



Sociolinguística cognitiva e o estudo da convergência/divergência entre o português europeu e o português brasileiro

Augusto Soares da Silva (Universidade Católica Portuguesa – Braga)*

RESUMO: este estudo apresenta elementos da emergente Sociolinguística Cognitiva e um programa de sociolinguística cognitiva e quantitativa do português europeu e brasileiro centrado na questão da convergência ou divergência entre as duas variedades. Primeiro, mostramos as razões da inevitabilidade da sociolinguística cognitiva dentro do paradigma da Linguística Cognitiva. A seguir, sumariamos os resultados da investigação sociolexicológica sobre o vocabulário do futebol e do vestuário e indicamos as extensões desta investigação para os domínios das palavras funcionais e das variáveis não-lexicais.

Palavras-chave: cognição social, linguística de corpus, português europeu/brasileiro, sociolinguística cognitiva, variação linguística

Introdução

A sociolinguística cognitiva é uma área ainda emergente no paradigma da Linguística Cognitiva. Na primeira parte deste estudo, mostraremos as razões da inevitabilidade da sociolinguística cognitiva e os seus contributos específicos para o estudo da variação linguística. Argumentaremos ainda que o desenvolvimento da sociolinguística cognitiva, centrada nos aspectos sociais da variação linguística, depende de uma metodologia empírica sustentada em *corpora* representativos e em métodos quantitativos e multivariacionais; e que esta sociolinguística cognitiva e quantitativa – com aplicação a outras questões, já presentes em estudos cognitivos anteriores, como a relação entre linguagem e cultura (estabelecida na noção de modelos cognitivos culturais) e entre linguagem e ideologia – muito contribuirá para a necessária compreensão sociocognitiva da linguagem, bem como da cognição como *cognição social*.

Precisamos de mais estudos empíricos sobre a variação do Português, como língua supranacional e transcontinental que é. Apresentaremos, na segunda parte deste estudo, um programa de sociolinguística cognitiva e quantitativa do Português Europeu e Brasileiro centrado nas questões da convergência e divergência entre as duas variedades nacionais e da estratificação interna da cada uma. Sumariaremos os resultados da investigação sociolexicológica sobre o vocabulário do futebol e da moda/vestuário e indicaremos as extensões desta primeira fase de investigação para os domínios das palavras funcionais e das variáveis não-lexicais (sintáticas e morfológicas).

1. Inevitabilidade e especificidade da Sociolinguística Cognitiva

A autodefinição da Linguística Cognitiva como *modelo baseado no uso* (Langacker (1988) (2000)) implica uma orientação sociolinguística para o estudo da variação da linguagem. A razão é simples: a variação é a consequência imediata e

* assilva@braga.ucp.pt



inevitável do uso. Paradoxalmente, ou talvez não, a variação linguística não tem sido objecto prioritário das mais conhecidas linhas de investigação linguístico-cognitiva. Ressalvem-se importantes estudos cognitivos (alguns cognitivo-funcionalistas) sobre a variação linguística, primariamente de três pontos de vista: diacrónico, com a investigação sobre a gramaticalização e a mudança semântica (por ex., HOPPER & TRAUGOTT (1993), TRAUGOTT e DASHER (2002), GEERAERTS (1997) BLANK & KOCH (1999)); tipológico e antropológico, incluindo os estudos sobre modelos cognitivos culturais e variação cultural (por ex., PALMER (1996), LAKOFF (1996), KÖVECSSES (2000) (2005), LEVINSON (2003), DIRVEN, FRANK e PÜTZ (2003)); e psicolinguístico, sobre a aquisição e o desenvolvimento da linguagem (TOMASELLO (2003)). Mas a variação intralinguística, ou diversidade sociolinguística, essa é ainda pouco estudada dentro do mesmo enquadramento. Constituem excepções os trabalhos de Geeraerts, Grondelaers & Speelman (1999), Geeraerts (2003, 2005), Kristiansen (2003) e, muito recentemente, a colectânea de estudos organizada por Kristiansen & Dirven (2006), institucionalizando a noção emergente de *sociolinguística cognitiva*.

Como se justifica a consequência lógica enunciada, sobretudo no contexto da cognição humana, tendencialmente vista como um fenómeno interno individual e invariável, e no contexto de uma teoria empenhada no estudo dos processos e funções cognitivos da linguagem, aparentemente mais individuais do que sociais? Antes de mais, convém esclarecer que a expressão *modelo baseado no uso* – introduzida por Langacker (1988, 2000), popularizada por Barlow & Kemmer (2000) e reforçada em publicações mais recentes como as de Tomasello (2000, 2003), Bybee & Hopper (2001), Croft & Cruse (2004) ou Geeraerts (2006) – pode interpretar-se de dois modos. Em primeiro lugar, no sentido da definição de Langacker (1988, 2000) de orientação de baixo-para-cima, maximalista e não-redutiva, em contraste com a perspectiva de cima-para-baixo, minimalista e redutiva da Gramática Generativa. Um modelo baseado no uso atribui importância substancial ao uso actual do sistema linguístico e ao conhecimento desse uso por parte dos falantes; utiliza *redes esquemáticas* inteiramente articuladas e enfatiza os esquemas ou padrões de baixo-nível (LANGACKER, 1987, p. 494). Em segundo lugar, no sentido de duas implicações maiores desta mesma definição: o método empírico de obtenção e observação dos dados a partir de *corpora* (sendo a experimentação um outro método empírico) e a variação como fazendo naturalmente parte do objecto de investigação.

Num estudo de pendor epistemológico e metodológico, Geeraerts (2005) explora este sentido implicativo para justificar a inevitabilidade de uma sociolinguística cognitiva dentro do paradigma da Linguística Cognitiva. Aponta duas razões principais pelas quais um modelo baseado no uso não pode evitar o estudo dos aspectos sociolinguísticos da variação. Por um lado, um *corpus* representativo nunca será absolutamente homogéneo, incluindo sempre alguma variação *lectal*, isto é, relacionada com qualquer tipo de variedades ou variantes de uma língua ou, simplesmente, *lectos* (dialectos, regiolectos, sociolectos, idiolectos, variedades nacionais, registos ou estilos, etc.). Mesmo que a análise não se centre na variação lectal, esta será sempre um factor, no sentido de que é necessário saber se a variação observável no *corpus* resulta ou não de factores lectais. Por outro lado, e como razão mais *positiva* do que a anterior, os aspectos sociais do significado



constituem uma forma específica de significado. Concretamente, é a variação lectal que dá origem a tipos de significado não-denotacional: o significado *emotivo* (de termos pejorativos, por exemplo), o significado *estilístico* e mais estritamente *sociolinguístico* (de termos populares ou eruditos, regionais ou sociais) e o significado *discursivo* (único de expressões como as interjeições e os marcadores discursivos; presente em termos como *senhor* e *senhora* usados em formas de tratamento). Sendo uma das marcas distintivas da Linguística Cognitiva (se não mesmo a principal), a sua orientação para o significado, estes tipos de significado *social*, e com eles a própria variação lectal, não podem ser descartados. E, mesmo que se contra-argamente com a ideia de que padrão e, por exemplo, dialectos constituem *diferentes* sistemas linguísticos, a verdade é que o conhecimento que os indivíduos têm de um mesmo sistema pode divergir consideravelmente e um mesmo indivíduo conhece e utiliza diferentes sistemas. Na realidade, uma comunidade linguística, um sistema linguístico, uma língua é um *diassistema* social. Não há, portanto, maneira de evitar a variação linguística a partir do momento em que se assuma séria e plenamente um modelo baseado no uso.

