

Gramática, competição e padrões de variação: casos com *ter/haver* e *de/em* no português brasileiro

Grammar, competition, and variation patterns: *ter/haver* and *de/em* in Brazilian Portuguese

Juanito Avelar
Universidade Estadual de Campinas

Abstract

I analyze two instances of linguistic variation in Brazilian Portuguese: the occurrence of *ter* 'to have' or *haver* 'there is/are' in existential constructions, and the use of *de* 'of' or *em* 'in' in prepositional phrases with adjectival function. Exploring Kato's (2005) proposal, I argue that *ter/haver* variation is a reflection of the competition between two grammars, with sociolinguistic motivations, while *de/em* variation results from morphosyntactic procedures within an only grammar.

Keywords

locative prepositions, verbos existenciais, linguistic change, sociolinguistics, generative syntax.

Resumo

Neste trabalho, analiso dois casos de variação lingüística no português brasileiro: *ter/haver* em construções existenciais e *de/em* em sintagmas preposicionados locativos adnominais. Explorando as propostas de Chomsky (1981), Kroch (1994) e Kato (2005), argumento que a variação entre *ter* e *haver* é reflexo da competição

entre duas gramáticas, com motivações sociolingüísticas, enquanto a variação entre *de* e *em* resulta de procedimentos morfossintáticos internos a uma mesma gramática.

Palavras-chave

preposições locativas, verbos existenciais, mudança lingüística, sociolingüística, sintaxe gerativa.

INTRODUÇÃO

Neste trabalho, observo dois casos de variação no português brasileiro (PB), trazendo para reflexão as seguintes questões: (a) quão reveladores de propriedades de uma gramática internalizada podem ser os padrões de distribuição, na fala e na escrita, envolvendo formas em variação, (b) em que medida a observação desses padrões pode contribuir para precisar a variação como um efeito proveniente de uma mesma gramática ou de gramáticas distintas e (c) que lição teórica a compreensão desses padrões pode trazer para a análise de estágios da língua dos quais dispomos apenas de fontes escritas, sem acesso à intuição dos falantes.¹ Os fatos que irei focalizar são a alternância de *ter* e *haver* em construções existenciais e das preposições *em* e *de* em constituintes preposicionadas adnominais, respectivamente como em (1) e (2) a seguir. Com base em dados de língua oral e escrita extraídos de documentos produzidos entre 1980 e 2005, bem como explorando juízos de (a) gramaticalidade, vou sugerir que a variação entre *ter* e *haver* é desencadeada pela ‘alimentação’ da **gramática periférica** no processo de escolarização (em oposição à **gramática nuclear**, construída no processo natural de aquisição da linguagem, nos termos de Chomsky, 1981 e Kato, 2005); diferentemente, a alternância entre *de* e *em* será tratada como resultante de procedimentos morfossintáticos distintos, mas internos à gramática nuclear.

- (1) a. *Tem* muitas praias bonitas no Rio de Janeiro.
b. *Há* muitas praias bonitas no Rio de Janeiro.
- (2) a. Todos os livros *na* biblioteca podem ser emprestados.
b. Todos os livros *da* biblioteca podem ser emprestados.

Parte das reflexões será dedicada a saber se seria possível chegar às mesmas conclusões sobre *ter/haver* e *de/em* se nos limitássemos às fontes

escritas, numa situação similar à vivenciada pelos pesquisadores que adotam pressupostos chomskianos. Com acesso apenas à língua-E (CHOMSKY, 1986), é impossível chegar a conclusões confiáveis sobre mudanças em língua-I sem uma metodologia que permita ir além da superfície da escrita. Pelo menos com relação a *ter/haver* e *de/em*, defendo que, conciliando a noção de *blocking effect* e conceitos como os de “competição de gramáticas”, no sentido de Kroch (1994), o investigador limitado à escrita poderia chegar às mesmas conclusões de um analista com acesso à intuição dos falantes.

O trabalho vem dividido da seguinte forma: na seção 1, apresento considerações gerais sobre a correlação entre variação, mudança e gramática; nas seções 2, 3 e 4, abordo os dados levantados na fala e na escrita, auxiliado por testes de aceitabilidade, bem como exploro as noções de gramática nuclear e gramática periférica para interpretar os padrões de distribuição obtidos; finalmente, na seção 5, estendo o debate à situação em que o investigador não teria à sua disposição acesso algum à intuição dos falantes ou a dados provenientes da linguagem falada.

1. CONSIDERAÇÕES SOBRE GRAMÁTICA, VARIAÇÃO E MUDANÇA

Um ponto nevrálgico nos estudos diacrônicos de base gerativista diz respeito à real importância que se deve atribuir às investigações sobre o passado. O gerativista historiador sempre é indagado acerca das vantagens que poderiam ser apontadas na observação de mudanças lingüísticas, com vistas, por exemplo, a decifrar o que existe na mente/cérebro dos humanos e lhes permite falar ou compreender as expressões da sua língua. Indagações desse tipo se justificam, em princípio, pelo fato de que bastaria aos analistas circunscrever suas investigações às gramáticas de indivíduos vivos para decifrar as propriedades da Gramática Universal. No estágio atual das investigações sobre a linguagem, parece não haver razão para crer que mudanças lingüísticas estejam atreladas a evoluções do ‘órgão mental’ da linguagem ou a modificações de propriedades da Gramática Universal. Só faria sentido, assim, implementar investigações diacrônicas se pudéssemos retornar a um ponto da evolução do *H. sapiens* em que este ainda não estivesse provido da linguagem tal qual a conhecemos hoje. Ou, ainda, se pudéssemos nos defrontar com a linguagem que pertencia, por exemplo, aos *H. neanderthalensis*, irmãos bastante próximos dos sapiens na evolução,

para detectarmos *se e em que termos* a linguagem destes era qualitativamente diferente da nossa (ver PINKER, 2002 e TATTERSALL, 2003). Desta forma, a rentabilidade teórica no retorno ao passado para detectar propriedades da linguagem, fora de uma perspectiva biológico-evolutiva, seria questionável, dado que a ferramenta inata subjacente à aquisição de uma língua em estágios anteriores deve ser da mesma natureza que a ferramenta identificada no estágio atual.

Às discussões dessa ordem podemos somar uma questão metodológica: o instrumento de análise do gerativista é a intuição do falante, cujo alcance é impossível se estivermos lidando com o passado. Sabemos que as fontes escritas não podem ser tomadas como um resultado direto da gramática naturalmente internalizada de quem as produziu. Por *gramática naturalmente internalizada*, estou entendendo o que Chomsky (1981) classifica como **gramática nuclear**, em oposição a uma **gramática periférica**. Seguindo os desdobramentos propostos por Kato (2005) em torno desses conceitos, uma gramática periférica “pode abrigar fenômenos de empréstimos, resíduos de mudança, invenções, de forma que indivíduos da mesma comunidade podem ou não apresentar esses fenômenos de forma marginal” (p. 3), contrariamente à gramática nuclear, resultante da aquisição natural da língua. Retornando à questão de uma maior ou menor transparência das fontes escritas para fatos da gramática internalizada, um texto desenvolvido por um falante do português brasileiro contemporâneo pode apresentar recursos que não fazem parte da sua gramática nuclear, mas que estão presentes em sua gramática periférica: ênclises, concordância do verbo com sujeito posposto, profusão de passivas sintéticas, alta frequência de sujeitos ocultos, inversão do sujeito em construções interrogativas, etc. Da mesma forma, estratégias recorrentes na fala, que compõem um reflexo mais claro da gramática nuclear, porém ainda rejeitadas no padrão culto, são geralmente ocultadas na escrita: constância de termos em posição de tópico inicial e tópico final, paradigma flexional empobrecido, preenchimento recorrente da posição de sujeito, ergativização e/ou detematização de verbos transitivos, próclise generalizada, preferência do uso de *a gente* ao uso de *nós*, ausência da marca de plural em substantivos, etc. Se o conservadorismo que caracteriza a escrita for válido para estágios anteriores, o observado em fontes do passado torna-se então um complicador para atingir a gramática internalizada de quem os produziu.

Os obstáculos empíricos e conceptuais, contudo, estão longe de intimidar o desenvolvimento de abordagens históricas no quadro gerativista. Principalmente após Lightfoot (1979), diversos trabalhos vêm argumentando que a observação de mudanças lingüísticas pode ser tão ou mais reveladora sobre a faculdade da linguagem que investigações centradas exclusivamente numa visão sincrônica. Kroch (1989), por exemplo, assevera que

In historical materials, we find a kind of information which is necessarily absent in synchronic data and which offers the prospect of an important contribution to general linguistics from history; that is, information about the time course of language change. With such process information, we may hope to learn how the grammars of languages change from one state to another over time; and from an understanding of the process by which the change, to learn more about their principles of organization. After all, perturbing a complex system and observing its subsequent evolution is often an excellent way of inferring internal structure. In addition, since the features of any language at a given point in time are the result of a complex interweaving of general principles of language and particular historical developments, knowledge of the historical process by which a language has reached a given state may be important to the proper assignment of responsibility to historical and general factors, and so to the proper formulation of linguistic theory.² (p. 1-2)

Complementando as vantagens apontadas por Kroch, podemos apontar a própria concepção de *variação interlingüística* que está na base da Teoria de Princípios e Parâmetros. À luz da Teoria, certos processos de mudança podem ser interpretados como frutos de uma alteração em um ou mais parâmetros da linguagem. Com algumas diferenças (por exemplo, propostas como as de Lightfoot, 1991, 1999 e Clark & Roberts, 1993), os investigadores defendem que, se o valor do parâmetro deve ser determinado durante a aquisição, uma mudança vai ser operada sempre que um indivíduo marcar pelo menos um desses parâmetros diferentemente da marcação apresentada pela gramática da geração anterior. Nesse quadro, os rumos de uma mudança podem ser reveladores de como os parâmetros se encontram organizados na Gramática Universal, aspecto que normalmente escapa a uma simples

observação de fatos sincrônicos. A questão da metodologia para atingir as propriedades de uma gramática nuclear exclusivamente a partir da escrita pode então ser adequadamente encaminhada se considerarmos que “the past is like the present and general principles derived from the study of living languages in the present will hold of archaic ones as well” (KROCH, 1989, p. 1), o que, em outras palavras, na linha de Labov (1975), nos permite fazer uso do presente para explicar o passado. No caso, assumindo os pressupostos da Teoria de Princípios e Parâmetros, as restrições da Gramática Universal devem ser válidas para qualquer estágio de uma língua; conseqüentemente, o reconhecimento dessas restrições poderá servir de base para conclusões sobre uma gramática naturalmente internalizada em qualquer estágio histórico. Dentro dessa perspectiva, um campo de investigação que se proponha a buscar os meios de ultrapassar a superfície do texto escrito é altamente desafiador. Se os pressupostos chomskianos forem adotados, a indagação primordial desse campo deve ser a de como os fatos identificados na superfície dos documentos podem nos conduzir para além desses próprios fatos, revelando aspectos da gramática nuclear.

Numa agenda com esse fim, um lugar especial deve ser dado à questão da variação lingüística, uma vez que o indício mais comum de uma mudança, emergente entre os documentos, é a ocorrência de formas variantes, com uma sobrepondo-se à outra ao longo do tempo. Ressalte-se que o reconhecimento de padrões de variação é relevante não apenas para estudos de mudança, mas também para as abordagens que procuram dar conta de aspectos pertencentes a um único recorte temporal. Por exemplo, descrever o comportamento de formas em variação pode ser necessário para elucidar se certos aspectos são internos a uma mesma gramática ou resultado da ação de gramáticas distintas, questão crucial para tomadas de decisão dentro de qualquer análise que adote a Teoria de Princípios e Parâmetros (ver, por exemplo, a discussão em Pintzuk *et al.*, 2000).

