* ou *solum* no português, apenas no latim, mas encontramos *sola* no português arcaico<sup>26</sup>, temos pistas para acreditar que a forma latina *sola* predominou sobre a neutra e a masculina e deu origem a *só*, palavra denotativa no português contemporâneo.

Vale ressaltar que encontramos mostras de uma forma *so* no português arcaico do século XIV, mas, como funcionava sempre como preposição, como em: “*nõ façam y outra malícia so pea da almotaçarya*” (port. não façam outra malícia sob pena de taxaço), excluímos dessa forma a responsabilidade por derivar a palavra denotativa *só*. Isso também porque a forma adjetiva *sola* funcionava, como vimos, esporadicamente como denotadora, o que não ocorreu com a preposição arcaica *so*. Portanto acreditamos que *sola*, por funcionar às vezes como denotação no português arcaico, deu origem ao atual *só* palavra denotativa.

### 2.3. A teoria variacionista

Para finalizar nosso eixo teórico, que parte da teoria da gramaticalização e perpassa a teoria sobre as conjunções coordenativas da língua portuguesa, precisamos discutir conceitos fundamentais da teoria da variação e mudança, já que esse modelo teórico possui uma metodologia de análise de fenômenos linguísticos que nos servirá para pensar as ocorrên-

---

<sup>26</sup> A pesquisa do português arcaico foi realizada no site <<http://www.corpusdoportugues.org>>, idealizado por Mark Davies e Michael Ferreira, pesquisadores do United States National Endowment for the Humanities.

cias de *mas* e *só que*, formas prototípicas para marcar o contraste no português brasileiro.

Como metodologia, usaremos o estudo da mudança no tempo aparente, modelo que pressupõe que

o comportamento linguístico de cada geração reflete um estágio da língua, com os grupos mais jovens introduzindo novas alternantes que, gradativamente, substituirão aquelas que caracterizam o desempenho linguístico dos falantes de faixas etárias mais avançadas (...). O processo de aquisição da linguagem se encerra mais ou menos no início da puberdade, estabilizando-se a partir desse momento ou, pelo menos, não sofrendo modificações significativas a partir de então. Assim a fala de um indivíduo com 75 anos, no ano 2000, representaria um estado de língua de sessenta anos atrás, ou seja, 1940. (PAIVA & DUARTE, 2003, p. 14)

Logo, esses estudos permitem distinguir mudanças que se produzem de forma gradual em toda a comunidade linguística e processos variáveis que implicam gradação etária. Desse modo, os movimentos de mudança podem ser apreendidos no seu curso de implementação. Nesse sentido, a observação dos dados presentes no corpus a ser coletado nos permitiria verificar a possível variação entre *mas* e *só que* e ainda constatar uma possível mudança em progresso, etapas previstas pela Sociolinguística Variacionista.

### 3. *Análise dos dados*

Como corpus que serviu de base para este artigo, selecionamos 103 gravações de fala espontânea realizadas com 161 falantes mineiros, moradores de Belo Horizonte e região metropolitana. Essas gravações constituem o produto principal do projeto C-ORAL-BRASIL, desenvolvido pelo Núcleo de Estudos em Linguagem, Cognição e Cultura (NELC) e pelo Laboratório de Estudos Empíricos e Experimentais da Linguagem (LEEL) da Faculdade de Letras da Universidade Federal de Minas Gerais, sob a coordenação dos professores Heliana Mello e Tommaso Raso.

Em nossa coleta, foram encontradas 91 ocorrências da perífrase *só que* e 1032 da conjunção *mas*. Como nosso objeto de estudo são essas formas na função de marcadoras de quebra de expectativa ou adversatividade, na tradição gramatical, foram excluídas as ocorrências desses itens que representassem outros tipos de usos, como as ocorrências de *só que* que não atuassem como perífrase, como no exemplo 8 a seguir, em

II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA  
XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA

que *só* é palavra denotativa de exclusão, apesar de aparecer diante de *que*:

(8) *querer só que a pessoa mude.* (CAB: [37])

e as ocorrências de *mas* como marcador de realce, por exemplo, em:

(9) *pegou ele / mas pegou até.* (MNV: [143])

Num primeiro momento, parece-nos previsível que a forma conservadora ocorra consideravelmente mais que a inovadora, inclusive porque esse é o caminho prototípico da mudança, pois, como dissemos, é natural ao item conservador desgastar-se justamente por ser muito usado, o que suscita o surgimento de novos termos que o substituam. Nesse caso, *mas* ocorre aproximadamente dez vezes mais que *só que*.