Inversamente, como argumenta Geeraerts (2005), o reconhecimento dos aspectos sociais da linguagem implica uma metodologia empírica de observação do uso dos dados, e a forma mais natural de os encontrar é em *corpora* textuais representativos; o que naturalmente implica a utilização e o desenvolvimento dos instrumentos e resultados da linguística de corpus e de métodos quantitativos de análise. O que é óbvio na óptica da investigação sociolinguística não o tem sido, pelo menos na prática, para os linguistas cognitivistas em geral. É conhecida uma certa divergência dentro da Linguística Cognitiva entre o ramo europeu, mais adepto da metodologia empírica de corpus, e o ramo americano, mais inclinado para a metodologia introspectiva, qual herança da Gramática Generativa. Por outro lado, entre os praticantes da metodologia do corpus é necessária uma maior sofisticação técnica em quantificação e estatística: trabalhos como os de Geeraerts, Grondelaers & Speelman (1999), Grondelaers, Speelman & Geeraerts (2002), Gries (2003) e Stefanowitsch & Gries (2003, 2005) são ainda excepcionais – ver em Tummers, Heylen & Geeraerts (2005) e Gries & Stefanowitsch (2006) o estado da arte dos métodos linguístico. Há ainda quem, como Itkonen (2003), num trabalho de epistemologia da linguística, assumindo uma concepção da linguagem como fenómeno social, estranhamente advogue uma metodologia intuitiva (ver discussão de GEERAERTS (2005)). Mas o facto mais relevante da situação actual é o de que as duas tendências emergentes no seio da Linguística Cognitiva – interesse crescente por métodos empíricos e maior atenção aos aspectos sociais da linguagem –, embora se impliquem mutuamente, como ficou mostrado, ainda não convergiram. Resultado: a investigação empírica, sustentada numa metodologia de corpus e quantitativa, dos aspectos sociais da variação linguística é relativamente rara dentro da Linguística Cognitiva.

Outras razões há ainda que implicam o desenvolvimento da sociolinguística cognitiva. E essas razões têm todas a ver com a própria natureza e o programa da Linguística Cognitiva: para além de modelo baseado no uso, a Linguística Cognitiva é também um modelo orientado para o significado, um modelo experientialista e um modelo recontextualizante (SILVA, 2004a, para uma síntese). Como modelo orientado para o significado, não poderá deixar de lado a variação sociolinguística, já



que esta constitui uma forma específica de significado, quer como diversos tipos de significado não-denotacional, acima referidos, quer como, e em termos de Gramática Cognitiva (LANGACKER, 1987, 1991, 1999), *subjectivização* (LANGACKER, 1990), isto é, o processo que consiste em pôr elementos do acto de fala “em palco”, como foco específico de atenção (focalizar locutor e/ou alocutário como membros de determinado grupo social ou numa determinada relação interactiva ou focalizar a situação do acto de fala no sentido de implicar uma atitude específica). Como modelo experiencialista, não poderá deixar de lado a experiência colectiva, social e cultural – tópico que exploraremos na secção seguinte. Finalmente, como modelo recontextualizante (GEERAERTS, 2003, SILVA, 2005c), em resposta à descontextualização da gramática exemplarmente efectuada pela Gramática Generativa, não poderá deixar de incorporar nenhuma das facetas do *contexto*: não só as bases cognitivas e experienciais dos falantes ou contexto cognitivo, como também o nível interaccional do uso linguístico ou contexto situacional e o ambiente sociocultural da língua ou contexto social. Por tudo isto e porque a Linguística Cognitiva ainda não respondeu devidamente a estas implicações variacionistas, é que a sociolinguística cognitiva constitui, no momento actual, um dos seus maiores desafios.

Mas o que há de específico na emergente sociolinguística cognitiva, ou que contributo pode ela oferecer à investigação sociolinguística em geral? Algumas respostas foram já dadas. Em síntese, podemos apontar três especificidades e contributos. Em primeiro lugar, a própria perspectiva cognitiva dos fenómenos variacionais; concretamente, a aplicação dos diversos modelos descritivos já existentes no estudo da variação linguística. Por exemplo, a abordagem da variação fonética pela teoria do protótipo (KRASTIANSEN, 2003) ou a utilização de metáforas conceptuais e modelos culturais na identificação de atitudes linguísticas (GEERAERTS, 2003). Em segundo lugar, a exploração da *cognição social*, de que nos ocuparemos a seguir; particularmente, a elucidação da interacção dialéctica entre o nível individual cognitivo e o lado social das normas colectivas. Finalmente, mas não menos importante, o desenvolvimento de métodos quantitativos baseados em *corpora* e de métodos de análise multivariacional da confluência de factores conceptuais, discursivos e variacionais dos fenómenos linguísticos.

2. Cognição social

As estruturas linguísticas exprimem conceptualizações e as conceptualizações realizadas na e através da linguagem têm uma base experiencial, isto é, estão intrinsecamente relacionadas ao modo como os seres humanos experienciam a realidade, tanto fisiológica como culturalmente. Constitui este *experiencialismo* a própria filosofia da Linguística Cognitiva, paradigmaticamente elaborada por Lakoff & Johnson (1999), em ruptura com o *objectivismo* de outras grandes correntes linguísticas, com a estruturalista e a generativista. Mas a interpretação de Lakoff & Johnson, com a qual mais se identificou a Linguística Cognitiva no seu conjunto, é parcial: a tese da *corporização* (“embodiment”) do pensamento e da linguagem ou a *filosofia na carne* foca a vertente individual e universal da cognição humana (o corpo é um universal da experiência humana), o



seu lado físico e neurofisiológico, recentemente explorado por Lakoff (2003) na sua Teoria Neural da Metáfora.¹ Ora, tendo a experiência humana uma dimensão também colectiva e interactiva, social, cultural e histórica e, portanto, variacional, impõe-se não reduzir a filosofia experiencialista e o princípio da corporização a operações neurais meramente individuais. É justamente isto que tem sido reclamado no seio da própria Linguística Cognitiva por autores como Sinha & Jensen de López (2000), Zlatev (2003), Harder (2003), Bernárdez (2004, 2005), entre outros: a natureza socialmente interactiva da linguagem e o seu ambiente cultural devem ser reconhecidos como elementos igualmente fundacionais da perspectiva cognitiva.²

O problema está então na própria concepção da cognição humana e na separação entre o individual e o colectivo, o interno e o externo, o pensamento e a acção, a cognição e a actividade – pensamento individual e acção colectiva, será esta a equação adequada? Poderemos conceber a cognição sem a interacção? Poderemos continuar a assumir, como é habitual na cultura ocidental (e não só), que o pensamento individual interno tem algum tipo de preeminência sobre a actividade supra-individual externa ou sobre o pensamento dirigido para a (inter)acção? Vários filósofos, psicólogos, neurocientistas e linguistas respondem que não. Lembremos a inseparabilidade da cognição e da emoção, demonstrada nos estudos de Damásio (1995, 2000). Efectivamente, tem havido, nos últimos vinte ou mais anos, um alargamento significativo do âmbito da cognição: desde uma perspectiva puramente interna com a primeira geração das ciências cognitivas, à perspectiva *corporizada* (VARELA, THOMPSON e ROSCH, 1991; EDELMAN, 1992; DAMÁSIO, 1995, 2000; LAKOFF e JOHNSON, 1999; GIBBS, 2005) aberta ao exterior, e, mais recentemente, à inclusão da situação, actividade ou interacção na cognição e, assim, à noção de *cognição situada*, *cognição distribuída* ou *cognição social* (ZLATEV, 1997, 1999; TOMASELLO, 1999; BERNÁRDEZ, 2004, 2005; ZIEMKE, ZLATEV e FRANK, no prelo; DIRVEN, FRANK, ZINKEN e ZLATEV, no prelo). Bernárdez (2005) fala da inevitabilidade de conceber a cognição e a linguagem em termos de actividade socialmente corporizada. A grande questão, ainda em aberto, é saber como é que especificamente interagem os dois tipos de factores da cognição e da linguagem – os factores individuais, neurofisiológicos e universais, de um lado, e os factores interindividuais, socioculturais e variacionais, do outro. Questão, afinal, antiga e recorrente das relações entre linguagem, cultura e cognição.