Este trabalho aborda duas instâncias de variação no português brasileiro contemporâneo – *ter/haver* e *de/em* – trazendo para discussão exatamente a idéia de que o comportamento demonstrado por formas variantes, num dado recorte temporal, pode ser revelador de propriedades internas à gramática nuclear. Vou me deter em dados de língua oral e escrita, bem como me apoiar em juízos de (a)gramaticalidade, para chegar a certas conclusões sobre o estatuto dessas variações dentro da gramática nuclear;

ao mesmo tempo, indago se seria possível chegar às mesmas conclusões se estivéssemos diante apenas de dados da escrita, na mesma situação que a vivenciada pelos historiadores da língua. Em tal condição estaria, por exemplo, um lingüista que, lá pelos idos de 2500, no aniversário de mil anos do Brasil, se interessasse pelo conhecimento de um falante do português brasileiro há cerca de quinhentos anos antes, na passagem do segundo para o terceiro milênio. Por alguma contingência histórica, esse lingüista teria acesso apenas a fontes escritas, como textos jornalísticos, literários e publicitários. Seu acesso à língua-I seria, portanto, indireto, de modo que uma de suas preocupações recairia na busca de um método que lhe possibilitasse adentrar a superfície da língua-E para detectar propriedades da gramática nuclear. O nosso lingüista do futuro terá predileção por formas em variação (talvez, por estar preocupado em detectar mudanças lingüísticas), e daí o seu interesse pelo uso de *ter* e *haver* entre as sentenças existenciais, em contextos como (3) a seguir, e pelo de *de* e *em* como núcleo de constituintes preposicionados adnominais com valor locativo, nos contextos em (4).

- (3) a. A partir de agora no Aeroporto de Congonhas **tem** um espaço exclusivo esperando por você. (Isto É, 28/07/2004, p. 82.)
b. Nunca **houve** tantos miseráveis em carros importados, trabalhando em grandes escritórios (Augusto Cury, p. 16.)
- (4) a. O espaço mais charmoso **de** Campos do Jordão está de volta... (Super Interessante, p. 70.)
b. Começava aí a surgir a noção da velocidade da luz como limite máximo **no** Universo. (Alexandre Cherman, p. 93.)

Retornando ao presente, vou procurar neste trabalho refletir acerca de quais instrumentos o analista pode se valer para apurar a representatividade de fatos identificados em fontes escritas como expressões mais ou menos reveladoras da gramática nuclear. A idéia é aproveitar a possibilidade de recorrermos a testes de aceitabilidade para comprovar imediatamente se os critérios empregados são eficazes no alcance do conhecimento internalizado em torno de cada variação. Como vou defender, o lingüista do futuro poderá chegar às mesmas conclusões que um lingüista da passagem do milênio se considerar que existem princípios gerais subjacentes a uma gramática internalizada, qualquer que seja o recorte temporal selecionado. Acredito

que, investigando fatos de língua escrita produzidos por uma gramática da qual dispomos de intuições confiáveis (que é a nossa situação diante do português brasileiro contemporâneo), é possível tirar conclusões rentáveis acerca de como uma gramática internalizada se reflete na escrita.

Como os dois fenômenos analisados dizem respeito a procedimentos de seleção vocabular, um dos instrumentos analíticos de que o lingüista de 2500 poderá se valer é o *blocking effect* (BEff), que deve ser válido em qualquer língua natural. O BEff atua de modo a impossibilitar que dois itens funcionais com as mesmas condições de realização coexistam numa mesma gramática (ver ARONOFF, 1976; EMBICK, 2005, entre outros). Kroch (1994) discorre largamente sobre as vantagens de considerar a ação desse efeito em processos de variação e mudança, destacando que o BEff não implica que as línguas sempre rejeitem os *doublets* (dois itens funcionalmente idênticos), mas que os *doublets* só vão existir como reflexo de uma “competição instável” entre duas opções gramaticais exclusivas entre si (p. 1-2). Assumindo a validade do BEff, o lingüista do futuro terá então de explicitar se as duas formas funcionais estão ou não ocorrendo sob as mesmas condições. Se Kroch (1994) estiver correto, nosso analista de 2500 precisará ou atestar os sinais de uma competição instável, originada pela sobreposição de gramáticas distintas, ou buscar indícios de que as duas formas funcionais não são *doublets*, mas elementos realizados sob condições intralingüísticas diversas em uma mesma gramática.

2. OS DADOS

Para a implementação da análise, foram levantados 511 dados de fontes escritas (140 construções existenciais e 371 adjuntos adnominais) e 1.010 de língua falada (511 construções existenciais e 499 adjuntos adnominais). Os dados de língua falada foram obtidos de entrevistas concedidas por indivíduos com e sem nível superior, respectivamente do projeto NURC-RJ (www.letras.ufrj.br/nurc-rj), com amostras da década de 1990, e do projeto PEUL (www.letras.ufrj.br/~peul), com amostras da década de 1980. Os inquéritos da amostra foram classificados de acordo com a faixa etária do falante (até 35, 36-55, 56 em diante), com quatro representantes (dois homens e duas mulheres) para cada faixa.⁴

Os dados de língua escrita foram extraídos de textos produzidos entre 2003 e 2005, a saber: textos dos jornais *O Globo* e *Extra*, de 12 e 13/08/2005, ambos do Rio de Janeiro; anúncios da revista *Veja* (13/04/2005 e 22/06/2005), *Isto É* (28/07/2004), *Época* (28/02/2005) e *Super Interessante* (abril de 2005); e os seguintes livros: (a) *Sobre os ombros de gigantes* (divulgação científica na área de física), de Alexandre Cherman, publicado em 2004; (b) *Amor é prosa. Sexo é poesia* (crônicas), de Arnaldo Jabor, publicado em 2004; (c) *Budapeste* (romance), de Chico Buarque, publicado em 2003; (d) *O Zahir* (romance), de Paulo Coelho, publicado em 2005; (e) *As melhores piadas do planeta... e da casseta também!*, do Casseta e Planeta, publicado em 2003; e (f) *Nunca desista de seus sonhos* (auto-ajuda), de Augusto Cury, publicado em 2004.^{5, 6}

Para o armazenamento e processamento dos dados, foi empregado o pacote de programas que integram a versão 2001 do Goldvarb, o que facilitou a depreensão de possíveis condicionamentos em cada variação. Sobre o padrão de distribuição das variantes, os dados levantados revelaram um contraste entre *ter*/*haver* e *de*/*em*: as frequências de *ter* e *haver* são diametralmente opostas nas línguas falada e escrita, enquanto as de *de* e *em* são as mesmas nas duas modalidades. Os gráficos a seguir ilustram o contraste: na língua falada, *ter* é realizado em 87% das construções existenciais levantadas, mas não passa de 14% nos dados da escrita; já entre as preposições, os percentuais se mantêm em torno de 72% para *de* e 28% para *em*, na fala e na escrita.⁷

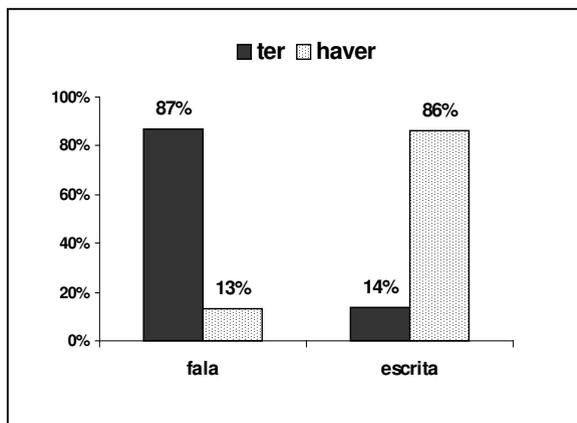


GRÁFICO 1 – Frequências de *ter* e *haver* em contextos existenciais na língua falada e na língua escrita.

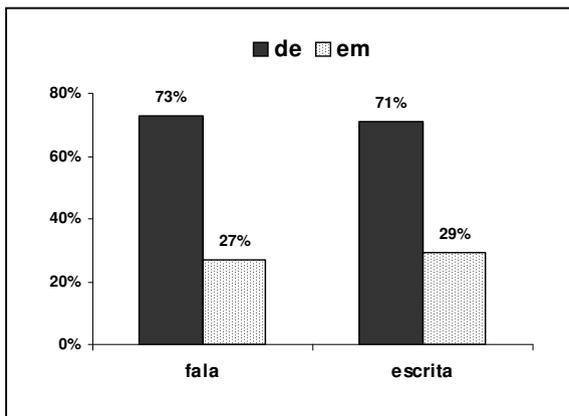


GRÁFICO 2 – Frequência de *de* e *em* em adjuntos como núcleos de adjuntos adnominais locativos na fala e na escrita.

O Goldvarb apontou condicionamentos de natureza extralingüística para *ter/haver*, e intralingüística para *de/em*, tanto na fala como na escrita. Vou abordar esses condicionamentos nas seções a seguir, procurando validar a idéia de que a alternância uniforme entre *de* e *em* e a polarização no comportamento de *ter* e *haver* apontam para estatutos diferenciados de variação, no sentido de que aspectos distintos intrínsecos à língua-I estão entrando em jogo em um e outro caso.

3. A VARIAÇÃO ENTRE *TER* E *HAYER*

Além do contraste observado entre o comportamento na fala e na escrita, a variação *ter/haver* mostra sensibilidade à faixa etária e ao nível de escolarização do falante, bem como ao tipo textual em que a sentença existencial é realizada. Observando os condicionamentos na fala, os dados revelam que, quanto mais velho o falante, menor o percentual de uso de *ter*, tanto na amostra do NURC quanto na do PEUL, como indicado no gráfico 3. Contudo, qualquer que seja a faixa etária, o percentual de *ter* é amplamente maior que o de *haver*. Quanto ao nível de escolarização, *ter* ocorre numa frequência maior entre os indivíduos sem nível superior, integrantes do PEUL. Apenas na primeira faixa etária é que o percentual de uso de *ter*, quase categórico, é praticamente o mesmo entre os falantes dos dois níveis, ficando

em torno de 97%. Já entre as fontes da escrita, os dados extraídos de piadas do Casseta e Planeta são os únicos que apontam para uma frequência maior de *ter* sobre *haver*, com 60% para o primeiro, como indicado no gráfico 4. Nos textos dos anúncios, de Chico Buarque e do Jornal *Extra*, as frequências de *ter* chegam, respectivamente, a 47%, 15% e 14%. Entre os demais documentos, não foram identificados casos de construções existenciais com *ter*.

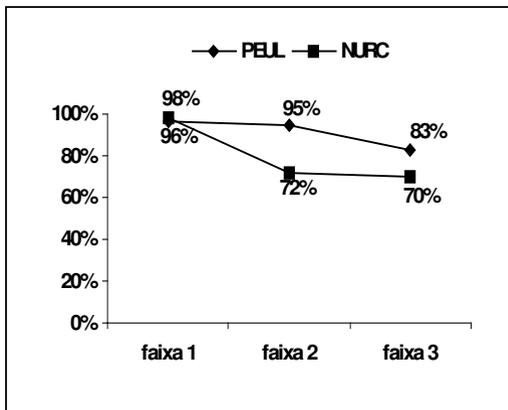


GRÁFICO 3: Percentual de uso de *ter* contra o de *haver* em construções existenciais na fala, considerando as três faixas etárias, entre os indivíduos do NURC e do PEUL.

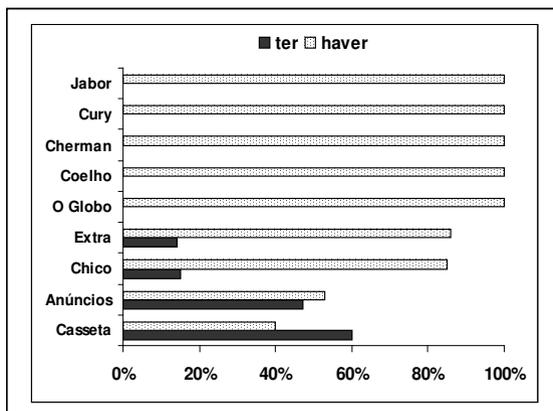


GRÁFICO 4: Frequências de *ter* e *haver* em contextos existenciais entre os dados extraídos de fontes escritas.

