

Publicado como:

BATORÉO, Hanna Jakubowicz & Augusto Soares da SILVA (2012). Estudar o português como língua pluricêntrica no enquadramento da Linguística Cognitiva com foco nas variedades nacionais do PE e PB". In: TEIXEIRA E SILVA, Roberval; YAN, Qiarong; ESPADINHA, Maria Antónia; LEAL, Ana Varani. (eds.) 2012. IISIMELP: A formação de novas gerações de falantes de português no mundo. China, Macau: Universidade de Macau. CD-ROOM.ISBN: 978-99965-1-035-9.

ESTUDAR O PORTUGUÊS COMO LÍNGUA PLURICÊNTRICA NO ENQUADRAMENTO DA LINGUÍSTICA COGNITIVA COM FOCO NAS VARIEDADES NACIONAIS DO PE E PB

Versão pré-final

Hanna Jakubowicz BATORÉO*

Augusto Soares da SILVA*

RESUMO

Nos últimos anos, tem-se vindo a observar, no âmbito de Linguística Cognitiva, a emergência do estudo da variação intralinguística, dando origem à especificidade de modelos e métodos da Sociolinguística Cognitiva (Kristiansen & Dirven 2008). Seguindo os pressupostos teóricos deste enquadramento, centramos o nosso estudo sobre o português perspetivado como uma língua pluricêntrica (Clyne 1992, Silva *et al.* 2011) e focando-o no estudo contrastivo de variáveis lexicais e gramaticais observadas no Português Europeu (PE) e no Português do Brasil (PB). O artigo abrange três partes: (i) o estudo da produtividade lexical no PE e PB (Batoréo & Casadinho 2009a, 2010, Batoréo 2011b); (ii) estudo de convergência e divergência entre o PE e PB com base na análise socioletométrica de *corpus* de Silva (2005, 2006, 2008, 2010, 2011, no prelo); (iii) estudo de percepção e atitudes no que diz respeito à convergência e divergência das variedades nacionais do português, com base no estudo de Silva (no prelo).

PALAVRAS CHAVE: *Linguística Cognitiva; Sociolinguística Cognitiva; Português Europeu (PE); Português do Brasil (PB); variação linguística; língua pluricêntrica*

INTRODUÇÃO

A Linguística Cognitiva estuda a linguagem como veículo de conhecimento em conexão com a experiência humana do mundo – um *modelo baseado no uso* –,

* Universidade Aberta, Departamento de Humanidades, Lisboa, Portugal, hanna@univ-ab.pt

* Universidade Católica Portuguesa, Faculdade de Filosofia, Braga, Portugal, assilva@braga.ucp.pt

analisando as unidades e as estruturas da linguagem como manifestações das capacidades cognitivas gerais, da organização conceptual e de princípios de categorização, assim como de mecanismos de processamento e de experiência cultural, social e individual (cf. Batoréo 2004; Silva e Batoréo 2010). Por ser um modelo baseado no uso, a Linguística Cognitiva dedica-se ao estudo da variação linguística, utilizando para tal os seguintes instrumentos de análise: prototipicidade, metáfora, metonímia, corporização/corporificação, perspetivação conceptual, subjetificação, bem como construção de modelos cognitivos e culturais.

Por conseguinte, e de um modo mais restrito no âmbito da recente Sociolinguística Cognitiva (Kristiansen & Dirven, 2008; Geeraerts, Kristiansen & Peirsman, 2010; Silva, 2009, para uma síntese), nos últimos anos, tem-se vindo a observar a emergência e o estudo da variação intralinguística, incluindo o estudo da variação em línguas pluricêntricas – no sentido institucionalizado por Clyne (1992) de línguas com diferentes variedades nacionais, cada qual com a sua norma própria (cf. Silva *et al.* 2011) –, a partir dos modelos descritivos da Linguística Cognitiva e perspetivando as variedades linguísticas como construções e representações sociocognitivas. Pretende-se, deste modo, verificar se as diferenças linguísticas locais e nacionais refletem diferenças culturais, assim como analisar até que ponto essas diferenças se correlacionam com diferenças cognitivas.

Seguindo a perspetiva atrás traçada, o estudo cognitivo do pluricentrismo linguístico caracteriza-se por explorar as interrelações entre as dimensões socioculturais, conceptuais e estruturais da variação e da mudança em línguas com diferentes variedades nacionais, perspetivadas como construções sociocognitivas.