Algumas articulações têm sido feitas. Langacker (1994) advoga uma perspectiva dinâmica e um entendimento da natureza cíclica do desenvolvimento cognitivo, balanceado entre capacidades psicológicas inatas (como a organização figura/fundo, a reificação conceptual, a categorização, a esquematização, a capacidade de ponto de referência) e estruturas mentais estabelecidas a partir da experiência prévia, umas pré-culturais (as que emergem bastante cedo) e outras marcadamente culturais. Assim se compaginam e se interligam, na cognição e na linguagem, factores *universais*, directamente ligados ao facto de os indivíduos terem a mesma estrutura biológica e interagirem num mundo basicamente igual para todos, e factores *culturalmente específicos*. Shore (1996) desenvolve uma teoria etnográfica da mente e uma teoria cognitiva da cultura ou, simplesmente, uma teoria da *cultura-na-mente* e mostra que os modelos culturais são uma parte integrante do processamento mental e são inevitavelmente condicionados por este. Zlatev (1997, 1999) combina a perspectiva biológica do significado com a tese da *corporização*



situada e propõe o conceito de *mimese* (uso intencional do corpo para fins representacionais) para a articulação entre o significado público e convencional e o significado privado e subjectivo. Tomasello (1999, 2003) investiga as origens culturais da cognição humana e desenvolve um modelo da aquisição da linguagem baseado no uso. Bernárdez (2004) explora as noções de *cognição colectiva* e *cognição-para-a-acção* e sugere que a maior parte da cognição humana depende da acção humana (pensar e fazer estão intimamente ligados, ao contrário do que a cultura ocidental faz acreditar).

Bernárdez (2005) sumaria importantes pontos da concepção sociocognitiva da linguagem: as línguas existem somente na forma de actividade social; a actividade linguística é essencialmente colectiva, social por natureza; conseqüentemente, a linguagem é um fenómeno inerentemente histórico e, como tal, fenómenos que hoje são examinados a nível individual (como a metáfora e a metonímia) são o resultado de uma cristalização social, histórica; a actividade linguística (o uso linguístico) determina formas e estruturas linguísticas, umas preferidas e outras não; através de um processo de integração cognitiva; as formas produzidas e preferidas pela actividade linguística são gradualmente fixadas na mente do indivíduo, a ponto de a sua conexão original com a actividade se perder e, desse modo, se tornarem parte da cognição individual – é assim que a actividade linguística social influencia os próprios processos cognitivos.

3. O estudo da convergência e divergência entre o Português Europeu e o Português Brasileiro

Desenvolvemos investigação sobre duas questões das relações entre o Português Europeu e o Português Brasileiro (doravante, PE e PB): a primeira e principal é saber se as duas variedades nacionais se encontram num processo de convergência ou de divergência; a segunda é comparar a estratificação interna actual das duas variedades, para saber se a distância entre padrão e registos subpadrão (intermédios) é maior numa variedade ou na outra e analisar os indicadores de subpadronização. Especificamente, são três as variáveis de análise para responder a estas questões diacrónica (a primeira) e sincrónica (a segunda), as quais estruturam o *corpus* CONDIVport em construção:

- geográfica: Portugal vs. Brasil (São Paulo e Rio de Janeiro)
- diacrónica: últimos 50 anos – 1950, 1970, 1990/2000
- estilística: jornais e revistas de qualidade > jornais e revistas populares > Internet *offline* (fóruns, *e-mails*) > Internet *online* ou *chats*

A investigação apoia-se na concepção geral e nos métodos quantitativos da investigação sociolexicológica desenvolvida pela unidade belga, liderada por Dirk Geeraerts, para o Neerlandês da Holanda e da Bélgica (GEERAERTS, GRONDELAERS e SPEELMAN, 1999). É utilizado o método *onomasiológico* para o estudo da variação, mais especificamente a variação onomasiológica envolvendo sinónimos denotacionais (referenciais), como *avançado*, *atacante* e *dianteiro* ou *casaco* e *blazer* (mas não *atacante* e *jogador* ou *casaco* e *blusão*) e unidades



gramaticais funcionalmente equivalentes, porque são os sinónimos denotacionais (lexicais ou gramaticais) os que mais diferenças sociolinguísticas (regionais, sociais, estilísticas, pragmático-discursivas, históricas) mostram e, assim, melhor revelam as relações entre variedades linguísticas diferentes. E, para medir a convergência e divergência e outras distâncias entre as duas variedades do Português, utilizamos o método quantitativo da *uniformidade linguística baseada em perfis* (SPEELMAN, GRONDELAERS e GEERAERTS, 2003): um método que se baseia nas noções de *perfil onomasiológico* ou conjunto de sinónimos usados para designar um conceito/função, diferenciados pela sua frequência relativa, e *uniformidade* (U) ou medida da correspondência entre dois conjuntos de dados, definidos em termos de perfis onomasiológicos. Por exemplo, a uniformidade de um conceito/função entre duas amostras, em que uma contém seis ocorrências do termo/forma A e quatro do termo/forma B e a outra três ocorrências de A e sete de B, resulta do número de pares comuns de nomeação desse conceito/função (sete pares), sendo, portanto, $U = 70\%$. Este mesmo resultado obtém-se somando as frequências relativas mais pequenas de cada termo/forma alternativo: 30% de A e 40% de B. Duas variedades encontram-se em convergência quando a medida de uniformidade U aumenta, e em divergência quando U diminui. E a distância maior ou menor entre registos formais e informais manifesta-se, respectivamente, em U menor e maior.

Como hipóteses sobre as relações entre PE e PB, admite-se (i) uma maior distância entre registos formais e informais no PB do que no PE; (ii) uma influência crescente do PB sobre o PE (e as variedades africanas); (iii) e, embora não se encontrem hipóteses claras nem, muito menos, bem fundamentadas na respectiva literatura, a ideia generalizada de uma fragmentação progressiva e inelutável e, portanto, a hipótese da divergência.

3.1. Resultados da investigação sociolinguística

A primeira fase da investigação incide no léxico.³ Escolhemos dois domínios: futebol e moda/vestuário, por serem populares e permeáveis à influência de línguas estrangeiras, com a diferença da emotividade e presumível influência brasileira do primeiro. A base empírica da investigação consiste em largos milhares de observações do uso de termos alternativos para nomear 43 conceitos nominais destes dois campos lexicais. Os materiais são extraídos de (i) oito jornais de desporto e 24 revistas de moda dos primeiros anos das décadas de 50, 70 e 90-2000; (ii) linguagem da Internet da conversação *online* de IRC ou *chats*; e (iii) etiquetas de lojas de vestuário de Braga, Lisboa, São Paulo e Rio de Janeiro. O *corpus* CONDIVport contém 4 milhões de palavras do registo formal (jornais de desporto e revistas de moda) e 15 milhões de palavras do registo informal (*chats* de futebol e etiquetas de lojas de vestuário). Está parcialmente disponibilizado no sítio da Linguateca, em <www.linguateca.pt/ACDC>.

São analisados 21 perfis onomasiológicos de futebol, num total de 183 termos, e 22 perfis onomasiológicos de vestuário, num total de 264 termos. Os primeiros compreendem os conceitos de



'árbitro', 'árbitro auxiliar', 'avançado', 'baliza', 'bola', 'defesa', 'equipa', 'extremo', 'falta', 'finta', 'fora-de-jogo', 'golo', 'grande penalidade', 'guarda-redes', 'jogada', 'jogo', 'médio', 'pontapé', 'pontapé de canto', 'pontapé livre', 'treinador'

Os segundos incluem os conceitos de

'blusa f', 'blusão m/f', 'calças m/f', 'calças curtas m/f', 'calças justas f', 'camisa m', 'camisola m/f', 'casaco f', 'casaco m', 'casaco curto f', 'casaco curto m', 'casaco de cerimónia m/f', 'casaco de malha m/f', 'casaco impermeável m/f', 'casaco quente (Inverno) m/f', 'conjunto f', 'fato m', 'jaqueta m/f', 'jeans m/f', 'saia f', 't-shirt m/f', 'vestido f'