O Goldvarb não apontou condicionamento algum de ordem intralingüística para *ter/haver*. Se pensarmos então em termos da oposição *gramática nuclear vs gramática periférica*, é plausível considerar que são elementos da periferia os que estão entrando em jogo para determinar essa variação. Os fatores extralingüísticos apontados dão sustento à idéia de que *haver* é uma variante de prestígio, sendo a forma preferida na língua escrita, muito embora não haja qualquer estigma para o uso de *ter* na língua falada (pelo menos, no meu dialeto). A distribuição de *ter* nos dados da escrita indicia que sua seleção, entre os textos, é condicionada pela necessidade de reproduzir elementos comuns da oralidade. No caso dos textos do Casseta e Planeta, com exemplos em (5) a seguir, o uso de *ter* é comum na representação dos diálogos que envolvem as personagens de uma piada. O verbo ocorre, inclusive, na representação da fala de um personagem português, apresentada no exemplo em (5d). Cabe lembrar que, no português europeu, *ter* é empregado apenas como verbo possessivo, não sendo admitido com valor existencial. Já *haver* ocorre nos três casos de construções existenciais identificadas fora de um diálogo, como no exemplo em (5e); dentro de diálogos, *haver* ocorreu apenas em um dos seis casos levantados, apontado em (5f).

- (5) a. - ...acho que um de nós devia subir no alto daquela palmeira para ver se **tem** algum navio que possa nos resgatar. (Casseta e Planeta, p. 13.)
- b. - Me contaram que aqui no inferno **tem** um telefone. (Casseta e Planeta, p. 32.)
- c. - ...**tinha** uma parte do homem que não encaixava em lugar nenhum! (Casseta e Planeta, p. 105.)
- d. - ...não **tem** problema nenhum, ó pá! (Casseta e Planeta, p. 14.)
- e. [Deus] decidiu mandar um *e-mail* aos 5% das pessoas boas que **havia** no mundo. (Casseta e Planeta, p. 87.)
- f. - ...se o senhor tivesse avisado, **haveria** manifestação, repúdio popular, passeata, repressão, pancadaria e morreria muito mais gente. (Casseta e Planeta, p. 28.)

Entre os anúncios, a frequência significativa de *ter* (embora menor que a de *haver*) pode ser desencadeada pela necessidade de uma aproximação maior com o leitor, parecendo consistir numa estratégia de persuasão por meio

do emprego de marcas recorrentes na oralidade. Casos de anúncios com os dois verbos estão exemplificados em (6) a seguir. Ressalte-se que, entre os anúncios, diferentemente do observado entre as piadas, não foi possível detectar o que ora leva ao uso de *ter*, ora ao uso de *haver*, com a escolha de uma ou outra forma, à primeira vista, parecendo se dar aleatoriamente.

- (6) a. Quando você sentir que **tem** alguém fazendo mais pelas pessoas, repare: a BASF está por perto. (Isto É, 28/07/2004, p. 5.)
- b. Não **tem** idade certa para ser feminina. (Veja, 13/04/2005, p. 39.)
- c. **Tem** também a Bohemia Escura, que mantém equilíbrio perfeito entre o doce e o amargo. (Veja, 13/04/2005, contracapa.)
- d. Sem água não **há** trabalho, não **há** saúde, não **há** vida. (Isto é, 28/07/2004, p. 63.)
- e. Para que não **haja** nenhuma dúvida, a partir deste mês você vai encontrar a sua Qualy com um alerta: “zero de gorduras trans”. (Veja, 13/04/2005, p. 27.)
- f. Rescenda a Paixão que **há** em você. (Época, 28/02/2005, p. 59.)

No texto de Chico, os dois casos (entre treze) de construções existenciais com *ter* estão apontados em (7a) a seguir. No Jornal Extra, dos quatro casos identificados com *ter* (entre 28), dois aparecem na transcrição da fala de pessoas entrevistadas (7b-c), um ocorre numa carta de leitor (7d) e um, na seção de programação cultural (7e). Este último dado, portanto, é o único que pode ser apontado como propriamente originado dentro da redação do jornal. No que diz respeito ao texto de Chico Buarque, o caso de *ter* surge na elaboração de um diálogo entre o narrador e uma personagem, o que novamente evidencia ser a necessidade de inserir marcas de oralidade o principal fator que condiciona o uso de *ter* na língua escrita.

- (7) a. Em Budapeste ela não conhecia ninguém, **tem** loja de departamentos em Budapeste? Não sei, deve **ter** confeitarias, excelentes museus. (Chico Buarque, p. 42.)
- b. – Mais de 30 mil vans estão rodando pelo estado entre legalizadas e não legalizadas. **Tem** até carro de passeio e moto fazendo transporte clandestinamente – acrescenta Antônio Tristão, do sindicato dos rodoviários. (Extra, 12/08/2005, Primeiro Caderno, p. 16.)

- c. – *Tem* carne sobrando no mercado interno. A tendência é de que o preço do produto se mantenha menor até o fim do ano – afirmou. (Extra, 13/08/2005, Primeiro Caderno, p. 11.)
- d. Na praia da brisa, em Guaratiba, *tem* muitos animais abandonados, na sua maioria cães e gatos. Eles não são vacinados e estão doentes. (Extra, 13/08/2005, Primeiro Caderno/Serviço, p. 13.)
- e. Hoje, *tem* tributo a Bob Marley, no Nectar, em Vargem Grande. (Extra, 13/08/2005, Primeiro Caderno/Geral, p. 4.)

Consideremos agora como abarcar a variação entre *ter* e *haver* no interior de uma descrição sobre o conhecimento internalizado de um falante do português brasileiro contemporâneo. Para capturar o comportamento das existenciais, o analista terá de elaborar um quadro formal que explicita certas propriedades, tais como o caráter impessoal dessas construções, efeitos de definitude, necessidade de associar o DP complemento a um constituinte predicativo (a chamada *coda*, no sentido de Milsark, 1974), ocorrência do argumento nominal em posição pós-verbal, etc. Juntamente com essas propriedades, o analista poderá achar necessário considerar a variação entre *ter* e *haver*. É o estatuto dessa variação como resultado ou não de algum procedimento interno à gramática nuclear do falante que será o ponto-chave da discussão.

Podemos tomar como ponto de partida a inexistência de condicionamentos internos para a variação. No que tange a procedimentos meramente gramaticais, esse aspecto parece evidenciar que a seleção de *ter* ou *haver* é aleatória. Esse caráter de aleatoriedade, contudo, é altamente suspeito, uma vez que o percentual de *ter* é drasticamente maior que o de *haver* na língua falada. Se a língua falada pode ser tomada como um acesso mais direto à gramática internalizada que a língua escrita, então devemos estar diante de uma situação em que *ter*, nessa gramática, exerce um papel diferenciado do exercido por *haver*. Para precisar essa diferenciação, vamos ter de recorrer aos chamados *dados negativos*, inexistentes na língua falada, que mostram que certos contextos existenciais licenciam o uso de *ter*, mas não o de *haver*. Contextos desse tipo foram discutidos em Avelar 2005a, onde sugiro que *haver*, no português brasileiro contemporâneo, não pode mais ser tratado como um verbo funcional (considerando a oposição entre categoria funcional e categoria substantiva, mais especificamente nos termos da Morfologia Distribuída (ver

HARLEY & NOYER, 2003 e EMBICK & NOYER, 2001)). Atentemos, por exemplo, para os casos em (8)-(9) a seguir: em (8), temos contextos que admitem tanto *ter* quanto *haver*; em (9), diferentemente, a inserção de *haver* torna a construção marginal ou inaceitável (pelo menos na minha intuição).

- (8) a. Sempre vai *ter* / *haver* políticos querendo tirar proveito dos eleitores.
 b. *Tinha* / *Havia* muitos livros bons na biblioteca.
 c. *Teve* / *Houve* uma confusão danada na hora da festa.
 d. “lá *tinha* (*havia*) várias faculdades” (Nurc-RJ 015/90)
 e. “eu talvez pudesse me interessar por um brinquedo de outro tipo... mas não *havia* (*tinha*)essa possibilidade...” (Nurc-RJ 023/90)
- (9) a. Sempre *tem* / ??*há* mulher desfilando pelada na Marquês de Sapucaí.
 b. *Tem* / **Há* gente na festa que odeia cheiro de cigarro.
 c. *Teve* / ??*Houve* muitos docinhos na festa que a Maria deu.
 d. A: *Tem* / **Há* pão? B: *Tem* / **Há*.
 e. “lá ... por exemplo, *tem* (**há*) jardim, tá, com lago em volta” (Nurc-RJ 003/90)
 f. “(o) respeito com pessoas mais velhas, sabe? isso *tem* (**há*) muito no Japão” (Nurc-RJ 012/90)

Um outro contexto que vale ser citado diz respeito à solicitação de um perueiro, que presenciei certa vez em Campinas, no interior de São Paulo, para que os passageiros da perua indicassem previamente o ponto do percurso em que iriam descer. A solicitação é reproduzida a seguir. É impensável uma situação em que o perueiro e o passageiro usassem *haver* em lugar de *ter* nas construções existenciais que ambos decidem empregar. Dados desse tipo evidenciam que, ao lado de uma absoluta superioridade quantitativa sobre *haver* na língua falada, *ter* também demonstra um diferencial qualitativo, ocorrendo em contextos corriqueiros de uso da língua nos quais o existencial clássico é dificilmente licenciado.

(10) Na peruca:

PERUEIRO: Tem (*Há) Avenida Brasil? Tem (*Há) Taquaral?
Tem (*Há) Tapetão? Tem (*Há) Barão Geraldo? Tem (*Há)
Hospital?

PASSAGEIRO: Hospital, tem eu. / * Hospital, há eu.

Partindo de dados desse tipo, argumento em Avelar (2005a) que *ter* e *haver* não estão em pé de igualdade. A oposição entre os dois verbos pode ser formalmente capturada pela idéia de que *haver*, no português brasileiro contemporâneo, deixou de ser uma categoria funcional, tendo sido reanalisado como um verbo existencial substantivo, incluindo-se na família de itens como *acontecer*, *ocorrer*, *surgir*, *existir*, etc. Ou seja, a rigor, não existe variação entre *ter* e *haver* na gramática internalizada de um falante do português brasileiro, mas o uso de *haver* nos mesmos termos em que se usam outros verbos apresentacionais da língua, licenciados normalmente em contextos de interpretação existencial. Sobre a distribuição desses verbos apresentacionais, um levantamento que apresento em Avelar (2005a) permite atestar que verbos como *acontecer* e *existir* são mais freqüentes, no falar carioca, que *haver*. Como observado nos casos em (11) a seguir, esses itens são normalmente licenciados em muitos contextos nos quais podemos empregar o verbo *ter*. Isso implica que, se assumirmos uma variação entre *ter* e *haver* de igual para igual, vamos ter também de incluir *acontecer* e *existir* no pacote das formas variantes, idéia que, até onde eu saiba, não aparece entre os estudos sobre as construções existenciais do português brasileiro.

- (11) a. *Teve/aconteceu* um acidente horrível na estrada.
b. Sempre *tem/acontece* alguma confusão nas festinhas da universidade.
c. Não *tem/existe* vida em outros planetas do sistema solar.
d. Nunca *teve/existiu(ram)* partidos políticos totalmente confiáveis no Brasil.

Se essa linha de raciocínio for válida, a questão de como abarcar a variação entre *ter* e *haver* num quadro formal que busque capturar as propriedades internas de sentenças existenciais parte de uma premissa equivocada. O modelo para capturar fatos da língua-I do falante no que diz

respeito às construções existenciais deve, de outra forma, se ater ao seguinte ponto: existem, de um lado, construções existenciais canônicas, construídas com o verbo *ter*; de outro lado, aparecem construções existenciais mais gerais, de uso normalmente apresentacional, com verbos como *haver*, *aparecer*, *acontecer*, *existir*, *surgir*, etc. Assim, não estamos diante de uma variação a ser capturada como um fato da gramática interna do falante, a sua gramática nuclear, mas simplesmente de um padrão frásico do português contemporâneo que elege como a sua forma verbal prototípica o verbo *ter*. No âmbito da gramática naturalmente internalizada, portanto, não existe variação entre dois verbos existenciais no português brasileiro, mas entre um padrão canônico de gerar uma sentença existencial, para o qual se recorre a um verbo funcional, e outros padrões diferenciados, com valores semântico-pragmáticos diversos, que se valem de verbos não-funcionais.