Na agenda sociocognitiva do estudo das línguas pluricêntricas, destacam-se os seguintes tópicos (ver a introdução de Silva, 2011): (i) correlação entre diferenças linguísticas nacionais e diferenças conceptuais e culturais; (ii) correlação entre variáveis lexicais, gramaticais e fonológicas e entre estas variáveis e as variáveis sociolinguísticas e estilísticas na convergência/divergência e na estratificação de variedades nacionais; (iii) modelos cognitivos e culturais de perceção e avaliação atitudinal das variedades nacionais; (iv) correlação entre distâncias linguísticas nacionais *objetivas* (observadas no comportamento linguístico dos falantes) e distâncias linguísticas *subjetivas* (distâncias percecionadas, atitudes linguísticas); (v) atitudes de purismo e independentismo e o seu impacto no desenvolvimento de variedades nacionais; (vi) planeamento e políticas multilaterais de língua; (vii) métodos empíricos e quantitativos

avançados, incluindo técnicas letométricas, capazes de dar conta da multidimensionalidade da variação linguística nacional.

O português é uma língua prototipicamente pluricêntrica, com duas normas nacionais bem estabelecidas (PE e PB) e outras normas em desenvolvimento. Sabemos que há diferenças entre o PE e o PB a todos os níveis da estrutura da língua. Contudo, sabemos pouco sobre a relação evolutiva entre PE e PB nos últimos anos: os estudos comparativos mais desenvolvidos (Brandão & Mota 2003, Peres & Kato 2004) não tratam da questão da convergência e divergência.

O presente artigo organiza-se em três partes: (i) na seção 1, será abordado o tema da produtividade lexical de um verbo, com base na análise de Batoréo & Casadinho (2009a, 2010) e Batoréo (2011b); (ii) na seção 2, será apresentado o estudo de convergência e divergência entre o PE e PB com base na análise socioletométrica de *corpus* desenvolvida por Silva (2005, 2006, 2008, 2010, 2011, no prelo); (iii) na seção 3, serão apresentados indicadores percetivos e atitudinais de convergência e divergência entre PE e PB com base no estudo de Silva (no prelo).

1. Produtividade lexical em PE e PB: estudo de um verbo (Batoréo & Casadinho, 2009b e 2010; Batoréo 2011a)

Na presente secção, é nosso objectivo demonstrar, com base no estudo de um item lexical – do verbo ‘*botar*’, bem como da palavra dele derivada ‘*botafogo*’ –, como funciona a produtividade lexical nas duas variantes nacionais do português dos dois lados do Atlântico, no PE e PB.

Tal como defendido pela Linguística Cognitiva, o estudo das línguas engloba tanto a perspectiva sincrónica como a diacrónica, sendo que a mudança semântica resulta da polissemia e reflete os efeitos de prototipicidade. Assim, por exemplo, a mudança na *intensão* de uma categoria lexical – seja ela, por exemplo, um nome comum ou um nome próprio – não envolve significados individuais e isolados, mas o conjunto de significados sincronicamente agrupados por *semelhanças-de-família*; é este conjunto que sofre alterações quer na forma de desenvolvimento a partir de vários significados coexistentes quer no desaparecimento de valores periféricos quer, ainda, na forma de reorganizações de protótipos. Por conseguinte, a não-discrição *intensional* determina a natureza *enciclopédica* da mudança semântica, o que significa que a mudança pode

envolver qualquer informação associada a um determinado item lexical, podendo surgir um novo significado tanto com base num já existente como com base num subconjunto deste, seja ele de carácter referencial, pragmático ou conotativo.

Os mecanismos conceptuais mais frequentes utilizados na categorização são a metáfora e a metonímia conceptuais, mecanismos e *modelos cognitivos* generalizados, convencionalizados e lexicalizados; além disso, existem também outros processos implicados na categorização, tais como a generalização ou a abstração, isto é, a *esquematisação*. Por conseguinte, categorização por *esquema* e categorização por *protótipo* não se excluem; antes se complementam. Na semântica de um item lexical, há sentidos ou referentes esquemáticos e específicos, prototípicos e periféricos. A sua estrutura tem, por isso, a forma de uma rede, sendo o modelo de representação da estrutura das categorias que combina esquemas e protótipos um *modelo em rede*.

Por conseguinte, na ótica da Linguística Cognitiva, a produtividade lexical de uma língua é determinada pelos (i) padrões de formação lexical nela utilizados, bem como pela sua operatividade e estabilidade, (ii) pelo tipo de processos de categorização (por protótipo e por esquema) e de processos cognitivos – principalmente metáfora e metonímia – disponibilizados a nível da neologia e dos processos da inovação neológica, assim como (iii) pelos graus de utilização e operacionalização lexical em diferentes níveis diatópicos, diastráticos e/ou diafásicos dos itens lexicais em variantes linguísticas da mesma língua. Nestas variantes, os referidos fatores podem tornar-se operacionais de modos e graus diferenciados, com ou até sem paralelismos.