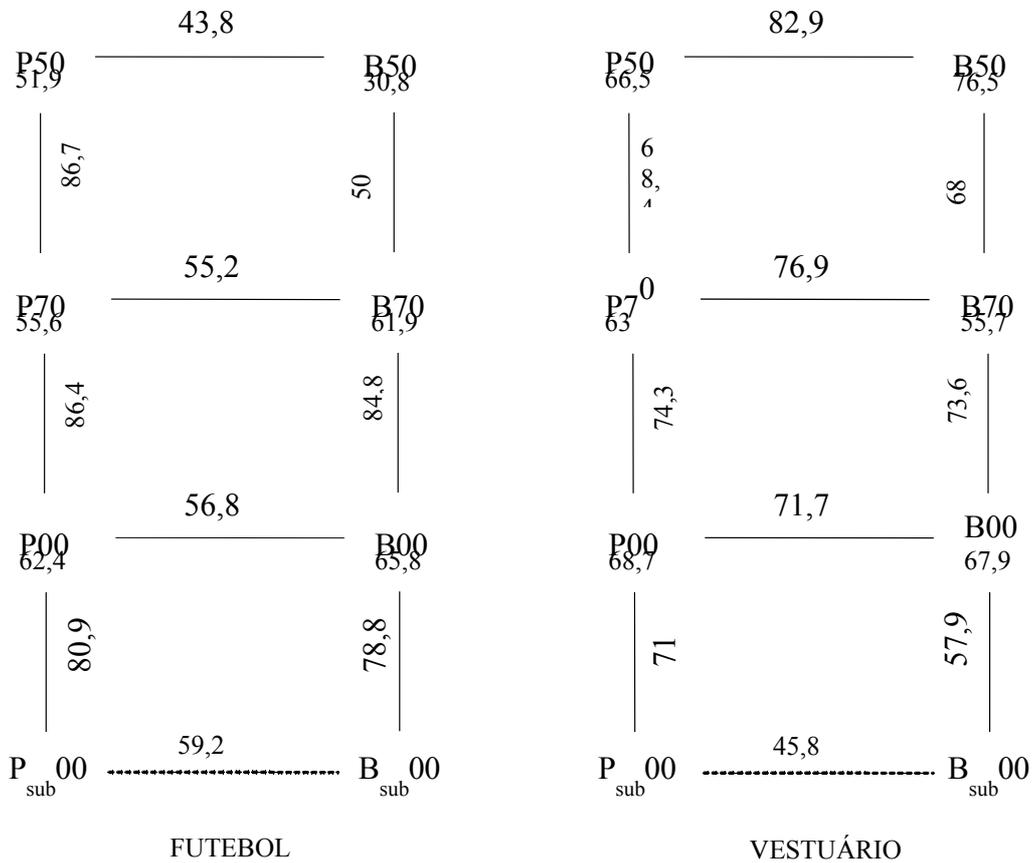
Apenas dois exemplos de perfis onomasiológicos:

GUARDA-REDES: *arquero, goal-keeper, goleiro, golquíper, guarda-meta, guarda-redes, guarda-vala(s), guardião, keeper, porteiro, quíper, vigia*

T-SHIRT M/F: *camisa, camiseta/e, camisetete, camisola, licra, singlet, tee-shirt, t-shirt.*

A base de dados de futebol contém 90.202 observações do uso dos respectivos termos nos jornais de desporto e 143.946 observações nos *chats* de futebol; a de moda contém 12.448 observações do uso dos respectivos termos de vestuário nas revistas de moda e 3.240 observações nas etiquetas de montras de lojas de vestuário.

Pelas limitações de espaço, apresentamos os resultados da pesquisa de modo muito sumário (para mais informações, ver Silva e Duarte (DUARTE, 2005 e SILVA, 2006). O Quadro 1 sintetiza os resultados, expressos em valores percentuais de uniformidade média (U') do conjunto dos perfis onomasiológicos de futebol e de vestuário analisados do português de Portugal (P) e do Brasil (B), relativamente às duas questões da investigação: convergência/divergência ao longo das décadas de 50 (P50/B50), 70 (P70/B70) e 90-00 (P00/B00) e estratificação actual ou distância entre padrão (P00, B00) e subpadrão (P_{sub}00, B_{sub}00). Nas linhas horizontais, encontram-se os valores de uniformidade U' entre as duas variedades; nas linhas verticais, os valores de uniformidade U' interna na década e de uniformidade U' interna de década para década dentro de cada variedade.



Quadro 1. Valores globais de uniformidade diacrónica entre PE e PB e de uniformidade actual entre padrão e subpadrão

Os resultados dos termos de futebol e dos termos de vestuário são diferentes, sobretudo em relação à questão principal da convergência ou divergência entre PE e PB. Os termos de vestuário confirmam a hipótese de divergência (diminuição do valor de uniformidade U' de 82,9% para 71,7%) e a hipótese de uma distância maior entre padrão e subpadrão na variedade brasileira (57,9% de U' , contra 71% no PE). Pelo contrário, os termos de futebol apontam no sentido da convergência (aumento do valor de uniformidade U' de 43,8% para 56,8%), mais clara de 50 para 70 e pouco significativa de 70 para 2000, mas, ao mesmo tempo, mostram uma uniformidade consideravelmente menor (grande distância) entre as duas variedades em todos os períodos, expressa numa média de 50% de não-uniformidade. Esta disparidade de resultados é um motivo acrescido para a análise de outros campos lexicais (estamos a analisar termos de saúde e tencionamos analisar termos de transportes e de banca e finanças).

Entre factores *internos* destes resultados diacrónicos, está a questão de saber que influência têm os estrangeirismos na tendência evolutiva geral. Medindo a proporção de termos com a característica (A) 'anglicismo', 'francesismo' ou 'estrangeirismo' (qualquer que seja a proveniência) no perfil onomasiológico de um



conceito e, depois, no conjunto dos perfis analisados (A'), nas amostras do PE e PB do futebol e do vestuário,⁴ os Quadros 2-4 mostram a influência do inglês no conjunto da amostra de futebol e do francês e do inglês no conjunto da amostra do vestuário e ainda o peso de todos os estrangeirismos no conjunto das respectivas amostras.

A' _{Ing} (P50)	7,1%	18,0%	A' _{Ingl} (B50)
A' _{Ing} (P70)	9,8%	17,1%	A' _{Ingl} (B70)
A' _{Ing} (P00)	10,2%	16,2%	A' _{Ingl} (B00)
A' _{estrang} (P50)	13,9%	23,5%	A' _{estrang} (B50)
A' _{estrang} (P70)	17,9%	22,8%	A' _{estrang} (B70)
A' _{estrang} (P00)	18,5%	23,3%	A' _{estrang} (B00)

Quadro 2. Percentagens de anglicismos e de todos os estrangeirismos nos termos de futebol

A' _{Ing.adapt} (P50)	6,0%	2,8%	A' _{Ing.adapt} (B50)
A' _{Ing.adapt} (P70)	7,9%	16,9%	A' _{Ing.adapt} (B70)
A' _{Ing.adapt} (P00)	8,9%	16,0%	A' _{Ing.adapt} (B00)

Quadro 3. Percentagens de adaptações/decalques de anglicismos nos termos de futebol

A' _{Fr} (P50)	17,6%	18,5%	A' _{Fr} (B50)
A' _{Fr} (P70)	15,9%	18,1%	A' _{Fr} (B70)
A' _{Fr} (P00)	10,2%	7,9%	A' _{Fr} (B00)
A' _{Ing} (P50)	3,3%	4,2%	A' _{Ing} (B50)
A' _{Ing} (P70)	5,8%	7,6%	A' _{Ing} (B70)
A' _{Ing} (P00)	16,9%	17,0%	A' _{Ing} (B00)
A' _{estrang} (P50)	22,4%	23,8%	A' _{estrang} (B50)
A' _{estrang} (P70)	22,1%	26,7%	A' _{estrang} (B70)
A' _{estrang} (P00)	28,2%	24,9%	A' _{estrang} (B00)

Quadro 4. Percentagens de francesismos, anglicismos e de todos os estrangeirismos nos termos de vestuário

Em relação ao futebol, a influência de anglicismos e outros estrangeirismos é claramente maior no PB do que no PE. A grande diferença percentual entre B50 e P50 resulta do facto de haver maior número e frequência de estrangeirismos conservando a sua forma original no PB. Estes resultados contribuem, em parte, para a distância maior entre as duas variedades nos anos 50 e, conseqüentemente, para a tendência convergente global. Um outro dado interessante tem a ver com a adaptação ou aportuguesamento dos termos estrangeiros (Quadro 3): a percentagem de adaptações de termos ingleses aumenta fortemente no PB de 50 para 70; pelo contrário, no PE essa percentagem mantém-se quase a mesma. Relativamente ao vestuário, verifica-se tanto no PE como no PB uma significativa diminuição de francesismos (mais acentuada no PB) e um claro aumento de anglicismos, por razões socioculturais bem conhecidas. A influência dos



estrangeirismos é também maior no PB, sobretudo nos anos 50 e 70. Todavia, não sabemos ainda qual o efeito destes resultados na tendência divergente global.