Se adotar essa visão, o analista estará livre da necessidade de explicar uma suposta falha de ação do *blocking effect* (BEff) entre as existenciais. O efeito somente permitiria uma variação entre *ter* e *haver* se as formas tivessem condições de seleção diferenciadas, determinadas por fatos internos à gramática (por exemplo, se uma fosse o verbo selecionado em existenciais no presente e no pretérito imperfeito, e outra, nos demais tempos). Não sendo este o caso, o BEff atua no sentido de eliminar uma das formas, por exemplo, alterando a condição funcional de uma delas. Poderíamos, nesses termos, arregimentar esforços para validar uma das três situações: ou (a) o BEff é uma falácia, sendo uma condição de estatuto teórico suspeito, ou (b) o BEff não atua no português brasileiro, com *ter* e *haver* variando na língua como dois verbos existenciais funcionais, ou (c) *ter* e *haver* não estão em pé-de-igualdade quanto ao caráter funcional, com o primeiro sendo o existencial prototípico, e o segundo, um existencial substantivo, de natureza não-funcional.

A situação em (c) corresponde à que defendo em Avelar (2005a). Consideremos brevemente, para evidenciar o contraste, as duas outras situações, que põem em xeque a atuação do BEff. A situação em (a) deve ser descartada, dada a existência de fatos empíricos significativos entre as línguas naturais, relatados na literatura (ver EMBICK, 2005 e KROCH, 1994), que dão sustento à validade do efeito. No que diz respeito à resposta (b), ela é particularmente problemática para dar conta de uma suposta variação entre verbos existenciais, dado que, também empiricamente, não há registros de

línguas (até onde eu saiba) em que dois verbos existenciais comprovadamente funcionais variem livremente. Num levantamento apresentado em Clark (1978), abarcando verbos existenciais, possessivos, copulativos e locativos em diversas línguas, o número de línguas que exibem duas formas possíveis para o verbo existencial é reduzidíssimo (duas línguas, num total de 34). Uma dessas línguas é o japonês, que emprega as formas *aru* e *iru* como verbos de expressão existencial. A ocorrência dessas formas, porém, não é aleatória, com *aru* sendo licenciado sempre que o argumento do verbo for inanimado, e *iru*, sempre que for animado (ver KISHIMOTO, 2000), situação que é compatível com o BEff.

Na própria história do português, encontramos evidências fortes para confirmar a validade do efeito. No português arcaico, *ser* era o item empregado para expressar cópula, locação e existência, e *haver*, a forma prototípica para a posse (MATTOS e SILVA, 1997). Por razões ainda não elucidadas, *haver* passa a ser usado como verbo existencial, e o que se nota entre os documentos não é uma convivência pacífica de *ser* e *haver* em tais contextos, mas o desaparecimento progressivo do existencial copular, evidenciando que as duas formas não puderam conviver lado a lado. Da mesma forma, *ser* foi “expulso” por *estar* dos contextos copulativos de atribuição transitória e da expressão de locação, assim como *haver* não pôde conviver ao lado de *ter* como possessivo. No passado da língua, portanto, identificamos fatos que mostram claramente a aplicabilidade do BEff; soaria estranho que a condição deixasse de ter efeito apenas no português brasileiro contemporâneo, mais especificamente sobre um tipo frásico determinado.

A hipótese em (c), então, é a que melhor indica a condição de *ter* e *haver*: as duas formas não estão competindo pelo posto de verbo existencial como duas categorias funcionais da língua. O que temos, de um lado, é o verbo *ter* “reinando” absoluto como existencial e possessivo; de outro, uma série de verbos existenciais, sem uso funcional, que apontam para interpretações específicas dentro da família de verbos com função apresentacional, entre os quais se inclui o item *haver*.

O que dizer, porém, do alto percentual de *haver* na escrita? Sua ampla frequência não estaria atestando que *haver* ainda detém, na gramática da língua, um estatuto funcional, tal como o que conferimos para *ter*? Seria possível afirmar que a escrita elege categorias funcionais diversas àquelas que

observamos na fala? Considero que um encaminhamento adequado para tais questões passa pela discussão desenvolvida por Kato (2005) sobre os domínios da gramática nuclear e da gramática periférica. Como referido na seção 1, a gramática periférica armazena resíduos de mudança, empréstimos, inovações lingüísticas, etc., com os indivíduos de uma mesma comunidade lingüística diferindo na manifestação ou não desses aspectos. Kato (2005) defende que uma das fontes provedoras dessa periferia é o processo de escolarização, que desencadeia o aprendizado da língua escrita. No caso de falantes do português brasileiro, a aprendizagem da língua escrita se daria nos mesmos moldes em que se verifica o aprendizado de uma segunda língua, com uma e outra sendo adquiridas por meio de aprendizagens “socialmente motivadas e não biologicamente marcada” (p. 9), diferentemente da gramática nuclear. Ambas teriam início depois de uma idade crítica para a aquisição, o que abre espaço para a emergência de diferenças individuais marcantes.

O uso de *haver* como existencial canônico não consiste, nessa perspectiva, num reflexo de procedimentos internos à gramática nuclear, mas do provimento da gramática periférica por elementos de prestígio na língua escrita. Se atentarmos para que, no Brasil, o aprendizado da língua escrita tende a ter como alvo ou estágios anteriores da língua ou a norma do português europeu (muito embora o resultado final não seja nem uma coisa nem outra, como destacado em Kato, 2005), fica fácil imaginar a razão da supremacia de *haver* e da tendência de supressão de *ter* na língua escrita. Se essa visão estiver correta, *haver* entra no acervo vocabular dos falantes, comportando-se quantitativamente como um verbo existencial funcional da escrita, nas mesmas condições em que entraria um verbo funcional de uma L2 (digamos, o *there is/are* do inglês). Provavelmente daí resulta o fato de a sua freqüência entre os falantes com nível superior ser maior que a freqüência entre os demais (ver o gráfico 3). Daí também o trânsito maior de *ter* entre as piadas e anúncios, dado que os produtores desses gêneros devem ter consciência de que, para alcançar os efeitos desejados, precisam recorrer a um padrão que reproduza marcas da oralidade.

E a que conclusão poderia chegar o lingüista de 2500, sem acesso a dados de fala, e muito menos à intuição dos falantes que produziram os textos que tem em mãos? Pude argumentar em favor de uma variação na qual entram em jogo elementos da periferia, e não estratégias internas à gramática

nuclear, porque estou lidando com um estágio da língua sobre o qual é possível esboçar hipóteses a partir de intuições satisfatórias. Mas o mesmo não será possível ao lingüista do futuro, que terá de lidar com “maus dados” para descrever parte do conhecimento de um falante do português brasileiro da virada do milênio. Retorno a essa questão na seção 5, quando abordo conjuntamente os casos de *ter/haver* e *delem*.

4. A VARIAÇÃO ENTRE DE E EM

Como já apontado, os padrões de distribuição para *de* e *em* como núcleos de adjuntos adnominais não mostram diferenças entre a fala e a escrita: nas duas amostras, as freqüências de *de/em* ficam em torno de 72%/28%, como indicado anteriormente no gráfico 2. Considerando primeiramente os dados de fala, o Goldvarb não detectou condicionamentos extralingüísticos, não havendo diferença significativa entre os indivíduos das três faixas etárias com e sem nível superior, como ilustrado no gráfico 5 adiante. Esse padrão uniforme não surpreende, uma vez que não existe prestígio ou estigma de uma forma sobre a outra que justifique uma ocorrência maior ou menor de um dos itens entre indivíduos de um segmento específico. Além disso, não há notícias na literatura de qualquer mudança recente envolvendo as duas posições como núcleos de adjuntos adnominais locativos, tal como o que sabemos acerca de *ter* e *haver* na história do português brasileiro.

Também entre os documentos da língua escrita não foram detectadas diferenças significativas de valores; o percentual de ocorrência de *de* mantém-se sempre mais elevado que o de *em*, como observamos no gráfico 6 a seguir. A freqüência de *de* frente à de *em* oscilou entre 63%, no texto de Chico Buarque, e 80%, na obra de auto-ajuda escrita por Augusto Cury. Essa oscilação não pode ser atribuída a fator de ordem extralingüística, parecendo antes estar radicada, como veremos, na recorrência de um “ambiente intralingüístico” que favorece a ocorrência de *em* e desfavorece a de *de*.

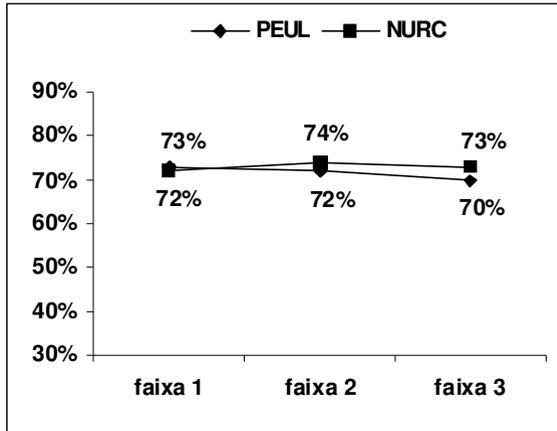


GRÁFICO 5: Frequência de *de*, com relação a *em*, como núcleo de adjuntos adnominais locativos na língua falada, considerando as três faixas etárias, entre os indivíduos com (NURC) e sem (PEUL) nível superior.

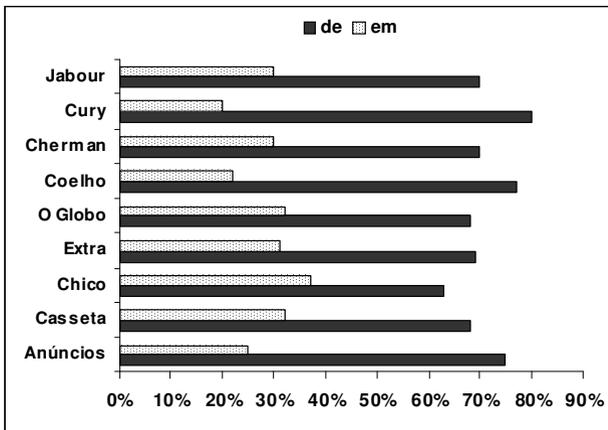


GRÁFICO 6: Frequências de *de* e *em* como núcleos de adjuntos adnominais locativos entre os dados extraídos de fontes escritas.

O Goldvarb identificou a especificidade semântica do argumento da preposição como sendo o fator intralingüístico relevante para determinar o comportamento. Quanto a essa característica, os dados foram divididos em cinco tipos: *material animado*, como em (12); *material inanimado*, como em (13); *localização espacial*, como em (14); *localização temporal*, como em (15); *evento*, como em (16); e *imaterial*, como em (17).

- (12) a. ...temos todos os problemas normais **de um casal**. (Paulo Coelho, p. 18.)
b. os sonhos ... renovam as forças **do ansioso**... (Augusto Cury, p. 12.)
- (13) a. Usando uma chave amarrada ao fio **da pipa**... (Alexandre Cherman, p. 76.)
b. ...perda de potência **no motor**... (Isto É, 28/07/2004, p. 35.)
- (14) a. Os dois foram vistos pela última vez em um café **na rua Faubourg Saint-Honoré**... (Paulo Coelho, p. 15.)
b. ...tinham professoras **no colégio** que eram, assim temidas né, pela disciplina... (Nurc-RJ, Inq. 001/masc/faixa 1.)
- (15) a. A grande obra **dos primórdios da eletricidade e do magnetismo** só seria escrita em 1600... (Alexandre Cherman, p. 72.)
b. Maxwell publicou um artigo ... que lhe valeu o Prêmio Adams **de 1857**. (Alexandre Cherman, p. 80.)
- (16) a. ...se eu lavar os pratos **do almoço**... (Nurc-RJ, Inq. 003/fem/faixa 1.)
b. Prova de natação **na paraolimpíada**. (Casseta e Planeta, p. 26.)
- (17) a. Escravos **de uma vida** que não tinham escolhido, mas que haviam decidido viver. (Paulo Coelho, p. 21.)
b. Ambos os estímulos foram registrados no mesmo lócus **do inconsciente** (Augusto Cury, p. 13.)

A ocorrência de *em* é condicionada pelos contextos que expressam localização espacial e evento, casos em que a frequência da preposição chega, respectivamente, a 56% e 48%; já a forma *de* têm trânsito corrente entre todos

os tipos, chegando a 88% entre os casos em que o argumento da preposição é material inanimado, como ilustrado no Gráfico 7 a seguir.