Para exemplificar os processos atrás referenciados (e conforme demonstrámos em Batoré & Casadinho, 2009b), o verbo ‘*botar*’ foi, ao longo dos séculos, um verbo espacial utilizado com o sentido de *pôr* e *colocar*, conforme atestam os textos de várias épocas, tais como, entre outros, os contos e lendas populares, os textos literários (por exemplo, de Gil Vicente ou de Fernão Mendes Pinto), etc. Este item deu também origem aos verbos utilizados hoje nos crioulos de base lexical portuguesa, tal como se pode observar, ainda, no crioulo de Malaca ou no(s) crioulo(s) são-tomense(s).¹

Hoje em dia, no entanto, o uso do verbo ‘*botar*’ é muito diferenciado nas duas variantes nacionais *standard* do português. No PB, o verbo é produtivo e frequente na

¹ Vejam-se os exemplos do Português de Malaca, em que ‘*botá*’ quer dizer *botar*, *lançar*, *aplicar*, *colocar* e ‘*botado*’ significa *atirado*, *estendido*. No léxico são-tomense, por outro lado, existe o verbo ‘*buta*’, no sentido de *lançar*, *pôr*, bem como diversas expressões deverbais de que este verbo faz parte, atribuindo-lhes carácter incoativo ou causativo.

linguagem do dia-a-dia com o sentido de *colocar*, *deitar*, *pôr*, embora haja tendência para o substituir pelo sinónimo contextual ‘*colocar*’, quando se pretende que o uso não seja rotulado como popular (hipercorreção). Pelo contrário, no PE padrão, não surgem ocorrências do verbo ‘*botar*’ na linguagem do dia-a-dia no sentido observado no PB, sendo substituídas pelos seus sinónimos contextuais ‘*deitar*’, ‘*pôr*’ ou ‘*colocar*’. O verbo ‘*botar*’ surge pontualmente apenas nas lexias com um certo grau de fixidez, em que se observam usos não literais, mas preferencialmente metafóricos. Verificámos, assim, basicamente a ocorrência de três paradigmas em que o verbo pode surgir: (i) no sentido de *falar*, *opinar*, *discursar*: ‘*botar discurso*’, ‘*botar faladura*’, ‘*botar palavra*’, ‘*botar opinião*’, ‘*botar sentença*’; (ii) no sentido de *atribuir o nome de X*: ‘*botar + [nome de X]*’ e nas expressões (iii) ‘*botar para cima*’ / ‘*botar para baixo*’.

As observações atrás apresentadas permitem-nos defender que, do ponto de vista diatópico, existe um grande desnível no emprego do verbo ‘*botar*’ nas duas variantes nacionais *standard* da Língua Portuguesa: este verbo é frequente e corrente no PB, mas raro e específico no PE, estando confinado sobretudo a usos fixos e metafóricos na língua padrão, bem como a regionalismos. No PE dialectal, o ‘*botar*’ mantém ainda certa produtividade, sobretudo nos dialectos setentrionais, por exemplo, na Beira Interior, no sentido corrente de ‘*colocar*’. Assim, na gíria covilhanense (Covilhã, Beira Interior), o verbo ‘*botar*’ significa *deitar*, *vazar* ou, então, *mandar embora*, como em “*Ela meteu-se lá dentro, mas botaram-na fora*”. Em Trás-os-Montes, o ‘*botar*’ pode surgir na expressão ‘*botar de baras*’, no sentido de *supor*, *prever*, como em “*Botei de baras que chegarias mais cedo e enganei-me*”. Nos dialetos das ilhas, o ‘*botar*’ surge também com sentidos específicos: na Madeira, refere uma determinada forma do acto de *bordar*, enquanto nos Açores significa *ir em direção a* / *ir dar a*.

É de destacar que a análise das contagens feitas com base na ocorrência do item em *subcorpora* do PE e do PB consultados no âmbito da *Linguateca* (cf. Batoréo e Casadinho 2010) permite verificar o que se entende pela *frequência de uso*, quando se refere que um item lexical é *mais frequente* do que o outro. No âmbito da *Linguateca*, foram consultados dois *subcorpora*: o *CETEMPUBLICO*, da variante do PE, com mais de 232 milhões de ocorrências e o *NILC/SãoCarlos*, da variante do PB, com mais de 42 milhões de ocorrências. A análise comparativa entre os dois *subcorpora* (tendo inclusive em conta a diferença relativa entre os seus tamanhos) permitiu verificar que o verbo ‘*botar*’ é 141,3 vezes mais frequente nos usos correntes da variante *standard*

brasileira do que na variante padrão europeia, o que explica e consubstancia as intuições e observações que dizem respeito às diferenças entre as duas variantes nacionais.