Estes resultados confirmam a maior tendência do PB para os estrangeirismos, quer para a sua importação directa, quer para a sua adaptação ou decalque; pelo contrário, no PE, e com maior visibilidade nos termos de futebol, maior é a tendência de os estrangeirismos serem substituídos por termos vernáculos.

Uma outra questão é a da esperada influência do PB no PE, justamente no domínio do futebol. O Quadro 5 sumaria os resultados percentuais de toda a amostra, em dois cálculos: as percentagens da esquerda têm em conta apenas os brasileirismos de que há menos dúvidas e as da direita, todos os brasileirismos que estão atestados nalgum dicionário de referência.

$A'_{\text{Braz}} (P50)$	0,8%	2,3%
$A'_{\text{Braz}} (P70)$	1,0%	3,4%
$A'_{\text{Braz}} (P00)$	1,1%	2,0%

Quadro 5. Percentagens de brasileirismos nos termos de futebol

Os resultados não são concludentes. Pelo menos, não permitem confirmar claramente a hipótese da influência crescente do PB no PE (lembramos que o futebol é indicado como exemplo típico desta influência). Naturalmente que há casos de influência (*bandeirinha, atacante, plantel, técnico, etc.*); mas também há brasileirismos sem qualquer atestação no PE (*goleiro, arqueiro, escanteio, impedimento, zagueiro, etc.*). Em contraste, a percentagem de lusitanismos no PB é praticamente nula, o que confirma a hipótese da não influência do PE no PB. Em conjunto, brasileirismos e lusitanismos contribuem para a grande distância entre as duas variedades em todos os períodos.

Uma outra questão diacrónica interna, mais complexa, é saber qual das variedades se aproxima ou se afasta da outra: por exemplo, se é o PB que se aproxima ou se afasta do PE e, portanto, se é do lado da variedade brasileira que a convergência ou a divergência se dá. Esta hipótese traduz-se na expectativa de que sejam mais fortes as mudanças dentro do PB do que dentro do PE, concretamente haja maiores aumentos ou maiores descidas de *uniformidade interna* dentro da variedade brasileira. Como podemos verificar pelo Quadro 1 (linhas verticais no registo padrão), os resultados apontam para a existência de mais mudanças no PB: maiores aumentos (no futebol) e descidas (no vestuário) de uniformidade interna no PB do que no PE, sobretudo dos anos 50 para os anos 70 (por outro lado, os valores de uniformidade interna das duas variedades nos anos 2000 e em ambos os campos lexicais são idênticos); maiores alterações dos valores de uniformidade entre diferentes épocas no PB do que no PE; e ainda (resultado não expresso no Quadro 1) maiores alterações das proporções de termos *exógenos* e *endógenos* no PB do que no PE.⁵ Pelo contrário, os resultados mostram uma situação de maior estabilidade do PE (ou, pelo menos, menos mudanças) em relação a estes três factores.⁶

Uma questão complementar, ainda em estudo, é saber se o movimento convergente e divergente do PB se manifesta numa orientação exógena e endógena de aproximação/afastamento a/de PE. Tendo em conta os resultados obtidos de



convergência no domínio do futebol e divergência no domínio da moda, esta hipótese traduz-se nas seguintes expectativas: o movimento convergente do PB no domínio do futebol manifestar-se-á num aumento de termos exógenos e de termos binacionais e numa diminuição de termos endógenos dentro do PB e ainda no facto de a variedade brasileira mais facilmente adoptar termos exógenos do que abdicar de termos endógenos; o movimento divergente do PB no domínio da moda manifestar-se-á num aumento de termos endógenos e numa diminuição de termos exógenos e de termos binacionais dentro do PB e ainda no facto de a variedade brasileira mais dificilmente adoptar termos exógenos do que abdicar de termos endógenos. Para o PE, a expectativa é a de que a parte dos termos exógenos, endógenos e bi-nacionais se mantenha idêntica nos diferentes períodos ou, pelo menos, não apresente uma evolução bem definida. Uma outra questão, também ainda em estudo, é analisar, através de técnicas estatísticas de *regressão linear* e *multilinear*, as correlações entre os factores internos estrangeirismos e termos endógenos/exógenos.

Quanto à questão sincrónica da estratificação (ver parte inferior do Quadro 1), os resultados do futebol, obtidos a partir da comparação entre os dados do registo formal dos jornais de desporto dos anos 90/2000 (P00 e B00) e os dados do registo informal dos *chats* (P_{sub}00 e B_{sub}00), não permitem confirmar claramente a hipótese de uma distância maior entre padrão e subpadrão no PB. De qualquer forma, a diferença de 2% a favor do PE talvez possa sugerir distância menor entre padrão e subpadrão no PE; o que poderá ser confirmado pelo maior número de discrepâncias entre padrão e subpadrão no PB. Por outro lado, os mesmos resultados mostram um grau relativamente elevado de uniformidade entre padrão (jornais) e subpadrão (*chats*) dentro de cada uma das variedades. Isto poderá indicar padronização lexical na conversação pela Internet: os utilizadores de *chat* tendem a ser conformistas quando se referem a conceitos tão populares como os do futebol. Já a uniformidade entre as duas variedades a nível do estrato subpadrão desce consideravelmente para um valor muito próximo do da uniformidade a nível do estrato padrão (59,2% e 56,8%, respectivamente).

Para a mesma questão, os resultados do vestuário, obtidos a partir da comparação entre os dados do registo formal e nacional das revistas de moda dos anos 90/2000 (P00 e B00) e os dados do registo informal e local das lojas de vestuário (P_{sub}00 e B_{sub}00), vêm confirmar claramente a hipótese de distância maior entre padrão e subpadrão no PB. Por outro lado, a distância entre as duas variedades é claramente maior a nível do estrato subpadrão (45,8%) do que a nível do estrato padrão (71,7%).

3.2. Extensão para a gramática

Como fase seguinte da investigação, pretendemos (i) estender a análise das variáveis lexicais (de palavras de conteúdo para palavras funcionais) para as variáveis não-lexicais (sintácticas e morfológicas) e (ii) ampliar e refinar a situação estratificacional de cada variedade em ordem ao estudo dos indicadores (lexicais e não-lexicais) de subpadronização. Mais especificamente, pretendemos analisar as correlações entre variáveis lexicais e variáveis não-lexicais e comparar o seu



impacto na convergência/divergência, na estratificação – particularmente as hipóteses de assimetria estratificacional (distância maior entre padrão e subpadrão no PB) e maiores alterações de uniformidade interna no PB (sinal de padronização tardia) – e na subpadronização das duas variedades nacionais. Será que palavras funcionais, como preposições, repetirão os resultados obtidos a partir dos termos (nomes) de futebol e vestuário? Será que as variáveis não-lexicais contribuirão de modo diferente das variáveis lexicais para a estrutura variacional? Quanto ao segundo objectivo, quantos mais níveis de estratificação deverão ser tomados em conta?, será a estratificação multidimensional?, como determinar o impacto mútuo dos vários factores variacionais?

Serão analisados 15 perfis de preposições e 20 perfis de construções sintácticas, seguindo os mesmos métodos onomasiológico e de uniformidade baseada em perfis (a 'sinonímia' de preposições e construções sintácticas é entendida no sentido de equivalência funcional). Os perfis preposicionais são restringidos aos mesmos tipos de complementos e ao mesmo contexto sintagmático, de forma a satisfazer a condição de sinonímia denotacional (p.ex. *reflectir/pensar em/sobre, interesse em/por, por/através de/via fax/internet/correio*). Os perfis construcionais são seleccionados de entre construções causativas/perceptivas (*fazer/mandar/deixar/ver y + INF flex/não-flex / INF não-flex + y / oração de que*), construções causativas sintéticas vs. analíticas, construções completivas e adverbiais finitas vs. infinitivas, construções de elevação, construções de sujeito nulo e de objecto nulo, construções alternativas de indicativo/conjuntivo, padrões alternativos de ordem de palavras, colocações alternativas do adjectivo atributivo e, eventualmente, outras construções. Analisaremos, ainda, formações diminutivas e aumentativas, sufixos derivacionais alternativos, formas alternativas dos pronomes possessivos e pessoais, verbos de participio duplo e formações de plural. As técnicas de regressão linear e multilinear permitirão analisar as correlações entre estas variáveis dependentes e entre elas e as variáveis independentes estilísticas (jornais de qualidade > jornais populares > Internet *offline* > Internet *online*, língua falada).