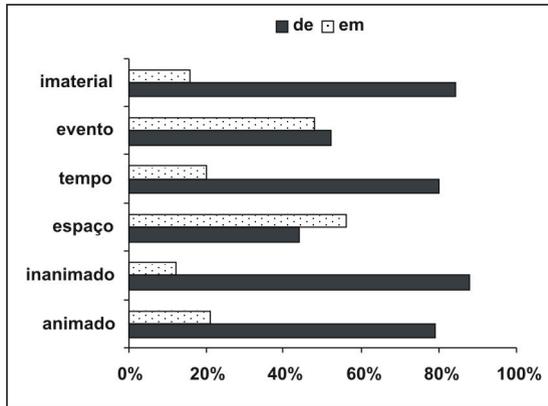


GRÁFICO 7: Frequências de *de* e *em* entre os tipos de argumento da preposição quanto à especificidade semântica, considerando conjuntamente dados da língua falada e da língua escrita.

Esses números podem explicar o porquê de a frequência de *de* não passar de 63% no texto de Chico Buarque, e chegar a 80% no de Augusto Cury. O trecho de Chico Buarque considerado para o levantamento dos dados contém uma descrição de lugares por onde o personagem principal transita, daí a profusão de expressões espaciais, favorecedoras de *em*, sendo empregadas como adjuntos adnominais, como nos casos que seguem em (18) a seguir. Diferentemente, o texto de auto-ajuda de Augusto Cury não explora descrições espaciais, recorrendo mais amplamente a metáforas de caráter bastante abstrato, o que se reflete na presença massiva de adjuntos adnominais preposicionados com o argumento interno apontando para um elemento imaterial, não favorecedoras do uso de *em*, como nos exemplos em (19). Os condicionamentos para a ocorrência de *de* e *em* não sofrem, assim, interferência de qualquer fator extralingüístico, mas apenas intralingüístico, atrelado à especificidade semântica do argumento interno da preposição.

- (18) a. ...ressentiria o cheiro do carpete alaranjado do hotel *em Melbourne*. (Chico Buarque, p. 21.)
- b. Senti vontade de ligar para alguém *no Brasil*... (Chico Buarque, p. 21.)
- c. ...ficou de me dar um guia de bons endereços *em Manhattan*. (Chico Buarque, p. 23.)
- d. Teria casado com ela, na capela de uma ilha *na Baía de Guanabara* (Chico Buarque, p. 30.)
- e. seria estúpido relatar ... a minha madrugada solitária *em Budapeste*. (Chico Buarque, p. 31.)
- (19) a. ...vivendo as batalhas *da existência*... (Augusto Cury, p. 9.)
- b. Desejos não resistem às dificuldades *da vida*... (Augusto Cury, p. 12.)
- c. ...ele detonava um gatilho psíquico que abria em milésimos de segundos a janela *da memória* em que a imagem doentia estava registrada (Augusto Cury, p. 13.)
- d. ...bastidores *da nossa mente* que afetam todo o processo de construção de pensamentos e geram os traumas psíquicos (Augusto Cury, p. 14.)
- e. Os maiores riscos para quem sonha são as pedras *do caminho*. (Augusto Cury, p. 19.)

Diante desse quadro, interessa-nos agora partir para o mesmo tipo de debate que foi desenvolvido para *ter e haver* na seção anterior: como capturar a variação entre *de* e *em* numa análise que procura caracterizar o conhecimento de um falante do português brasileiro sobre o uso dessas preposições em adjuntos adnominais? Como ponto de partida para a resposta, podemos explorar a idéia de que não estamos diante uma variação que envolve a interferência de condições da gramática periférica, do contrário a alternância entre uma e outra preposição estaria sujeita a condicionamentos extralingüísticos.

O caminho mais viável, nesse sentido, é considerar que estamos diante de uma variação interna à gramática nuclear. Se quisermos investir nessa hipótese, devemos nos perguntar por que o BEff não mostra sinal algum de estar se importando com essa variação. Uma possível resposta seria que, da

mesma forma que o proposto para *ter* e *haver*, no caso de *de* e *em* teríamos uma diferença de estatuto funcional, e daí as duas formas não estariam no mesmo pé de igualdade; por extensão, o BEff não atuaria no bloqueio da variação, dada a suposta diferença de funcionalidade entre as duas formas. Contudo, considero essa visão problemática, pelo menos no estágio atual do debate acerca do estatuto mais ou menos funcional das preposições. Embora *de* pareça ser semanticamente mais esvaziada que *em* (AVELAR, 2004), não é claro como capturar formalmente esse esvaziamento por meio de uma gradação de funcionalidade (TSENG, 2000); além disso, quando se trata de contextos que não o de adjuntos adnominais, a preposição *em* mostra um caráter tão semanticamente esvaziado quanto o de *de*, como nos casos em (20) a seguir; da mesma forma, a preposição *de* pode apontar, em certos contextos, para um conteúdo semântico bem delineado, veiculando a noção de origem, ponto de partida ou outro afim, como nas construções em (21). Descarto então qualquer resposta que passe pela idéia de o BEff estar “desligado” nos contextos de variação *de/em* pelo suposto fato de um item ser mais funcional que o outro.

- (20) a. A moça só fala **em** casamento.
 b. Eu pensei **em** viajar logo no primeiro dia de férias.
- (21) a. A criança fugiu **do** cachorro.
 b. O rapaz viajou **do** Rio para São Paulo numa motocicleta.

Mantendo a idéia de explorarmos procedimentos internos à gramática nuclear, uma outra via passa pela consideração de dados negativos, impossíveis de serem levantados nas amostras da fala e da escrita. Os sintagmas adnominais nucleados por *de* e *em* exibem um comportamento sintático diferenciado (AVELAR, 2005b,c), cuja observação pode ser determinante para caracterizarmos a variação no interior da gramática nuclear. Um dos contrastes diz respeito, por exemplo, à extração de constituintes interrogativos, como vemos em (22)-(23) a seguir. Nos dados em (22), é possível extrair o constituinte interrogativo nucleado por *de*, como em (22c), sem que a construção se torne agramatical; em (23), diferentemente, é impossível extrair o constituinte interrogativo nucleado por *em*, como vemos em (23c), de modo a manter o sentido que esse constituinte veicula quando em posição

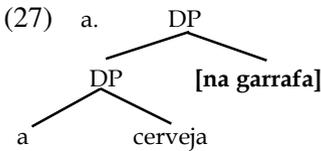
adnominal. A leitura possível para (23c) é aquela em que, absurdamente, a criança se encontra dentro da caixa enquanto come o bombom, caso em que o constituinte preposicionado é interpretado como adjunto adverbial, e não adnominal.

- (22) a. A criança comeu o bombom *da* caixa.
 b. A criança comeu o bombom *de qual* caixa?
 c. *De qual* caixa que a criança comeu o bombom?
- (23) a. A criança comeu o bombom *na* caixa.
 b. A criança comeu o bombom *em qual* caixa?
 c. * *Em qual* caixa que a criança comeu o bombom?

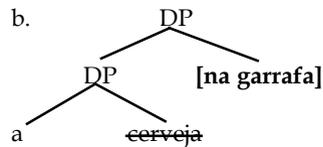
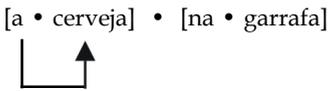
Esses e outros contrastes são discutidos em Avelar (em preparação; ver também 2004, 2005b,c), onde defendo que o comportamento diferenciado entre esses sintagmas está atrelado a padrões distintos de configuração entre a adjunção adnominal com *de* e a adjunção adnominal com *em*. É impossível trazer para este trabalho, dados os seus limites, o conjunto de argumentos que estabeleço em torno dessa hipótese, mas vou apresentar brevemente a discussão de Raposo (1999) em torno do contraste entre *de* e outras preposições da língua, que tomo como ponto de partida para a argumentação. Raposo nota que os artigos-demonstrativos *o(s)*, *a(s)* exibem um contraste de sensibilidade morfo-fonológica quando são modificados por constituintes preposicionados: se a preposição for *de*, o artigo-demonstrativo é licenciado, como nos casos em (24a-26a); se a preposição for *em*, *com* ou *para*, respectivamente como em (24b)-(26b), o uso do artigo-demonstrativo resulta numa construção marginal.

- (24) a. O rapaz bebeu *a* (cerveja) *da* garrafa.
 b. O rapaz bebeu *a* *(cerveja) *na* garrafa.
- (25) a. O bandido roubou *o* (dinheiro) *da* velhinha.
 b. O bandido roubou *o* ??(dinheiro) *com* a velhinha.
- (26) a. A costureira consertou *a* (saia) *da* Maria.
 b. A costureira consertou *a* *(saia) *pra* Maria.

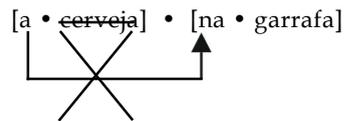
O autor interpreta a oposição considerando três fatores: (a) artigos-demonstrativos são fonologicamente dependentes e necessitam de um “hospedeiro”; (b) explorando o modelo de fase (CHOMSKY, 1998/2000), os itens *com*, *em* e *p(a)ra* nucleiam constituintes formados numa fase que não à do DP/NP modificado; e (c) *de* é inserido pós-sintaticamente, a caminho da Forma Fonológica. A conjugação desses fatores permite uma análise como a que segue. Como ilustrado em (27) adiante, o sintagma com *em* (o que vale também para um sintagma com *para* ou *com*) se encontra com todas as suas matrizes fonológicas linearizadas ao entrar nas dependências do DP, por corresponder a uma fase. Se o constituinte *cerveja* for realizado, o determinante pode se afixar a ele na morfologia, como indicado em (27a); diferentemente, se *cerveja* estiver elíptico (digamos, por apagamento durante a linearização), como em (27b), o artigo-demonstrativo ficará sem hospedeiro: o constituinte à direita, por ser uma fase, se fecha para qualquer operação pós-sintática, sendo impossível migrar para seu domínio em busca de “apoio morfo-fonológico”. A expressão **a na garrafa* é mal formada, portanto, pelo fato de requerimentos morfo-fonológicos do artigo-demonstrativo não terem sido satisfeitos.



Afixação no componente morfológico:

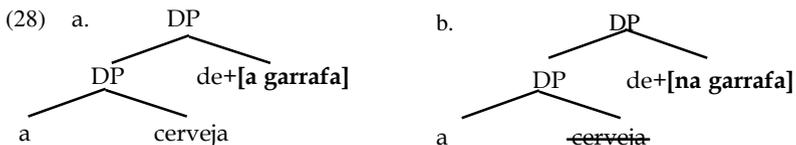


Afixação no componente morfológico:

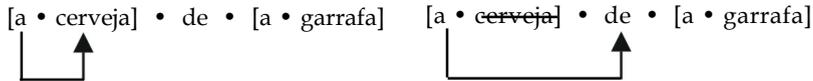


Diferentemente, tratando *de* como um morfema dissociado (no sentido de que sua inserção é pós-sintática), teríamos a situação em (28). O que vai se adjungir ao DP *a cerveja*, nesse caso, é outro DP, que traria/receberia Caso inerente genitivo. O DP adjungido também se constituiria como fase, e o morfema correspondente a *de* seria inserido pós-sintaticamente, em sua margem. Nas condições apontadas, os elementos do DP *a cerveja* serão

linearizados juntamente com *de*, o que abre espaço para que a matriz fonológica dos termos internos a esse DP e do morfema dissociado interajam na morfologia. O artigo-demonstrativo vai então poder se afixar à preposição, como indicado em (28b): embora tenha entrado na derivação para se associar a uma fase pronta, o item *de* é externo a essa fase e adjacente ao DP na estrutura linearizada, condição que licencia a afixação (para maiores detalhes em torno de operações pós-sintáticas, ver Embick & Noyer, 2001).



Afixação no componente morfológico: Afixação no componente morfológico:



O estatuto morfológicamente diferenciado da preposição *de* pode ser reforçado por diferenças semânticas entre esse item, por um lado, e as demais preposições da língua, por outro. Embora estejamos focalizando neste trabalho apenas construções com *de* e *em*, o fato é que a preposição *de* pode substituir outras preposições, tal como nos casos em (29)-(31) a seguir. A rigor, ao lado da variação *de/em*, encontramos variações do tipo *delcom*, *delpara*, *delsobre*, etc.

- (29) a. A empregada passou aquela camiseta **de / com** botão.
b. Aquele rapaz **de / com** cabelos longos perguntou pela Ana.
- (30) a. Roupa **de / p(a)ra** criança custa muito caro.
b. A viagem **de / p(a)ra** Brasília vai ser antes da viagem **de / p(a)ra** Manaus.
- (31) a. A notícia **do / sobre** o acidente vai ser um choque para a família da vítima.
b. Aquele história **de / sobre** impeachment parece ser boato da oposição.