É interessante verificar que os falantes do português – tanto do PE como do PB – pouca ou nenhuma consciência têm das enormes diferenças do uso observadas entre as duas variantes-padrão. Se, em Portugal, o verbo ‘*botar*’ é remetido para a especificidade do emprego dos falantes brasileiros, por um lado, e por outro, para os usos regionais e populares (o que tem o seu eco até a nível lexicográfico), no Brasil parece não existir a consciência da diferença da produtividade deste item lexical nos dois lados do Atlântico. Aliás, no Brasil, além de ser muito produtivo na oralidade, o verbo faz parte de muitas expressões idiomáticas, tais como: ‘*botar as barbas de molho*’, ‘*botar as cartas na mesa*’, ‘*botar (pôr) lenha na fogueira*’, casos nos quais, em Portugal, este verbo é substituído por ‘*pôr*’ ou um sinónimo. Por outro lado, existe, no Brasil, uma certa consciência do carácter popular do verbo, o que pode levar os falantes a assumirem reacções de hipercorreção, substituindo-o por um dos seus sinónimos contextuais como ‘*colocar*’.

Uma das expressões historicamente documentadas com o verbo ‘*botar*’ é ‘*botar fogo*’, dando origem a uma nova palavra: o ‘*bota-fogo*’ e, seguidamente, o ‘*botafogo*’ (Batoréo 2011), criando uma rede de sentidos formados sucessivamente sobretudo por metonímia, mas também por metáfora, ao longo dos séculos (cf. Quando 3 e 4 na *op. cit.*). Inicialmente, o item ‘*bota-fogo*’ designava (i) um instrumento militar, uma haste com um pavio, com a qual o artilheiro detonava os canhões, a partir dos quais se formavam as respetivas profissões. Foi esta designação que – por metonímia – deu origem ao nome atribuído a um famoso (ii) galeão conhecido pelo seu poder de fogo, robustez e eficácia. A partir daí, muitos dos fidalgos portugueses passaram a utilizar este nome por metáfora como (iii) título nobre, tal como aconteceu com (iv) um oficial de artilharia que – por metonímia – ganhou o apelido de ‘*Botafogo*’ e incorporou-o ao seu sobrenome. Galardoado com (v) terras situadas na área da baía da Guanabara pelos seus feitos heróicos, passou-lhes – por metonímia – o seu novo nome, batizando, assim, (v-a) uma enseada, (v-b) uma praia e, posteriormente, (v-c) um bairro do Rio de Janeiro. A partir das actividades desportivas desenvolvidas nesta praia teve – por metonímia – a sua origem (vi) o nome do conhecido clube desportivo carioca de *Botafogo de Futebol e Regatas*, sendo os seus associados, jogadores ou adeptos conhecidos por *botafoguenses*.

Os processos acima especificados e envolvidos na estruturação conceptual de ‘*bota(-)fogo*’ podem ser representados em forma de uma rede de conceptualizações e

interpretações que constituem uma estrutura aberta sujeita a possíveis extensões futuras. Basta imaginarmos, por exemplo, que o clube carioca *Botafogo* publique um jornal desportivo – por metonímia – com o mesmo nome ou, por outro lado, se torne ainda mais conhecido por causa de um jogador que queira – também por metonímia – adotar o nome do clube como seu próprio apelido. Tendo em consideração esta possível extensão semântica da rede já existente, consideramo-la como um esquema produtivo e sempre em aberto.

A análise da linguagem jornalística do Portugal contemporâneo com base na *Linguateca* permite-nos verificar que o uso dos itens escolhidos ‘*fogo*’ e ‘*bota(-)fogo*’ no PE, bem como a respetiva frequência estão bastante afastados dos sentidos apresentados e analisados com base nas aceções atestadas nos dicionários. Os dados reunidos mostram que o conhecimento ativo dos falantes do PE ilustrado a nível de linguagem jornalística se limita aos usos sincrónicos do nome próprio relativo à designação do clube de futebol ou, em menor grau, à designação toponímica de um bairro carioca. Os usos da palavra comum ‘*bota-fogo*’ estão dele totalmente ausentes. Com base nestes resultados podemos concluir que a riqueza polissémica do item ‘*bota-fogo*’ na qualidade do nome comum está totalmente caído em desuso, tal como, por exemplo, o conhecimento do nome próprio quer no seu sentido histórico da designação do galeão quer no seu sentido antropomórfico, permanecendo vivos apenas os usos relativos à atividade futebolística ou a designações topográficas.