4. Situação da presente investigação relativamente ao estado da arte

A presente investigação assume posições diferentes relativamente a três tradições de estudos sobre a variação linguística: estudos sobre a variação do Português, particularmente os que têm comparado as variedades europeia e brasileira; investigação sobre línguas pluricêntricas, isto é, com diferentes variedades nacionais; e quadro teórico e metodológico do estudo da variação linguística.

No contexto dos estudos sobre variação linguística do Português, a investigação sociolinguística iniciada em Silva (2004b, 2005a, b, 2006, SILVA e DUARTE, 2005) e presentemente estendida ao domínio gramatical ocupa uma posição específica em três aspectos. Primeiro, há um rico acervo de estudos descritivos e históricos sobre as variedades europeia e brasileira, destacando-se dois projectos luso-brasileiros: um funcionalista (BRANDÃO e MOTA, 2003) e outro generativista (PERES e KATO, 2004). Mas estes trabalhos estão centrados em aspectos fonéticos e sintácticos e,



além disso, não tratam consistentemente a questão da convergência ou divergência entre as duas variedades nacionais. Segundo, a nossa abordagem utiliza uma perspectiva teórica e métodos quantitativos diferentes dos que caracterizam os estudos existentes. Terceiro, falta evidência empírica a favor ou contra as hipóteses (ainda mais intuitivas do que científicas) da divergência entre as duas variedades e da maior distância entre padrão e subpadrão no PB. E, como vimos, os nossos estudos sociolinguísticos não as confirmam inteiramente.

No contexto da investigação sobre línguas pluricêntricas, a especificidade da nossa investigação reside em dois pontos. Primeiro, focalizamos a convergência/divergência das duas variedades e a estratificação interna de cada uma, em vez de as considerarmos globalmente ou somente a nível do registo padrão. Mais especificamente, procuramos obter uma visão quantificada da convergência/divergência e da estrutura estratificacional interna. Segundo, a metodologia que utilizamos enquadra-se mais na perspectiva da sociolinguística quantitativa do que da sociologia da linguagem. Certamente que ambos os aspectos não estão ausentes da investigação actual sobre a pluricentricidade linguística, mas o que procuramos é dar-lhes maior importância do que a que geralmente é atribuída nos estudos existentes.

No contexto teórico e metodológico do estudo da variação linguística, e ao lado da investigação sociolinguística conduzida pela referida unidade belga (GEERAERTS, GRONDELARES e SPELMAN, 1999), o nosso estudo, naquele inspirado, ocupa uma posição específica em vários aspectos, sobretudo em relação aos estudos existentes sobre o Português. Primeiro, a nossa investigação insere-se no quadro teórico da Linguística Cognitiva. Esta perspectiva cognitiva está praticamente ausente nos estudos sociolinguísticos da língua portuguesa. A importância que este novo paradigma tem atribuído ao estudo da flexibilidade e variabilidade como propriedades da categorização linguística (lembramos a teoria do protótipo), o reconhecimento do papel determinante da conceptualização da realidade social na constituição dos factos sociolinguísticos (lembramos a teoria dos modelos culturais) e os contributos específicos da emergente e inevitável sociolinguística cognitiva, identificados nas duas primeiras secções deste estudo, explicam a sua relevância para a investigação sociolinguística em geral. Segundo, a presente investigação combina o interesse pela variação geográfica e o interesse pela variação estilística ou situacional, integrando, assim, a abordagem dialectométrica (VIERECK, 1996) e a abordagem estilométrica (BIBER, 1988; FINEGAN e BIBER, 1994; BIBER, CONRAD e REPPEN, 1998). Terceiro, a nossa investigação caracteriza-se pelo uso sistemático da perspectiva onomasiológica da variação. Assumimos que a convergência/divergência e as diferenças estratificacionais podem ser mais bem estudadas se analisarmos as expressões alternativas de determinado significado lexical ou determinada função linguística. Trabalhar, não com a frequência isolada de uma variável, mas com a sua posição relativa entre as suas expressões alternativas efectivamente usadas pelos falantes é o método mais adequado para classificar e comparar variáveis. De um ponto de vista linguístico, a ideia crucial é a de que os valores não-denotacionais (ou significado *social*) das expressões linguísticas (como, por exemplo, os valores estilísticos) revelam-se em diferenças de comportamento entre sinónimos denotacionais. Além disso, a perspectiva onomasiológica permite um mecanismo



seguro de controle capaz de evitar os potenciais perigos estatísticos resultantes de uma distribuição assimétrica de conceitos ou funções. Finalmente, a investigação utiliza diversas e adequadas técnicas analíticas quantitativas, nomeadamente técnicas de regressão e multivariacionais, que permitem analisar com rigor e consistência as correlações entre variáveis lexicais e não-lexicais e a interacção entre factores internos (estruturais ou semânticos) e externos (contextuais ou sociolinguísticos) da variabilidade linguística. Estes métodos quantitativos são pouco utilizados (ou mesmo desconhecidos) na sociolinguística portuguesa. Mais significativo ainda, são poucos os estudos que exploram as relações entre os diferentes factores da variabilidade linguística.

Conclusão

Esperamos ter conseguido mostrar, na primeira parte deste estudo, a inevitabilidade da emergente *sociolinguística cognitiva* e os contributos específicos para a investigação sociolinguística. Assumindo-se a Linguística Cognitiva como modelo baseado no uso, então não poderá evitar o estudo da variação *lectal*, quer porque esta estará necessariamente presente no seu mais natural objecto de análise (o *corpus*), quer porque nela se consubstancia uma forma específica de significado linguístico, de maneira nenhuma desprezível num modelo orientado para o significado. E assumindo a Linguística Cognitiva uma concepção social da linguagem – de forma a não cair no perigo do solipsismo epistemológico ou do que Sinha (1999) designa como “solipsismo neural” –, então não poderá prescindir da metodologia empírica sustentada em *corpora* representativos e em métodos quantitativos multivariacionais. A convergência destas duas implicações e necessidades, poucas vezes conseguida nos estudos cognitivos existentes, marcará uma das principais extensões da Linguística Cognitiva nestes primeiros anos do séc. XXI – o desenvolvimento da Sociolinguística Cognitiva. Com ela, daremos mais um grande passo em frente na compreensão do funcionamento da mente e da linguagem – a *cognição social* através da linguagem.

Esperamos ter conseguido mostrar também as vantagens da perspectiva cognitiva, do método da variação onomasiológica entre sinónimos denotacionais e de métodos quantitativos multivariacionais no estudo da convergência e divergência entre o PE e o PB e da estratificação interna de cada uma das variedades. Pudemos verificar que as duas variedades nacionais divergem no léxico do vestuário (e a distância actual entre padrão e subpadrão é aí claramente maior no PB), mas convergem no léxico do futebol (mas com uma distância bastante grande em todos os períodos). Em ambos os campos lexicais, é no PB que se verificam mais mudanças. A hipótese da divergência (bem como a hipótese contrária) carece de maior evidência empírica. Precisamos de consolidar a análise sociolexicológica, incluindo outros campos lexicais, e sobretudo estendê-la para o domínio das palavras funcionais e das variáveis não-lexicais. A análise das correlações entre variáveis lexicais e não-lexicais e entre estas variáveis dependentes e as variáveis independentes estilísticas ou outras sociolinguísticas permitirá um conhecimento mais exacto da verdadeira situação actual de divergência ou convergência entre PE e PB e da estrutura estratificacional de cada variedade. Com a sociolinguística



cognitiva e quantitativa, pretendemos contribuir para o desenvolvimento da sociolinguística e da sociolectometria baseada em *corpus* do Português de âmbito binacional (europeu e brasileiro) e estamos certos **de** que os resultados obtidos garantirão uma base empírica segura para a construção de melhores políticas da língua portuguesa, como língua plurinacional e transcontinental.