Um outro contraste igualmente interessante entre *de* e as demais preposições diz respeito à possibilidade de alterarmos a posição dos constituintes nas relações intermediadas por *de*, tal como em (32a-b)-(34a-b) adiante, sem que o papel semântico dos membros da relação seja afetado. Notemos, por exemplo, que a relação semântica entre *as praias famosas* e *a cidade* em (32a-b) (bem como entre *aquele bolo* e *(o) fubá* em (33a-b) e *aqueles livros* e *(a) bolsa* em (34a-b)) é preservada, não importando se um ou outro constituinte ocorre externa ou internamente à preposição. Contrariamente, se formas como *com*, *em* ou *p(a)ra* forem empregadas, a posição de cada constituinte precisa ser fixa, sem o quê a construção resultante será inaceitável e/ou produzirá outro significado. Especificamente em (32c), *a cidade* ocorre como o argumento interno, e a preposição usada deve ser *em*, e não *com* ou *p(a)ra*, o que certamente está associado ao fato de a condição requerida ser a de que o argumento tenha a interpretação de *localização* ou *todo/continente*. Já se o argumento interno for *as praias famosas*, como em (32d), a preposição deve ser *com*, de modo a satisfazer o papel de *localizado* ou *partel/conteúdo*. As mesmas observações cabem para os casos em (33)-(34).

- (32) a. *As praias famosas da cidade* vão ficar lotadas no verão.
 b. *A cidade das praias famosas* vai receber muitos turistas no verão.
 c. *As praias famosas *com/em/*p(a)ra a cidade* vão ficar lotadas no verão.
 d. *A cidade com/*em/*p(a)ra as praias famosas* vai receber muitos turistas.
- (33) a. *Aquele bolo de fubá* ficou gostoso.
 b. *O fubá daquele bolo* estava estragado.
 c. *Aquele bolo com/*em/*p(a)ra fubá* ficou gostoso.
 d. *O fubá *com/em/p(a)ra aquele bolo* estava estragado.
- (34) a. *A Ana não consegue carregar aquela bolsa de livros.*
 b. *Os livros daquela bolsa* são todos da Ana.
 c. *A Ana não consegue carregar aquela bolsa com/*em/#p(a)ra livros.*
 d. *Os livros *com/em/#p(a)ra aquela bolsa* são todos da Ana.

Se a proposta de Raposo (1999) estiver no caminho correto, os contrastes semânticos entre os sintagmas construídos com *de* e com as demais preposições podem ser capturados nos seguintes termos: uma vez que, quando nucleando um adjunto adnominal, o item *de* é inserido pós-sintaticamente, a caminho da Forma Fonológica, o seu conteúdo não pode ser interpretado na Forma Lógica (dado o pressuposto minimalista de que os procedimentos do componente morfo-fonológico não têm efeito no componente semântico). Sendo esse o caso, o significado dessa preposição (origem, ponto de partida, ou outros afins; ver as construções em (21)) não interfere no tipo de relação semântica estabelecida entre o DP modificado e o DP modificador. As demais preposições da língua, entre as quais se inclui a forma *em*, integrarão a estrutura já na computação sintática, o que determina que seu conteúdo seja legível no componente semântico e sua ocorrência se dê apenas em contextos especificados, daí os contrastes observados naqueles casos em (30)-(34).

Temos, assim, uma resposta para o porquê de o BEff não estar se importando com a variação entre *de* e *em*: as duas formas não são selecionadas sob as mesmas condições, a primeira consistindo num morfema dissociado, inserido na estrutura apenas depois da computação sintática, e a segunda, num item que compõe a estrutura desde a etapa estritamente sintática da derivação. A forma *de*, nesse sentido, pode ser tratada como uma espécie de “preposição curinga”, entrando na estrutura quando a modificação adnominal envolve a adjunção sintática de um DP a um DP, como em Raposo (1999); diferentemente, a preposição *em* veicula um conteúdo semântico específico, compondo a estrutura do constituinte já na sintaxe, sendo licenciada apenas nos contextos em que se requer uma interpretação locativa ou outras afins. A rigor, portanto, o que na superfície vemos como um fenômeno de variação entre dois itens corresponde, na gramática nuclear, a dois mecanismos diferenciados de adjunção, um resultando na realização de *de*, e outro, na de *em*.

Esta análise pode encaminhar uma resposta para o comportamento uniforme da distribuição de *de* e *em* na fala e na escrita, diferentemente do comportamento exibido por *ter* e *haver*. No caso de *de/em*, não entra em jogo qualquer propriedade da gramática periférica, mas mecanismos exclusivos da gramática nuclear, “construídos” no processo de aquisição natural da língua. Esses mecanismos devem ser os mesmos entre os indivíduos da comunidade e, como tal, não estão sujeitos a fatores externos à gramática. O comportamento esperado em tal caso é que a distribuição dessas preposições

entre os diferentes segmentos da comunidade seja uniforme, independentemente de nível de escolarização, idade ou classe social, bem como de se tratar ou não de língua falada ou escrita. No caso de *ter/haver*, contrariamente, os padrões diferenciados derivam de diferenças individuais, atrelados a informações da gramática periférica, construída numa etapa que não mais diz respeito à aquisição natural da língua, mas a um aprendizado consciente, não-natural e social, na linha do defendido em Kato (2005).

5. NO FUTURO...

Consideremos agora a situação do nosso lingüista de 2500, sem acesso aos dados de fala e sem qualquer possibilidade de obter os chamados dados negativos, que poderiam evidenciar o conhecimento internalizado dos falantes. A questão que nos interessa é exatamente a de saber se seria possível chegar às mesmas conclusões propostas para *ter/haver* e *delem* olhando apenas para a escrita. O ponto de partida para o lingüista do futuro deverá ser o levantamento de hipóteses com base nos padrões de distribuição identificados nas fontes escritas. Especificamente sobre *ter* e *haver*, as hipóteses poderiam ser as seguintes:

- (35) (a) havia uma variação entre *ter* e *haver* em contextos existenciais do português brasileiro na passagem do milênio. Essa variação pode ser reflexo de um processo de mudança em progresso, com uma forma invadindo o domínio da outra, o que exigirá a observação de outros recortes temporais para a sua confirmação;
- (b) ocorria uma variação entre *ter* e *haver*, mas *haver* era o verbo existencial canônico, dado que a sua ocorrência era amplamente maior que a de *ter*;
- (c) *ter* e *haver* estavam em variação estável no português brasileiro na passagem do milênio, com o primeiro sendo empregado em situações mais informais, e o segundo, em situações formais;
- (d) *ter* e *haver* estavam em variação, compondo uma mudança em progresso, com *ter* sendo o verbo inovador, o que se reflete pelo seu uso mais informal; na passagem do milênio, contudo, essa mudança ainda estava num estágio incipiente, já que a frequência de *ter* não passava dos 14%.

Todas essas situações são incompatíveis com o comportamento de *ter/haver* na língua falada; conseqüentemente, se investir em qualquer das quatro, o lingüista se equivocará acerca do conhecimento internalizado dos falantes, cuja descrição é o seu principal objetivo. No que diz respeito à fala, a hipótese correta deveria ser a de que *ter* é o verbo canônico, fato completamente obscurecido pelas distribuições identificadas na língua escrita. Sobre a gramática nuclear, se o raciocínio desenvolvido na seção 3 estiver correto, não deverá existir variação entre *ter* e *haver* como verbos funcionais em pé de igualdade, mas um verbo existencial funcional (*ter*) convivendo ao lado de outros verbos existenciais substantivos, entre os quais se inclui *haver*.

Um outro caminho que o lingüista poderia seguir antes de formular suas hipóteses diz respeito à idéia de que o passado pode ser explicado a partir do presente. Dada a natureza do objeto em foco (a variação entre dois itens supostamente funcionais), um pressuposto do qual o lingüista poderá se valer é o do BEff, cuja ação deve perpassar a gramática internalizada de qualquer indivíduo falante de toda e qualquer língua em qualquer que seja o recorte temporal. Obviamente, aliado a esse pressuposto, o lingüista do futuro deverá congregiar o máximo de informações sociolingüísticas e/ou pragmático-discursivas acerca da produção escrita na passagem do milênio. Por exemplo, ele precisará ter em mente que, no Brasil do ano 2000, valorizava-se um padrão apartado da língua falada, que mirava em muitos aspectos a gramática do português europeu ou estágios anteriores da língua. Da mesma forma, ele deverá conhecer que os diferentes gêneros textuais poderiam afrouxar ou radicalizar esse afastamento da língua falada, tendo em vista os diferentes objetivos a que um determinado texto se propunha. O casamento de pressupostos sobre princípios gerais da gramática com o conhecimento de informações extralingüísticas, visando à depreensão de fatos da gramática nuclear subjacente a um texto, pode-nos levar a concluir, juntamente com Kroch (1994), que

[...] the best explanation for the occurrence of doublets is sociolinguistic: Doublets arise through dialect and language contact and compete in usage until one or other form wins out. Due to their sociolinguistic origins, the two forms often appear in different registers, styles, or social dialects; but they can only coexist stably

in the speech community if they differentiate in meaning, thereby ceasing to be doublets. Speakers learn either one or the other form in the course of basic language acquisition, but not both. Later in life, on exposure to a wider range of language, they may hear and come to recognize the competing form, which for them has the status of a foreign element. They may borrow this foreign form into their own speech and writing for its sociolinguistic value or even just because it is frequent in their language environment. Over time, however, as dialects and registers level out through prolonged contact, the doublets tend to disappear.⁸ (p. 6)

Em síntese, o analista deverá ter em mente que é impossível a existência de *doublets* numa mesma gramática nuclear. Se duas formas aparecerem em variação, pelo menos duas possibilidades precisam ser consideradas: ou (a) pertencem a gramáticas distintas, numa situação de “competição de gramáticas”, ou, ainda, numa disputa que envolve a gramática nuclear e a gramática periférica, seguindo a discussão de Kato (2005), ou (b) pertencem a uma mesma gramática, com a ocorrência de uma e outra forma não se dando em condições intralingüísticas idênticas. Somente na primeira possibilidade é que estaríamos diante de *doublets*, realizados sob condições idênticas ou similares, mas necessariamente pertencentes a gramáticas distintas.

Qual deve ser o plano de ação a ser aplicado pelo lingüista do futuro para saber se está diante da situação indicada em (a) ou em (b)? Tendo acesso a informações que permitam caracterizar os documentos escritos quanto a aspectos sociais e pragmático-discursivos, o analista poderá aproveitar o padrão de variação que caracteriza o comportamento das duas formas para alicerçar suas hipóteses. Vamos considerar que existam três padrões distintos: **(i) variação com condicionamentos exclusivamente extralingüísticos**, **(ii) variação com condicionamentos exclusivamente intralingüísticos** e **(iii) variação com condicionamentos extralingüísticos e intralingüísticos**. O analista poderá apostar alto na idéia de que está diante de *doublets* se os resultados apontarem para o padrão em (i); se for o caso do padrão em (ii), a melhor aposta será a de que as formas em variação são estratégias internas a uma mesma gramática. Já na situação em (iii), que não será considerada neste trabalho (embora, do ponto de vista metodológico, seja a mais instigante e desafiadora), o analista deverá pensar duas vezes antes de apostar suas fichas.

Adotando esses pressupostos, o lingüista do futuro inevitavelmente deverá concluir que o comportamento de *ter/haver* na escrita não pode ser fruto de procedimentos da gramática nuclear, uma vez que a variação não está sujeita a condicionamentos intralingüísticos. Sobre a forma verbal que deve receber o *status* de existencial funcional da gramática nuclear, o lingüista terá de considerar, pelo menos, duas possibilidades. Numa delas, alguns falantes trariam *haver* como o funcional em sua gramática nuclear, e outros falantes, o verbo *ter*, o que poderia ser fruto de diferenças dialetais no português brasileiro da passagem do milênio. Na outra possibilidade, o que se verificaria na amostra seria a predominância de uma forma mais conservadora, que atuava como um item funcional em estágios anteriores da língua, mas não no atual, e é recuperada como tal por ainda ter prestígio na língua escrita. O lingüista poderia preferir a primeira opção pelo fato de alguns textos trazerem exclusivamente o verbo *haver*, o que estaria indicando que os autores destes documentos particulares não dispõem de *ter* como verbo existencial em seu dialeto. Já os autores cujos documentos mostram a variação entre os dois verbos apresentariam *ter* como existencial; particularmente entre estes, o emprego ainda bastante recorrente de *haver* se daria pelo prestígio deste verbo. Dentro dessa visão, o dialeto destes últimos seria mais inovador que o daqueles. Em outras palavras, essa posição levaria o lingüista a concluir que o dialeto de Chico Buarque, dos autores dos anúncios de revistas e da turma do Casseta e Planeta detinha, na passagem do milênio, um caráter mais inovador que o de Paulo Coelho, Augusto Cury, Alexandre Cherman, e todos os demais.