Na sequência dos estudos acima referidos dedicados a redes conceptuais na arquitetura da semântica lexical, com base no exemplo da análise dos itens lexicais ‘*botar*’ e ‘*bota(-)fogo*’, demonstrou-se que estamos perante um sistema de estruturação conceptual, que envolve quer as capacidades cognitivas gerais quer os conhecimentos que temos sobre o mundo (integrando, assim, uma semântica enciclopédica, que abrange tanto os nomes comuns como os nomes próprios) quer, ainda, mecanismos do imaginário como metáfora e metonímia conceptuais. O significado surge, por conseguinte, não como objeto mental estável, mas resulta da construção de rede(s) de conceptualizações e interpretações, efetuada de acordo com critérios precisos e linguisticamente bem definidos no âmbito da Linguística Cognitiva, mas diferenciada em função dos factores específicos de cada uma das culturas nacionais em que se desenvolve, originando, assim, as especificidades no PE e no PB.

O significado estudado com base nos exemplos de uso linguístico demonstra que os usos sincrónicos comprovados pelos *corpora* do PE de hoje na reflectem as redes

conceptuais histórica e socialmente determinadas. Os empregos da linguagem-em-uso privilegiam certas aceções como mais frequentes, deixando outras fora do uso, com emprego delimitado diatópica e diastraticamente, e/ ou com frequência reduzida. Deste modo, o conhecimento da estruturação das conceptualizações dos itens lexicais permite-nos conhecer também a estrutura cognitiva do português que os nossos usos atestados hoje em dia podem, por vezes, nem sempre deixar transparecer.

2. Convergência e divergência entre PE e PB: análise socioletométrica de *corpus* (Silva 2005, 2006, 2008, 2010, 2011, no prelo).

Desenvolvemos um projeto de investigação centrado na questão diacrónica da convergência e divergência entre o português europeu (PE) e o português brasileiro (PB) nos últimos 60 anos (Silva 2005, 2006, 2008, 2010, 2011, no prelo). O projeto utiliza métodos avançados de análise de corpus, designadamente métodos socioletométricos para medir convergência e divergência entre variedades de uma língua.

O projeto situa-se no domínio de investigação sobre línguas *pluricêntricas*, no sentido institucionalizado por Clyne (1992: 1), já referido na *Introdução*. No contexto da investigação sobre o pluricentrismo, a especificidade do presente projeto reside em dois aspetos. Primeiro, centramo-nos na variação do significado, pouco estudada em Sociolinguística, e na interrelação entre os aspetos sociais e os aspetos conceptuais da variação pluricêntrica. Deste modo, o projeto inscreve-se no quadro da Sociolinguística Cognitiva (ver *Introdução*). Segundo, utilizamos métodos socioletométricos que permitem medir distâncias linguísticas e correlacioná-las não só com distâncias espaciais mas também com todos os tipos de variáveis sociolinguísticas. Especificamente, aplicamos a socioletometria desenvolvida pela unidade belga de Lexicologia Quantitativa e Linguística Variacional para o neerlandês como língua pluricêntrica (Geeraerts, Grondelaers & Speelman 1999).

O projeto incide na variação *onomasiológica* entre sinónimos denotacionais (lexicais e construcionais), isto é, expressões referencialmente equivalentes e utiliza métodos socioletométricos baseados em *perfis onomasiológicos*, entendidos como conjuntos de termos/construções alternativos juntamente com as suas frequências. A variação onomasiológica que envolve sinónimos denotacionais como, por exemplo, *avançado* e *atacante* para o conceito ‘avançado’ (o primeiro termo é mais usado no PE,

ao passo que o segundo é mais usado no PB) – ou *variação onomasiológica formal*, em contraste com a *variação onomasiológica conceptual*, a qual envolve diferenças conceptuais, como *avançado* e *jogador* (o primeiro termo é mais específico, ao passo que o segundo é o termo taxionomicamente hiperonímico) – é particularmente interessante do ponto de vista sociolinguístico porque a escolha entre sinónimos denotacionais dá pistas sobre as relações entre variedades linguísticas.