No C-oral Brasil, os falantes foram estratificados segundo sua idade, sexo e nível de escolaridade. Quanto ao fator idade, dividiram-se os entrevistados em cinco faixas: A: 18 a 25 anos; B: 26 a 40 anos; C: 41 a 60 anos e D: maiores de 60 anos. Quanto ao fator sexo, foram divididos em masculino e feminino. Por fim, quanto ao nível de escolaridade, foram agrupados nas seguintes categorias: faixa 1: ausência de escolarização ou escolarização até o nível primário incompleto (não mais de 7 anos de escolarização); faixa 2: até o terceiro grau completo, desde que exerça uma profissão que não requer o terceiro grau; faixa 3: falantes que possuem terceiro grau e exercem uma atividade que exige esse título.

Em nossas 103 entrevistas, que totalizaram 161 falantes, encontramos, como dito, 1032 ocorrências de *mas* e 91 de *só que*. Nesse corpus, havia 71 homens. Dentre eles, 20 pertenciam à faixa etária A, 22 à faixa B, 22 à faixa C e 7 à D. Além disso, 12 possuíam até 7 anos de escolarização, 28 possuíam até 3º grau completo e 31 possuíam 3º grau completo e exerciam profissão que dependia dessa titulação. Quanto às mulheres, foram entrevistadas 90 ao todo. Dentre elas, 29 pertenciam à faixa etária A, 28 à faixa B, 28 à faixa C e, por fim, 5 à faixa D. Além disso, 13 possuíam ensino fundamental incompleto, 44 possuíam ensino superior completo sem exercer profissão ligada a ele e 33 possuíam ensino superior completo e exerciam profissão ligada a esse nível.

Após verificar comparativamente as ocorrências de *só que* e *mas* segundo essa estratificação, encontramos os seguintes resultados: com relação ao fator sexo, homens realizaram *mas* 421 vezes e *só que* 46 vezes; as mulheres realizaram *mas* 611 vezes e o *só que* 45 vezes. Com relação à idade, a faixa A realizou 337 vezes o *mas* e 31 vezes *só que*; a faixa B

realizou *mas* 275 vezes e *só que* 35 vezes; a faixa C realizou *mas* 348 vezes e *só que* 17 vezes, e, por fim, a faixa D realizou *mas* 72 vezes e *só que* 8 vezes. Já com relação ao fator nível de escolaridade, a faixa 1 realizou *mas* 176 vezes e *só que* 3 vezes; a faixa 2 realizou *mas* 486 vezes e *só que* 55 vezes e a faixa 3 realizou *mas* 370 vezes e *só que* 33 vezes.

Portanto, nota-se que em todos os fatores extralinguísticos considerados, a forma simples de adversatividade é mais frequente que a perifrástica. Entretanto cabe-nos elucidar algumas observações. Na tabela a seguir, explicitamos a distribuição dos itens analisados em relação à variável sexo.

<b>Distribuição de <i>mas</i> e <i>só que</i> por sexo</b>	<b>mas</b>	<b>%</b>	<b>só que</b>	<b>%</b>
Sexo masculino	421	40.8	46	50.5
Sexo feminino	611	59.2	45	49.5
Total	1032	100.0	91	100.0

**Quadro 1: Distribuição de *mas* e *só que* em relação ao sexo.**

Com base no quadro, parece-nos relevante ressaltar sobretudo o fato de as mulheres terem realizado em maior quantidade *mas* do que os homens. É evidente que isso se deve ao fato de termos mais falantes do sexo feminino (90) do que do sexo masculino (71). Entretanto, como essa diferença substancial não se mostrou também em relação a *só que*, podemos considerar preliminarmente que essa forma não é estigmatizada socialmente, pois é utilizada invariavelmente por homens e mulheres.