Contudo, o mínimo de conhecimento sobre a prática da escrita na passagem do milênio, envolvendo concepções do “bom escrever” e relações entre autor e leitor, poderiam livrar o lingüista desse equívoco. E não só: buscar informações sobre a data e o local de nascimento dos autores, bem como sobre onde viveram a infância, poderia trazer pistas sobre se estamos diante ou não de dialetos diferenciados. Em relação a esta parte, excluindo os autores dos jornais e anúncios (de cuja identidade não tenho conhecimento), todos os demais são nascidos (ou foram criados) no Rio ou em São Paulo, regiões geograficamente bastante próximas. Esse fator deveria ser levado em conta na tentativa de depreender uma possível emergência de diferenças dialetais drásticas, a ponto de separar seus falantes pelo uso de categorias funcionais distintas. Se o lingüista do futuro tiver em mãos esses

conhecimentos, ele poderá notar que (1) piadas trazem situações do cotidiano, envolvendo necessariamente reproduções de diálogos e, por isso, seu registro deve ser um reflexo do oral; (2) como pretendem convencer, anúncios são mais eficazes se se aproximam do discurso natural de quem os lê; (3) na passagem do milênio, a língua escrita era conservadora, valorizando aspectos da norma portuguesa e tendendo a rejeitar inovações da fala. A partir daí, o lingüista estará a um passo de concluir que, na língua falada e na gramática nuclear, o verbo existencial canônico é *ter*, e que dificilmente diferenças dialetais estariam interferindo na distribuição dos verbos.

O lingüista do futuro poderia ainda ampliar os recortes temporais, analisando documentos de estágios anteriores do português brasileiro, na tentativa de identificar em que ponto da história teria se dado a emergência do verbo possessivo como um item existencial. A esse respeito, os estudos de Callou & Avelar (2002,2003,2005) mostram que *ter*, na língua escrita, já era usado como existencial na segunda metade do século XIX. O gráfico a seguir aponta a distribuição dos dois verbos em contextos existenciais ao longo de oito séculos, considerando documentos produzidos exclusivamente no Brasil a partir do século XVII. Ressalte-se que o percentual de 8% apontado para *ter* em contextos existenciais nos séculos XVII e XVIII são índices suspeitos: trata-se dos chamados contextos existenciais opacos, em que é impossível determinar se a sentença aponta para uma interpretação existencial ou possessiva (Callou & Avelar, 2005). A partir do século XIX, contudo, os contextos de ocorrência de *ter* são inequivocamente existenciais, com o verbo alcançando um total de 22% entre os dados na escrita.

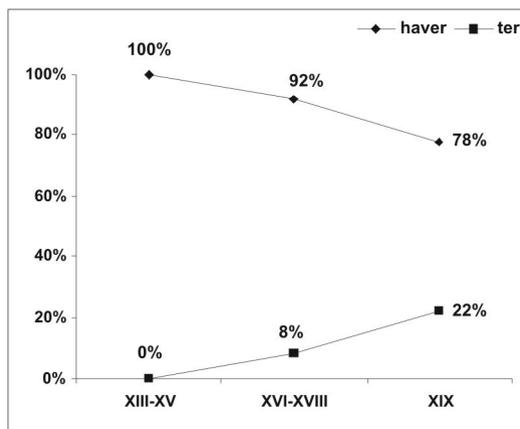


GRÁFICO 8: Percentual de *ter* e *haver* com uso existencial em oitos séculos de história do português, considerando-se documentos escritos exclusivamente no Brasil a partir dos séculos XVII (Callou & Avelar, 2005)

Interessantemente, numa amostra abarcando anúncios e editoriais de jornais do século XIX provenientes de diferentes estados do país, é num anúncio de um jornal paulistano de 1879, reproduzido em (36) a seguir, que Callou & Avelar (2002) identificam o primeiro caso de *ter* com valor inequivocamente existencial. O anúncio forja um diálogo bastante informal para convencer o leitor de que o melhor lugar para a compra de tranças é num tal Salão Oliveira. Trata-se de um contexto de produção que reúne estratégias identificadas tanto nas piadas do Casseta (diálogos) como nos anúncios de revistas (lista de qualidades do produto a ser vendido).

(36) ***Correio Paulistano, 08 de janeiro de 1879***

Ora Nenê? | - A onde comprou essas | tranças de cabelos | tão bonitos? – Foi no SALÃO OLIVEIRA, | RUA DIREITA Número 42; | só lá é que **tem** um grande sortimento de postiços | de todas as qualidades, | e que vende tudo mais | barato do que em outra | qualquer parte. | - Olhe Nenê!!! Eu garanto-lhe que lá não **tem** | tranças de imitação!! É tudo | de cabellos finissimos; também | **tem** outra casa. | - Aonde? | - Na RUA DA IMPERATRIZ Número 30; | que bonitos

posições lá **tem!**! - Mas olha Nenê, a firma das | duas casas é | OLIVEIRA & CUNHA|. (Guedes & Berlinck 2000, p. 408)

Ainda sobre os números no Gráfico 8, se não levasse em conta o conjunto de pressupostos em torno de propriedades de uma gramática nuclear, o lingüista do futuro poderia olhar para o gráfico e inferir que, do século XIX ao século XXI, o verbo *ter* estava “capengando” como existencial: o decréscimo na frequência desse verbo, de 22% para 14% (ver o Gráfico 1, na seção 2) entre as fontes escritas, poderia ser tomado como um reflexo de mudanças que estariam ocorrendo na língua falada, na qual o uso de *ter* teria supostamente decrescido. Embora seja para isso que a língua-E dos textos escritos esteja apontando, sabemos que essa visão é equivocada, pelo menos no que diz respeito aos séculos XX e XXI (ver CALLOU e AVELAR, 2000). Até poderíamos indagar se, no século XIX, não estaríamos de fato diante de uma verdadeira variação, com a gramática nuclear abarcando dois verbos existenciais, já que não temos acesso à língua falada daquele período. Mas lembremos que o BEff só admitiria tal situação se os dois verbos estivessem sujeitos a situações intralingüísticas específicas, condição para a qual não temos nenhuma evidência, o que nos leva a descartar a hipótese de variação na gramática nuclear entre os dados daquele estágio. Além disso, dispomos de informações bastante valiosas, por parte dos observadores da língua naquele período, de que *ter* já era usado como verbo existencial. Por exemplo, Júlio Ribeiro, em sua *Grammatica Portuguesa*, afirma que o verbo *ter* vinha “se tornando geral no Brasil, até mesmo entre as pessoas illustradas” (1914, p. 296). Esse é um forte indício de que, na língua falada da época e, provavelmente, na gramática nuclear, a situação de *ter* já era a mesma que a identificada no início do terceiro milênio, com a variação *ter/haver* resultando antes de uma disputa entre a gramática nuclear e a gramática periférica. Como vemos, se assumir a aplicabilidade do BEff, levando em conta informações sociolingüísticas sobre as fontes analisadas, o lingüista do futuro não se enganará tão facilmente pelo que irá encontrar na superfície dos textos.

Tratemos agora, sob o mesmo enfoque, dos fatos envolvendo *de/em*, tendo em mente a informação de que apenas condicionamentos intralingüísticos atuam nessa variação. Se adotar os mesmos pressupostos que levou em conta para *ter/haver*, o lingüista acabará por inferir que a variação entre *de* e *em* é

resultado de operações internas a uma mesma gramática, mas que, dado o *blocking effect*, essas operações devem corresponder a procedimentos diferenciados. Esse é exatamente o quadro a que chegamos para a variação *delem* na seção 4, explorando dados negativos. As particularidades da conclusão, contudo, deveriam contar com uma exploração mais detalhada dos dados. Por exemplo, para elaborar a idéia de que a preposição *de* funciona, naqueles dados, como uma espécie de “preposição curinga”, com sua matriz fonológica sendo inserida pós-sintaticamente, o lingüista deverá perceber que *de* pode ocorrer não apenas em contextos em que a forma *em* pode ser usada, mas também formas como *com*, *para*, *sobre*, etc. Novamente, o conhecimento de princípios gerais de funcionamento da gramática pode entrar em jogo nas tomadas de decisão do analista. Se explorar a idéia de que todas as línguas naturais empregam a estratégia de conectar um DP a outro DP (por adjunção ou não) e de que, superficialmente, operações desse tipo se manifestam, em algumas línguas, pela manifestação de uma categoria curinga, semanticamente esvaziada (cuja função, digamos, seja a de atribuir Caso para um dos DPs), o lingüista poderá chegar a alguma hipótese interessante sobre a condição interna da variação *delem*. Olhando para línguas contemporâneas, temos evidências bastante fortes para relações dessa natureza: línguas como o inglês dispõem de um índice genitivo para marcar um determinado tipo de relação estabelecida entre dois DPs (*Peter's car*, *the USA president's airplane*, etc.); outras línguas, como o árabe marroquino (*ktab Nadia* ‘o livro da Nádia’ (OUHALLA, 1998)) e o pima bajo (*huun uus* ‘planta do milho (lit.)’ (FERNANDEZ, 2003)) podem gerar o mesmo tipo de relação pela simples justaposição dos dois DPs; entre as línguas românicas, o sinal de que dois DPs se combinaram para uma relação genitiva como a do inglês pode ser exatamente a manifestação da preposição *de/di* (ver Kayne, 2002, para discussões de casos no francês). Por essa via, o lingüista poderá apostar na identificação de um tipo de operação similar entre os dados do início do milênio, com a preposição *de*, dado o seu comportamento peculiar, sendo a manifestação dessa operação na superfície da língua.

Um outro ponto a ser frisado é que pode não ficar claro para o lingüista do futuro se o variação *delem* consiste no reflexo de uma mudança em progresso. Se for esse o caso, os procedimentos morfossintáticos que subjazem à ocorrência de uma preposição poderão estar tomando o lugar dos procedimentos que subjazem à ocorrência de outra, com o lingüista tendo em

mãos apenas um recorte revelando um determinado estágio da mudança. Contudo, podemos questionar a validade ou o lugar da noção de *mudança em progresso* no interior de uma abordagem desenvolvida à luz da Teoria de Princípios e Parâmetros: a rigor, não existe coisa alguma que possa ser chamada de *mudança* no interior de uma gramática nuclear. A mudança lingüística que interessa ao gerativista é aquela que resulta do fato de um indivíduo, no processo de aquisição, formar a sua gramática nuclear com propriedades diversas às da gramática pertencente à geração que lhe fornece os dados primários. Não faz sentido, sob essa ótica, falar de uma mudança em progresso ocorrendo na gramática internalizada. Daí Lightfoot (1999) afirmar que a mudança gramatical é como o pôr-do-sol: embora qualquer pessoa já tenha visto o sol se pondo, sabemos que esse acontecimento é uma ilusão, uma vez que é o próprio planeta, e não o astro-rei, que realiza um movimento em torno do próprio eixo, dando origem ao que chamamos de *amanhecer* e *anoitecer*. Da mesma forma, não existe entidade orgânica alguma chamada *língua*, supostamente sujeita a um conjunto de mudanças, pairando sobre uma comunidade, mas crianças marcando os parâmetros da sua gramática diversamente à marcação das gerações anteriores; conseqüentemente, a mudança lenta e gradual que depreendemos a partir das fontes escritas é, na verdade, o reflexo de uma ruptura abrupta. A variação não deve então ser tomada como indício de uma mudança em progresso, mas de uma mudança que já se operou entre as gramáticas internalizadas de diferentes gerações da comunidade, e que só aos poucos vai ganhando espaço entre as fontes escritas. O lingüista do futuro deverá assim reformular a sua indagação no sentido de saber se as gerações anteriores usavam *de/em* diferentemente da geração que forneceu os dados usados para a pesquisa, descartando a idéia de mudança em progresso no âmbito da gramática nuclear. Seria preciso descobrir, por exemplo, se os pais ou avós de Chico Buarque, Paulo Coelho, da turma do Casseta e demais autores realizavam os casos de adjunção adnominal com *de* e *em* diferentemente dos seus filhos e netos. Sobre esse aspecto em particular, não se dispõe atualmente, até onde eu saiba, de números sobre a variação entre *de* e *em* em outros estágios do português; também desconheço qualquer abordagem que considere a hipótese de essas duas preposições terem experimentado algum tipo de mudança no interior de adjuntos adnominais.