São utilizadas duas técnicas socioletométricas: medidas de uniformidade e medidas do impacto de traços linguísticos. A *uniformidade* é a medida da (di)similaridade entre os perfis onomasiológicos nas diferentes variedades linguísticas e é calculada somando as frequências relativas mais pequenas de cada palavra ou construção alternativa. Convergência e divergência entre duas variedades traduzem-se em aumento ou diminuição da uniformidade, respetivamente. A medida de uniformidade interna calcula a uniformidade dentro de uma variedade linguística: essa uniformidade diminui à medida que mais termos/construções competirem entre si para designar o mesmo conceito/função e/ou à medida que mais termos/construções dominantes existirem dentro de um perfil onomasiológico. As medidas de traços calculam a proporção de termos/construções com determinada característica (por exemplo, endo/exogeneidade ou estrangeirismo) no perfil onomasiológico de um conceito/função. Estes métodos implicam um vasto corpus e milhares de observações do uso linguístico.

A análise baseia-se no *corpus* CONDIVport, atualmente com 4 milhões de palavras do registo formal e 15 milhões do registo informal, disponibilizado pela *Linguateca* (Silva 2008b). O corpus está estruturado na base de três variáveis: (i) geográfica, com textos de Portugal e do Brasil (São Paulo e Rio de Janeiro); (ii) diacrónica, das décadas de 50, 70 e 2000; e (iii) estilística, do registo formal de jornais e revistas e do registo informal de *chats* da internet.

A investigação sociolexicológica já realizada (ver, sobretudo, Silva 2010) com base em 43 conceitos nominais – 21 perfis onomasiológicos do campo lexical do futebol e 22 perfis onomasiológicos do campo lexical do vestuário – e em largos milhares de observações do uso destes termos (183 termos de futebol e 264 de vestuário) no corpus CONDIVport (jornais de desporto e revistas de moda, *chats* associados a clubes de futebol e etiquetas de roupas de lojas de vestuário) permite concluir que a hipótese da divergência entre PE e PB se confirma no campo lexical do vestuário, mas não no do futebol. Os termos de vestuário são mais representativos do vocabulário comum e, por isso, os resultados do vestuário estarão mais próximos da realidade sociolinguística; a

ligeira convergência no campo do futebol será um efeito da globalização e da padronização do vocabulário do futebol.

As Figuras 1 e 2 sistematizam os resultados da análise sociolexicológica. São apresentadas as percentagens de uniformidade no *corpus* de futebol (90.202 observações) e no *corpus* de vestuário (12.451 observações) nos três períodos estudados (1950-1970-2000). O primeiro número de cada par de resultados representa a uniformidade não ponderada (U) e o segundo a uniformidade ponderada (U'), sendo U' mais significativo do que U; os números nas linhas horizontais indicam a percentagem de uniformidade entre PE e PB nos anos 50, 70 e 2000; os números nas linhas verticais e diagonais indicam a percentagem de uniformidade de período para período e entre diferentes períodos; e os números associados a cada variedade e período indicam a percentagem de uniformidade interna.

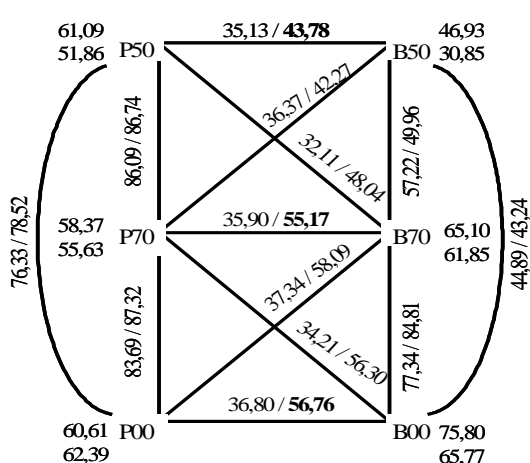


Figura 1. Termos de futebol

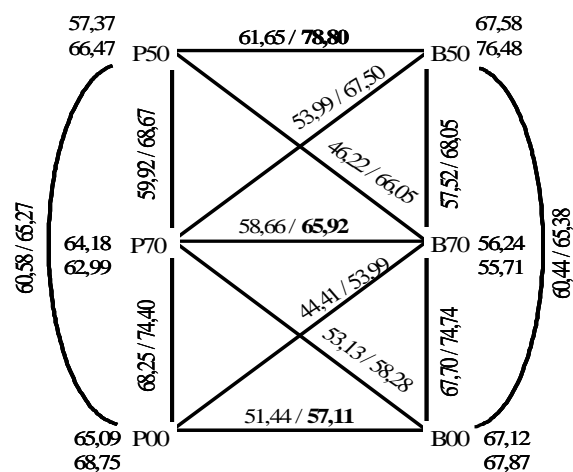


Figura 2. Termos de vestuário

Relativamente aos termos de futebol, as percentagens de U' mostram ligeira convergência de 50 para 70 (U' aumenta 11,39%) e estabilidade a partir de então, mas as de U não indicam qualquer alteração significativa. Ao mesmo tempo, os números sugerem uma grande distância entre as duas variedades nos três períodos. Quanto aos termos de vestuário, as percentagens de U e U' evidenciam divergência ao longo do tempo: U' diminui 12,88% de 1950 para 1970 e 8,81% de 1970 para 2000.