Observemos agora a distribuição de nossos objetos de estudo em relação à variável faixa etária:

<b>Distribuição de <i>mas</i> e <i>só que</i> por faixa etária</b>	<b>mas</b>	<b>%</b>	<b>só que</b>	<b>%</b>
Faixa A	337	32.7	31	34.1
Faixa B	275	26.6	35	38.5
Faixa C	348	33.7	17	18.7
Faixa D	72	7.0	8	8.8
Total	1032	100.0	91	100.0

**Quadro 2: Distribuição de *mas* e *só que* em relação à faixa etária.**

Nota-se que o uso de *só que* é predominante entre falantes de idade mediana, cuja faixa etária está no intervalo de 26 a 40 anos. Isso demonstra que a variante que estamos considerando com inovadora provavelmente surgiu no português brasileiro há mais de 20 anos, pois nossos falantes de 40 anos já deviam usá-la antes de terem quinze anos, idade em que o ser humano consolida seu repertório linguístico.

## II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA

Algo curioso encontrado nos dados é o pouco uso da quebra de expectativa por falantes idosos, já que essa faixa realizou comparativamente menos tanto *mas* quanto *só que*.

Analisemos por fim nossas variantes em relação ao nível de escolaridade dos falantes, critério distribuído na tabela a seguir:

Distribuição de <i>mas</i> e <i>só que</i> por nível de escolaridade	mas	%	só que	%
Nível 1	176	17.1	3	3.3
Nível 2	486	47.1	55	60.4
Nível 3	370	35.9	33	36.3
Total	1032	100.0	91	100.0

**Quadro 3: Distribuição de *mas* e *só que* em relação ao nível de escolaridade**

Observa-se que a forma perifrástica marcadora da adversatividade ocorre mais entre falantes mais escolarizados, o que confirma nossa hipótese de que *só que* não é uma forma estigmatizada socialmente. Um fator interessante em relação aos dois itens em análise é o fato de eles serem mais produzidos por falantes de escolarização intermediária, isto é, o nível 2. Entretanto sobre isso devemos considerar que falantes dos níveis 2 e 3 possuem em geral a mesma formação acadêmica, diferenciando-se apenas pelo fato de estes trabalharem em sua área de formação superior e aqueles não. Desse modo não poderíamos afirmar que o uso predominante de *mas* no nível 2 em relação ao 3 se deve à escolarização.

Por fim, nota-se que os dados coletados apresentam aspectos interessantes que merecem ser melhor avaliados. Neste artigo, procuramos apenas apontar indícios, o que não nos exime, obviamente, de uma análise mais pormenorizada futuramente.

#### 4. Considerações finais

Para finalizar este artigo, parece-nos pertinente retomar alguns pontos cruciais. Partimos da conceituação de gramaticalização a fim de demonstrar que a perífrase *só que* do português brasileiro é fruto desse processo, que uniu o item *sola* do latim arcaico, posteriormente transformado na partícula denotadora *só*, e a conjunção *que*, sendo esta muito produtiva para formar perífrases conjuncionais em nossa língua. Em seguida, levantamos a hipótese de que há duas formas variantes no português brasileiro para marcar a quebra de expectativa ou adversatividade na

tradição gramatical, quais sejam: *mas*, a forma conservadora, sintética e mais usual, e *só que*, a forma inovadora, perifrástica e menos comum.

Parece-nos adequado dizer que pode haver indícios de mudança linguística nesse caso, uma vez que a variante conservadora é bem mais frequente que a inovadora, portanto, mais passível ao desgaste e à erosão para ceder lugar à forma inovadora. Algo que pode inclusive contribuir para esse processo é o fato de *só que* não ser uma forma estigmatizada socialmente, posto que é livremente utilizada por falantes dos sexos feminino e masculino e por falantes mais escolarizados, fatores que podem acelerar o estabelecimento de *só que* no português brasileiro falado e sua concorrência mais expressiva com *mas*.