ABSTRACT: this paper presents elements of the emerging Cognitive Sociolinguistics and a program of cognitive and quantitative sociolinguistics for European and Brazilian Portuguese which focuses on the issue of convergence or divergence between the two national varieties. First, we show the reasons why cognitive sociolinguistics is inevitable within the paradigm of Cognitive Linguistics. Following that, we summarize the results of our sociolexicological research about the vocabulary of football and clothing and we point out the extents of this research for the fields of functional words and non-lexical variables.

Keywords: cognitive sociolinguistics; corpus linguistics; European/Brazilian Portuguese; language variation; social cognition

Notas explicativas:

Para algumas discussões críticas, ver o terceiro fascículo do volume 13 da revista *Cognitive Linguistics*, de 2002, particularmente a crítica de Rakova (2002), continuada em Rakova (2003), e as respostas de Johnson & Lakoff (2002); ver também Sinha & Jensen de López (2000), Martins (2003), Silva (2004c), Haser (2005) e, de outros pontos de vista, Hirose (2002), Riegler (2002) e Semin & Smith (2002).

² De referir o grupo liderado por Margarida Salomão, na Universidade Federal de Juiz de Fora (Brasil), cuja investigação se orienta por esta visão sociocognitiva da linguagem.

³ Trata-se do projecto “Convergência e Divergência no Léxico do Português”, financiado pela Fundação para a Ciência e a Tecnologia (Ref^a POCTI/LIN/48575/2002), no âmbito do POCTI do fundo comunitário europeu FEDER, com a duração de dois anos (2004 – 2006) e em fase de conclusão. Integram o grupo de investigação, para além do autor deste texto, José João Dias Almeida, Alberto Manuel Simões, Ana Margarida Abrantes, Marlene Danaia Duarte, José Luiz de Lucca e Ana Margarida Nunes.

⁴ Neste cálculo, a atribuição da característica A não é uma questão binária (presença ou ausência), mas um *continuum*: é atribuído o valor mais alto (1) aos estrangeirismos que mantêm a forma original e o valor mais baixo (0,25, por exemplo) a fortes adaptações e a decalques.

⁵ Termos *endógenos* são os termos típicos do PE ou do PB e termos *exógenos* são os termos típicos do PE presentes no PB ou do PB presentes no PE.

⁶ Em relação aos termos de futebol, é de notar uma tendência evolutiva nas duas variedades no sentido da diminuição da variação onomasiológica, como resultado de (i) integrações de estrangeirismos ou sua substituição por termos vernáculos, (ii) outras mudanças e (iii) ainda a massificação do futebol e “simplificação” da sua linguagem. Mas é o PB que mostra uma diminuição maior da variação onomasiológica. Isto poderá indicar maior standardização do vocabulário do futebol no PB nos últimos anos, a nível do registo formal da linguagem dos jornais de desporto. Mas poderá também (ou apenas?) indicar menor preocupação com o apuramento estilístico da língua nos jornais brasileiros de hoje, em comparação com os jornais portugueses actuais.

Referências

BARLOW, M. e KEMMER, S. (eds.). *Usage-based Models of Language*. Stanford: CSLI Publications. 2000.



- BERNÁRDEZ, E. Intimate enemies? On the relations between language and culture. In: SILVA, A. S. da, TORRES, A. e GONÇALVES, M. (orgs.), *Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva*. Vol. I. Coimbra: Almedina, 2004. p. 21-45.
- BERNÁRDEZ, E. Social cognition: variation, language, and culture in a cognitive linguistic typology. In: Ruiz de Mendoza, F. J. e Peña Cervel, S. (eds.), *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interaction*. Berlin/New York: Mouton de Gruyter, 2005. p 191-222.
- BIBER, D. *Variation across Speech and Writing*. Cambridge: Cambridge University Press. 1988.
- BIBER, D., CONRAD, S. e REPPEN, R.. *Corpus Linguistics. Investigating Language, Structure and Use*. Cambridge: Cambridge University Press. 1998.
- BLANK, A. e KOCH, P. (eds.). *Historical Semantics and Cognition*. Berlin/New York: Mouton de Gruyter. 1999.
- BRANDÃO, S. F. e MOTA, M. A. (orgs.). *Análise Contrastiva de Variedades do Português. Primeiros Estudos*. Rio de Janeiro: In-Fólio. 2003.
- BYBEE, J. e HOPPER, P. (eds.). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins. 2001.
- CROFT, W. e CRUSE, D. Alan. *Cognitive Linguistics*. Cambridge: Cambridge University Press. 2004.
- DAMÁSIO, A. *O Erro de Descartes. Emoção, Razão e Cérebro Humano*. Mem Martins: Publicações Europa-América. 1995.
- _____. *O Sentimento de Si. O Corpo, a Emoção e a Neurobiologia da Consciência*. Mem Martins: Publicações Europa-América. 2000.
- DIRVEN, R., FRANK, R. e PÜTZ, M. (eds.). *Cognitive Models in Language and Thought. Ideology, Metaphors and Meanings*. Berlin/New York: Mouton de Gruyter. 2003.
- DIRVEN, R., FRANK, R., ZINKEN, T. e ZLATEV, J. (eds.). no prelo. *Body, Language, and Mind*. Vol. 2. Berlin/New York: Mouton de Gruyter.
- EDELMAN, G. M. *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York: Basic Books. 1992.
- FINEGAN, E. e BIBER, D. (eds.). *Sociolinguistic Perspectives on Register*. New York: Oxford University Press. 1994.
- GEERAERTS, D. *Diachronic Prototype Semantics. A Contribution to Historical Lexicology*. Oxford: Clarendon Press. 1997.
- _____. Cultural models of linguistic standardization. In: Dirven, R., Frank, R. e Pütz, M. (eds.), *Cognitive Models in Language and Thought. Ideology, Metaphors and Meanings*. Berlin/New York: Mouton de Gruyter, 25-68. 2003 (este artigo foi publicado novamente na obra referida: Silva, A. S. da, Torres, A e Gonçalves, M.) (orgs.), *Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva*. Vol. I. Coimbra: Almedina, 47-84) 2004
- _____. Lectal data and empirical variation in Cognitive Linguistics. In: Ruiz de Mendoza, F. J. e Peña Cervel, S. (eds.), *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interactions*. Berlin/New York: Mouton de Gruyter, 163-189. 2005.
- _____. (ed.). *Cognitive Linguistics: Basic Readings*. Berlin/New York: Mouton de Gruyter. 2006.



- GEERAERTS, D., GRONDELAERS, S. e SPEELMAN, D. *Convergentie en Divergentie in de Nederlandse Woordenschat*. Amsterdam: Meertens Instituut. 1999.
- GIBBS, R. W. *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press. 2005.
- GRIES, Stefan Th. *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*. London: Continuum Press. 2003.
- GRIES, Stefan Th. e STEFANOWITSCH, A. (eds.). *Corpora in Cognitive Linguistics. Corpus-based Approaches to Syntax and Lexis*. Berlin/New York: Mouton de Gruyter. 2006.
- GRONDELAERS, S., SPEELMAN, D. e GEERAERTS, D. Regressing on *er*. Statistical analysis of texts and language variation. In: Morin, A. e Sébillot, P. (eds.), *6th International Conference on the Statistical Analysis of Textual Data*. Rennes: Institut National de Recherche en Informatique et en Automatique, 335–346. 2002.
- HARDER, P. The status of linguistic facts. Rethinking the relation between cognition, social institution, and utterance from a functional point of view. *Mind and Language* 18, 52-76. 2003.
- HASER, V. *Metaphor, Metonymy, and Experientialist Philosophy: Challenging Cognitive Semantics*. Berlin/New York: Mouton de Gruyter. 2005.
- HIROSE, N. An ecological approach to embodiment and cognition. *Cognitive Systems Research* 3, 289-299. 2002.
- HOPPER, P. J. e TRAUGOTT, E. C. *Grammaticalization*. Cambridge: Cambridge University Press. 1993.
- ITKONEN, E. *What is Language? A Study in the Philosophy of Linguistics*. Turku: Åbo Akademis tryckeri. 2003.
- JOHNSON, M. e LAKOFF, G. Why cognitive linguistics requires embodied realism? *Cognitive Linguistics* 13-3, 245-263. 2002.
- KÖVECSES, Z. *Metaphor and Emotion. Language, Culture, and Body in Human Feeling*. Cambridge-Paris: Cambridge University Press. 2000.
- KÖVECSES, Z. *Metaphor in Culture: Universality and Variation*. Cambridge: Cambridge University Press. 2005.
- KRISTIANSEN, G. How to do things with allophones: Linguistic stereotypes as cognitive reference points in social cognition. In: Dirven, R., Frank, R. e Pütz, M. (eds.), *Cognitive Models in Language and Thought: Ideologies, Metaphors, and Meanings*. Berlin/New York: Mouton de Gruyter, 69-120. 2003.
- KRISTIANSEN, G. e DIRVEN, R. (eds.). *Cognitive Sociolinguistics*. Berlin/New York: Mouton de Gruyter. 2006.
- LAKOFF, G. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press. 1987.
- _____. *Moral Politics: What Conservatives Know that Liberals Don't*. Chicago: The University of Chicago Press. 1996.
- _____. The brain's concepts. The role of the sensory-motor system in reason and language. *Working papers of the NTL Group at the International Computer Science Institute (ICSI)*. Berkeley: University of California, Berkeley. 2003.
- LAKOFF, G. e JOHNSON, M. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books. 1999.
- LANGACKER, R. W. *Foundations of Cognitive Grammar, Vol. 1: Theoretical Prerequisites*. Stanford: Stanford University Press. 1987.