6. CONSIDERAÇÕES FINAIS

Se o proposto neste trabalho estiver no caminho correto, os passos teórico-metodológicos que levariam o linguísta do futuro a concluir que o caso de *ter/haver* é reflexo de uma competição entre gramáticas são os mesmos que lhe permitiriam considerar que a variação *de/em* provém de operações internas a uma mesma gramática. Creio que exercícios dessa ordem, fazendo-nos confrontar dados de fala e escrita provenientes de período sobre o qual temos intuições satisfatórias, pode ser profícuo na tentativa de estabelecer dispositivos de análises que nos permitam ir para além da superfície dos textos e atingir a gramática internalizada de quem os produziu. É crucial, para o sucesso dessa atividade, casar a abordagem sociolingüística, que nos autoriza a olhar para o texto como um produto sociocultural, com uma abordagem que explore princípios gerais da linguagem. A convergência dessas dimensões, às vezes tomadas como inconciliáveis, consiste num instrumento poderoso para o trabalho com “maus dados”. Aproveitar o que temos no presente para verificar os resultados dessa relação pode ser, como espero ter contribuído para mostrar, uma ferramenta lucrativa nas investigações sobre o passado da língua.

NOTAS

¹ Quero aqui agradecer a todos que contribuíram valiosamente para o desenvolvimento deste trabalho, em especial: Dinah Callou, Maria Eugênia Duarte, Mary Kato, Charlotte Galves, Maria Clara Paixão de Sousa e Jairo Nunes.

² “Em materiais históricos, encontramos um tipo de informação que está necessariamente ausente em dados sincrônicos e oferece uma importante contribuição à lingüística geral a partir da história; isto é, informação sobre a mudança lingüística no curso do tempo. Com tal informação, esperamos depreender como as gramáticas mudam de um estado a outro ao longo do tempo, bem como aprender mais sobre seus princípios de organização, a partir da compreensão do processo pelo qual elas mudam. Perturbar um sistema complexo e observar sua subsequente evolução é freqüentemente um caminho excelente para alcançar a estrutura interna. Em adição, uma vez que as características de qualquer língua em um dado ponto no tempo são o resultado de um complexo entrecruzamento de princípios gerais da linguagem e desenvolvimentos históricos particulares, o conhecimento do processo histórico pelo qual uma língua tenha alcançado um determinado estágio pode ser importante para demarcar

apropriadamente a responsabilidade de fatores históricos e gerais, e portanto para a formular apropriadamente a teoria lingüística.” (tradução minha)

³ “[...] o presente é como o passado, e os princípios gerais que derivam do estudo das línguas atualmente vivas também são válidos para línguas antigas.” (tradução minha)

⁴ Como as amostras do NURC-RJ e PEUL provêm de décadas distintas, um indivíduo integrante do PEUL estava cerca de dez anos mais velho quando as entrevistas da amostra NURC-RJ foram realizadas. Como não estou interessado em detectar mudanças em tempo aparente, essa diferença não interfere sobre os aspectos em que irei me deter.

⁵ Das revistas *Veja*, *Isto é* e *Época*, extraíram-se todos os casos de *ter/haver* em existenciais, enquanto da *Isto é* e *Super Interessante*, todos com *delem*. Entre os jornais, extraíram-se todas as construções relevantes com *ter/haver* e *delem*. Entre os livros, foram extraídos os 40 primeiros casos de *delem*, e entre 10 e 17 (primeiros casos) de *ter/haver*.

⁶ O fato de os dados da escrita pertencerem ao século XXI e os de fala serem provenientes do século XX poderia ser alvo de críticas, no sentido de que possíveis processos de mudança na fala, ao longo dos vinte e cinco anos considerados, pode ter promovido diferenças significativas na escrita. Contudo, uma vez que o objetivo final não é atestar mudanças em progresso, mas refletir acerca de como propriedades de uma gramática nuclear emergem nos padrões de distribuição atestados em língua-E, diferenças pequenas de datação não têm efeitos significativos na conclusão final.

⁷ A codificação das construções existenciais contou com os seguintes grupos de fatores, tomando a especificidade do verbo (*ter/haver*) como *variável dependente*: (a) marcação de definitude no argumento do verbo, (b) marcação de número no argumento, (c) especificidade semântica do argumento (material animado, material inanimado, evento, espaço, abstrato, outros), (d) tempo e modo verbal, (e) amostra (NURC ou PEUL), (f) faixa etária, (g) gênero, (h) fonte (língua falada ou escrita) e (i) fonte escrita de proveniência dos dados. Quanto a *delem*, os dados foram codificados considerando-se os seguintes grupos de fatores, com a especificidade da preposição (*de* ou *em*) correspondendo à variável dependente: (a) definitude e (b) especificidade semântica (material animado, material inanimado, evento, espaço, imaterial/abstrato) do argumento da preposição, (c) definitude, (d) função sintática e (e) especificidade semântica do constituinte modificado pelo sintagma preposicionado (idem b), (f) gênero, (g) faixa etária, (h) fonte (fala ou escrita), (i) fonte escrita de proveniência dos dados, (j) amostra (NURC ou PEUL).

⁸ “A melhor explicação para a ocorrência de *doublets* é sociolingüística: *doublets* surgem por meio do contato de línguas e dialetos e competem em uso até que uma forma vença. Devido às suas origens sociolingüísticas, as duas formas freqüentemente aparecem em diferentes registros, estilos ou dialetos sociais; podem, porém, coexistirem de forma estável na comunidade de fala se passarem a ter significados diferentes, deixando portanto de funcionarem como *doublets*. Os falantes aprendem uma ou outra forma (mas não ambas) durante o processo básico de aquisição da língua. Mais tarde, quando em exposição a um espectro mais amplo da linguagem, os falantes podem ouvir e reconhecer a forma competidora, que tem para eles o estatuto de um elemento estranho. Eles podem usar esta forma estranha em sua própria fala e escrita, por causa do valor sociolingüístico demonstrado por essa forma ou apenas porque ela é freqüente em seu ambiente lingüístico. Ao longo do tempo, entretanto, quando dialetos e registros se estabilizam através de um contato prolongado, os *doublets* tendem a desaparecer.” (tradução minha)

REFERÊNCIAS

- ARONOFF, M. *Word formation in generative grammar*. MIT Press, 1976.
- AVELAR, J. *De verbo funcional a verbo substantivo: uma hipótese para a supressão de haver no português brasileiro*, 2005a. (a sair na Revista *Letras de Hoje*, publicação da PUC-RS).
- _____. *Constituintes preposicionados, derivação por fase e critérios de interpretação temática*. Comunicação apresentada no Seminário do GEL, realizado em julho de 2005, na Universidade Federal de São Carlos, 2005b.
- _____. *Sobre PPs adnominais no português brasileiro*. Comunicação apresentada no XI Seta, realizado em outubro de 2005, no IEL/UNICAMP, 2005c.
- _____. (em preparação). *Adjuntos locativos no português brasileiro*. Tese de Doutorado. IEL/UNICAMP.
- _____. *Dinâmicas morfosintáticas com ter, ser e estar no português brasileiro*. 2004. Dissertação (Mestrado) – IEL/UNICAMP.
- CALLOU, D.; AVELAR, J. *Sobre a emergência do verbo possessivo em contextos existenciais na história do português*, 2005. (Ms)
- _____. *‘Ter’ and ‘haver’ in the history of Portuguese: the appearance of ‘ter’ in existential environments*. Comunicação apresentada no N.WAVE 32, na Universidade da Pensilvânia, outubro/2003.

CALLOU, D. Estruturas com 'ter' e 'haver' em anúncios do século XIX. In: ALKMIM, T. *Para a história do português brasileiro*. São Paulo: Humanitas-USP, 2002. v. III, p. 47-68.

_____. Sobre *ter* e *haver* em construções existenciais: variação e mudança no português do Brasil. *Gragoatá* 9. p. 85-114, 2000.

CHOMSKY, N. Minimalist inquiries: the framework. In: MARTIN, R. *et al.* *Step-by-step: essays in minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: The MIT Press. p. 89-155, 1998/2000.

_____. *Knowledge of language*. New York: Praeger, 1986.

_____. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CLARK, E. Locational: existential, locative and possessive constructions. In: GREENBERG, J. (Org.). *Universals of human languages*. *Syntax*. Stanford: Stanford University Press, 1978. V. 4.

CLARK, R.; ROBERTS, I. A computational model of language learnability and language change. *Linguistic Inquiry* 24. p. 294-345, 1993.

EMBICK, D. *Blocking effects and Analytic/Synthetic Alternations*, 2005. (Ms).

EMBICK, D.; NOYER, R. Movement operations after syntax. *Linguistic Inquiry*, 32. p. 555-595, 2001.

FERNANDEZ, Z. *Construcciones posesivas en Pima Bajo*. *Liames*, 3, p. 23-37, 2003.

GUEDES, M.; BERLINCK, R. *E os preços eram commodos... anúncios de jornais brasileiros – Século XIX*. São Paulo: Humanitas/FFLCH-USP, 2000.

HARLEY, H.; NOYER, R. *Distributed Morphology*. 2003. Disponível em: <<http://linguistics.arizona.edu/~harley>> Acesso em: 09 set. 2003

KAYNE, R. *On some prepositions that look DP-internal: English of and French de*, 2002. Ms.

KATO, M. A gramática do letrado: questões para a teoria gramatical. MARQUES, M. *et al.* (Org.). *Ciências da linguagem: trinta anos de investigação e ensino*. Braga: CEHUM (U. do Minho), 2005. p. 131-145.

KISHIMOTO, H. Locational verbs, agreement, and object shift in Japanese. *The Linguistic Review*, 17. p. 53-109, 2000.

KROCH, A. *Morphosyntactic variation*, 1994. (Ms)

_____. *Reflexes of grammar in patterns of language change*. *Language Variation and Change*, 1. p. 199-244, 1989.

LABOV, W. The use of the present to explain the past. In: *Proceedings of the 11th International Congress of Linguistics*. Bologna: Il Molino, 1975. p. 825-851.

LIGHTFOOT, D. *The development of language: acquisition, change, and evolution*. Malden, MA: Blackwell, 1999.

_____. *How to set parameters: arguments from language change*. Cambridge, MA: MIT Press, 1991.

_____. *Principles of Diachronic Syntax*. New York: Cambridge University Press, 1979.

MATTOS e SILVA, R. V. Observações sobre a variação no uso dos verbos *ser*, *estar*, *haver*, *ter* no galego-português ducentista. *Estudos lingüísticos e literários*, 19. p. 253-285, 1997.

MILSARK, G. *Existential sentences in English*. Cambridge: Department of Linguistics/MIT, 1974.

OUHALLA, J. *Possession in sentences and noun phrases*, 1998. (Ms)

PINKER, S. *O instinto da linguagem*. São Paulo: Martins Fontes, 2002.

PINTZUK, S.; TSOULAS, G.; WARNER, A. Syntactic Change: Theory and Method. In: PINTZUK, S.; TSOULAS, G.; WARNER, A. (Org.). *Diachronic syntax – Models and mechanisms*. Oxford University Press, 2000. p. 1-22.

RAPOSO, E. *Towards a minimalist account of nominal anaphora*, 1999. (Ms)

RIBEIRO, J. *Grammatica Portuguesa*. Rio de Janeiro: Livraria Francisco Alves, 1914.

TATTERSALL, I. Não estávamos sozinhos. *Scientific American Brasil*, Edição especial, n. 2. p. 29, 2003.

TSENG, J. 2000. *The representation and selection of prepositions*. 2000. Tese (Doutoramento) – University of Edinburgh.