#### REFERÊNCIAS BIBLIOGRÁFICAS

BECHARA, Evanildo. *Moderna gramática portuguesa*. 37. ed. Rio de Janeiro: Nova Fronteira, 2009.

CEZARIO, Maria Maura da Conceição; ALONSO, Karen Sampaio Braga. Estudos em gramaticalização: uma homenagem a Mário Eduardo Martelotta. In: RODRIGUES, Violeta Virgínia. (Org.). *Gramaticalização, combinação de cláusulas, conectores*. Rio de Janeiro: UFRJ, 2013.

KRUG, Manfred G. *Emerging English Modals. A Corpus-based Study of Grammaticalization*. New York: Mouton de Gruyter, 2000.

LONGHIN-THOMAZI, Sanderléia Roberta. *A gramaticalização da perífrase conjuncional 'só que'*. 2002. Tese (de doutorado). – Universidade Estadual de Campinas, Campinas.

MARTELOTTA, Mário Eduardo. Unidirecionalidade na Gramaticalização. In: VITRAL, Lorenzo Teixeira; COELHO, Sueli. (Orgs.). *Estudos de gramaticalização em português: metodologias e aplicações*. Campinas: Mercado de Letras, 2010.

MARTELOTTA, Mário Eduardo; VOTRE, Sebastião Josué; CEZARIO, Maria Maura da Conceição. (Orgs.). *Gramaticalização no português do Brasil: uma abordagem funcional*. Rio de Janeiro. Universidade Federal do Rio de Janeiro: Grupo de Estudos Discurso & Gramática, 1996.

NEVES, Maria Helena de Moura. *A gramática funcional*. São Paulo: Martins Fontes, 1997.

II CONGRESSO INTERNACIONAL DE LINGUÍSTICA E FILOLOGIA  
XX CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA

PAIVA, Maria da Conceição de; DUARTE, Maria Eugênia Lamoglia. A mudança linguística: tempo real e tempo aparente. In: \_\_\_\_; \_\_\_\_\_. (Orgs.). *Mudança linguística em tempo real*. Rio de Janeiro: Contra Capa, 2003.

\_\_\_\_\_; \_\_\_\_\_. Quarenta anos depois: a herança de um programa na sociolinguística brasileira. In: WEINREICH, Uriel; LABOV, William; HERZOG, Marvin I. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola, 2006.

POGGIO, Rosauta Maria Galvão Fagundes. *Processos de gramaticalização de preposições do latim ao português: uma abordagem funcionalista*. Salvador: Edufba, 2002.

RASO, Tommaso; MELLO, Heliana. (Orgs.). *C-oral-Brasil I: corpus de referência do português brasileiro falado informal*. Belo Horizonte: UFMG, 2012.

SAID ALI, Manuel. *Gramática histórica da língua portuguesa*. 6. ed. São Paulo: Melhoramentos, 1966.

SOUSA, Lilian; VITRAL, Lorenzo Teixeira. Formas reduzidas do item 'não' no português brasileiro. In: VITRAL, Lorenzo Teixeira; COELHO, Sueli. (Orgs.). *Estudos de processos de gramaticalização em português: metodologias e aplicações*. Campinas: Mercado das Letras, 2010.

VITRAL, Lorenzo Teixeira. *O papel da frequência na identificação de processos de gramaticalização*. Scripta, Belo Horizonte: PucMinas. 2006.

\_\_\_\_\_. A forma cê e a noção de gramaticalização. In: \_\_\_\_; RAMOS, Jânia Martins. *Gramaticalização: uma abordagem formal*. Rio de Janeiro: Tempo Brasileiro; Belo Horizonte: FALÉ/UFMG, 2006.

\_\_\_\_\_; VIEGAS, Maria do Carmo; OLIVEIRA, Alan Jardel de. Inovação *versus* mudança: a interseção gramaticalização / teoria da variação e mudança. In: VITRAL, Lorenzo Teixeira; COELHO, Sueli. (Orgs.). *Estudos de gramaticalização em português: metodologias e aplicações*. Campinas: Mercado das Letras, 2010.