- _____. A usage-based model. In: Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*. Amsterdam: John Benjamins, 127-161. 1988.
- _____. Subjectification. *Cognitive Linguistics* 1-1, 5-38. 1990.
- _____. *Foundations of Cognitive Grammar*, Vol. 2: *Descriptive Application*. Stanford: Stanford University Press. 1991.
- _____. Culture, cognition, and grammar. In: Pütz, M. (ed.), *Language Contact and Language Conflict*. Amsterdam: John Benjamins, 25-53. 1994.
- _____. *Grammar and Conceptualization*. Berlin/New York: Mouton de Gruyter.
- LANGACKER, Ronald W. 2000. A dynamic usage-based model. In: Barlow, M. e Kemmer, S. (eds.), *Usage-based Models of Language*. Stanford: CSLI Publications, 1-63. 1999.
- LEVINSON, S. C. *Space in Language and Cognition*. Cambridge: Cambridge University Press. 2003.
- MARTINS, H. Sobre linguagem e pensamento no paradigma experiencialista. *Veredas. Revista de Estudos Lingüísticos* 6, 75-90. 2003.
- PALMER, G. B. *Toward a Theory of Cultural Linguistics*. Austin: University of Texas Press. 1996.
- PERES, J. A. e KATO, M. A. (orgs.). *Studies in the Comparative Semantics of European and Brazilian Portuguese*. Issue of *Journal of Portuguese Linguistics* 3. 2004.
- RAKOVA, M. The philosophy of embodied realism: A high price to pay? *Cognitive Linguistics* 13-3, 215-244. 2002.
- _____. *The Extent of the Literal. Metaphor, Polysemy and Theories of Concepts*. New York, NY: Palgrave Macmillan. 2003.
- RIEGLER, A. When is a cognitive system embodied? *Cognitive Systems Research* 3, 339-348. 2002.
- SEMIN, G. R. e SMITH, E. R. Interfaces of social psychology with situated and embodied cognition. *Cognitive Systems Research* 3, 385-396. 2002.
- SHORE, B. *Culture in Mind: Cognition, Culture, and the Problem of Meaning*. Oxford: Oxford University Press. 1996.
- SILVA, A. S. da. Introdução: linguagem, cultura e cognição, ou a Linguística Cognitiva. In: SILVA, A. S. da, TORRES, A. e GONÇALVES, Miguel (orgs.), *Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva*. Vol. I. Coimbra: Almedina, 1-18. 2004a.
- _____. Léxico e variação Portugal/Brasil: Para uma sociolexicologia cognitiva do Português. *Revista Portuguesa de Humanidades* 8, 99-117. 2004b.
- _____. Protótipos, imagens e metáforas, ou o experiencialismo da linguagem e do pensamento. In: DINIS, A. e CURADO, J. M. (orgs.), *Consciência e Cognição*. Braga: Publicações da Faculdade de Filosofia da Universidade Católica Portuguesa, 79-96. 2004c.
- _____. Para o estudo das relações lexicais entre o Português Europeu e o Português do Brasil: Elementos de sociolexicologia cognitiva e quantitativa do Português. In: DUARTE, I. e LEIRIA, I. (orgs.), *Actas do XX Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: Associação Portuguesa de Linguística, 211-226. 2005a.
- _____. Palavras e conceitos no tempo: Para uma onomasiologia diacrónica e cognitiva do Português. In: RIO-TORTO, G. M.; Figueiredo, O. e SILVA, F. (orgs.),



Estudos em Homenagem ao Professor Doutor Mário Vilela. Porto: Faculdade de Letras do Porto, 121-139. 2005b.

_____. Gramática, cognição e contexto: descontextualizar ou recontextualizar? In: GONÇALVES, M., SILVA, A. S. da, COUTINHO, J., MARTINS, J. C. e FERREIRA, M. J. (orgs.), *Gramática e Humanismo. Colóquio de Homenagem a Amadeu Torres*. Vol. I. Braga: Faculdade de Filosofia da Universidade Católica Portuguesa, 665-685. 2005c.

_____. O léxico do futebol no Português Europeu e no Português Brasileiro: convergência ou divergência? *Diacrítica. Ciências da Linguagem* 20-1, 167-196. 2006.

SILVA, A. S. da e DUARTE, M. D. O léxico do vestuário no Português Europeu e no Português Brasileiro: convergência ou divergência? *Revista Portuguesa de Humanidades* 9, 117-136. 2005.

SINHA, C. Grounding, mapping, and acts of meaning. In: JANSSEN, T. e REDEKER, G. (eds.), *Cognitive Linguistics: Foundations, Scope, and Methodology*. Berlin/New York: Mouton de Gruyter, 223-255. 1999.

SINHA, C. e LÓPEZ, JENSEN DE, Kristina. Language, culture, and the embodiment of spatial cognition. *Cognitive Linguistics* 11, 17-41. 2000

SPEELMAN, D., GRONDELAERS, S. e GEERAERTS, D. Profile-based linguistic uniformity as a generic method for comparing language varieties. *Computers and the Humanities* 37, 317-337. 2003.

STEFANOWITSCH, A. e GRIES, Stefan Th. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8, 209-243. 2003.

_____. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1, 1-43. 2005.

TALMY, L. *Toward a Cognitive Semantics*. 2 vols. Cambridge: The MIT Press. 2000.

TOMASELLO, M. *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press. 1999.

_____. First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11, 61-82. 2000.

_____. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press. 2003.

TRAUGOTT, E. C. e DASHER, R. B. *Regularity in Semantic Change*. Cambridge: Cambridge University Press. 2002.

TUMMERS, J., HEYLEN, K. e GEERAERTS, D. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1, 225-261. 2005.

VARELA, F., THOMPSON, E. e ROSCH, E. *The Embodied Mind*. Mass.: Massachusetts Institute of Technology. 1991.

VIERECK, W. English Dialectology and Dialectometry. In: KLEMOLA, Juhani; KYTO, M. e RISSANEN, M. (eds.), *Speech Past and Present: Studies in English Dialectology in Memory of Ossi Ihalainen*. Frankfurt am Main: Peter Lang, 447-469. 1996.

ZIEMKE, T.; ZLATEV, J. e FRANK, R. (eds.). no prelo. *Body, Language, and Mind. Vol. 1: Embodiment*. Berlin/New York: Mouton de Gruyter.



ZLATEV, J. *Situated Embodiment: Studies in the Emergence of Spatial Meaning*. Stockholm: Gotab. 1997.

_____. Situated embodied semantics and connectionist modeling. In: ALLWOOD, J. e GARDENFORS, P. (eds.), *Cognitive Semantics: Meaning and Cognition*. Amsterdam: John Benjamins, 173-194. 1999.

_____. Beyond cognitive determination. Interactionism in the acquisition of spatial semantics. In: LEATHER, J. e VAM DAM, J. (eds.), *Ecology of Language Acquisition*. Amsterdam: Kluwer Academic Publishers, 83-107. 2003.