As linhas verticais e diagonais mostram que não há nenhuma orientação específica de uma variedade em relação à outra: PE e PB divergem um do outro no vocabulário do vestuário; as maiores mudanças do PB no vocabulário de futebol não significam uma aproximação ao PE e a influência do PB no PE no vocabulário do futebol é menor do

que o que se esperava. Um terceiro resultado, expresso nas percentagens de uniformidade interna, evidencia maiores mudanças no PB em ambos os campos lexicais. Um quarto resultado confirma que a influência estrangeira do inglês e de outras línguas é maior na variedade brasileira: o PB importa um maior número de estrangeirismos e adapta-os e integra-os mais facilmente do que o PE. Como último resultado, os termos de vestuário confirmam a hipótese de uma distância sincrónica maior entre estrato padrão e estrato subpadrão no PB do que no PE: a uniformidade entre os dados das revistas de moda de 2000 e os dados das etiquetas de lojas de vestuário é de 57,94% no PB e de 70,99% no PE. No vocabulário do futebol, não há esta assimetria entre os dados dos jornais de 2000 e os dados dos *chats*: a uniformidade é de 80,93% no PE e 78,76% no PB, o que sugere padronização dos termos de futebol na linguagem da internet.

Como extensão do estudo anterior, pretendemos investigar até que ponto variáveis lexicais e não lexicais se correlacionam como indicadores de convergência/divergência entre PE e PB. Analisaremos 15 variáveis morfológicas e sintáticas. Dividimo-las em três grupos. Primeiro, construções preposicionais como *falar de/em/sobre/acerca de e interesse em/por*, situadas a meio caminho entre as variáveis lexicais e as não lexicais. Segundo, variantes morfológicas como *nós* vs. *a gente* como pronomes de 1.pl, *seu* vs. *dele* como possessivo de 3.sg, variação de concordância nominal e verbal com *a gente* e alternâncias sufixais, como *-ção* vs. *-mento* em substantivos deverbais. Terceiro, diversas construções sintáticas: três construções de complemento infinitivo com verbos causativos e percetivos (*A Maria fez/deixou/mandou/viu os miúdos correrem* vs. *correr* vs. *correr os miúdos*), alternância entre infinitivo simples e flexionado (*Incentivou os trabalhadores a fazer* vs. *fazerem greve*), alternância (in)transitiva (*O carro furou o pneu* vs. *O pneu do carro furou*), construções intransitivas com/sem *se* (*O departamento reuniu-se* vs. *reuniu*), presença/ausência de preposição em construções relativas preposicionadas e alternância com *onde* (*não há lugar em que/que/onde a polícia não entre*), alternância entre sujeito preenchido e sujeito nulo, variação da ordem de palavras em construções nome-adjetivo (*ideia falsa* vs. *falsa ideia*) e alternância entre SV e VS.

Num estudo piloto, foram analisadas as seguintes variáveis não lexicais:

(i) 10 perfis preposicionais (1.730 observações), como *falar de/sobre/acerca de/em*, N_{mental} *acerca de/sobre, interesse em/por; precisar, necessitar de/Ø*, etc.;

(ii) 3 construções de complemento infinitivo dos verbos causativos *fazer, mandar, deixar* e percetivos *ver, ouvir, sentir* (1.385 observações): VSV, VOV e VV;

(iii) 2 construções adjetivas com *verdadeiro, falso, bonito, lindo, recente*: N + A e A + N (763 observações).

Os resultados mostram que as construções preposicionais (1), infinitivas (2) e adjetivas (3) replicam os resultados dos termos de vestuário, confirmando assim a hipótese da divergência:

- (1) $U'(P50,B50) 68,74\% > U'(P70,B70) 54,39\% > U'(P00,B00) 48,64\%$
- (2) $U'(P50,B50) 72,62\% > U'(P70, B70) 56,51\% > U'(P00,B00) 50,33\%$
- (3) $U'(P50,B50) 95,83\% > U'(P70, B70) 94,54\% > U'(P00,B00) 83,73\%$

A leitura de (1)-(3) é igual à das Figuras 1 e 2, apresentadas acima: por exemplo, (1) a uniformidade ponderada (U') entre as variedades portuguesa e brasileira nos anos 50 é maior do que a mesma uniformidade nos anos 70 e esta, por sua vez, é maior do que a mesma uniformidade nos anos 2000.

3. Convergência e divergência entre PE e PB: percepção e atitudes (cf. Silva no prelo)

Verifiquemos agora se existe uma correlação entre os indicadores *objetivos* extraídos do *corpus* (cf. Seção 2) e os indicadores *subjetivos* expressos nas atitudes dos falantes (para mais desenvolvimento, ver Silva no prelo). Através de um inquérito dirigido a 120 estudantes portugueses e brasileiros, compreendendo uma componente comportamental e uma componente cognitiva sobre o uso e a origem de termos de vestuário (15 perfis onomasiológicos), podemos verificar que as intenções atitudinais dos informantes vão no mesmo sentido divergente que observámos no corpus de vestuário. A proporção dos termos que os informantes consideram como endógenos, isto é, típicos da sua variedade nacional, e binacionais, usados em ambas as variedades nacionais, isto é, a proporção da *atitude divergente* (93,94% para os informantes do PE e 96,43% para os informantes do PB) é bastante mais elevada do que a proporção dos termos considerados como exógenos (típicos da outra variedade nacional) e binacionais, isto é, a proporção da *atitude convergente* (49,60% para os informantes do PE e 46,52% para os informantes do PB):

(4)PE: 93,94% (t. endógenos+binacionais) > 49,60% (t. exógenos+binacionais)

(5)PB: 96,43% (t. endógenos+binacionais) > 46,52% (t. exógenos+binacionais)

Um outro estudo atitudinal tem por objeto a literatura recente portuguesa e brasileira sobre política de língua e questões de padronização linguística (Bagno 2001, Faraco 2001, Coelho 2005, Gama 2007, entre outros). Seguindo a distinção de Geeraerts (2003) entre modelos cognitivos *românticos* e *racionalistas* da variação linguística, podemos identificar quatro atitudes convergentes/divergentes românticas e racionalistas sobre as relações entre PE e PB. As atitudes mais radicalmente convergentes e divergentes são *românticas* e são mais frequentes no Brasil. A atitude *convergente romântica* exprime-se, no Brasil, num normativismo conservador e dogmático (impondo um padrão decalcado do PE literário), na ideologia da exclusão social e na ideia de que “*o PE é puro e inalterado e o PB é emprestado e corrompido*”; em Portugal, traduz-se na atitude neocolonialista de que a miscigenação conduz à corrupção e ao empobrecimento de um hipotético português autêntico. Em contrapartida, a atitude *convergente racionalista* manifesta-se na ideia da “unidade na diversidade” e na proclamação, mais frequente em Portugal do que no Brasil, da unidade linguística como meio de afirmação política e económica no contexto transcontinental e global.

A atitude *divergente romântica* está presente nos que reclamam a existência de uma *língua brasileira*, (Bagno 2001). como manifestações sociais a legendagem de entrevistas a jovens portugueses no programa de televisão do canal MTV, e a legendagem e tradução de filmes portugueses para PB, como o filme “Capitães de abril” dirigido pela atriz portuguesa Maria de Medeiros. A atitude *divergente racionalista*, mais explícita no Brasil do que em Portugal, é assumida por todos os que veem na valorização, estudo e ensino da(s) norma(s) urbana(s) brasileira(s) um instrumento de participação política e educativa e a base de um nacionalismo cívico e de uma democracia liberal.

Até que ponto estas quatro atitudes poderão influenciar a divergência entre PE e PB? O processo gradual de divergência receberá um forte ímpeto, no Brasil, tanto da atitude romântica de afirmação da língua brasileira como da atitude racionalista de nacionalismo cívico e de democratização do ensino. Esta atitude racionalista poderá ter efeitos convergentes: a diminuição da acentuada diglossia brasileira e o reforço e aceleração do processo de padronização podem incluir mudanças “de baixo para cima”

que eliminam características do PB popular. Um terceiro cenário é ainda possível: interrupções ou reversões no esperado desenvolvimento divergente poderão ser desencadeadas pela atitude racionalista de preservação da unidade da língua como forma de afirmação política e económica.

4. Notas finais

No presente artigo, procurámos dar conta dos estudos desenvolvidos ao longo dos últimos anos sobre o português como língua pluricêntrica, na perspectiva de Linguística Cognitiva.

Assim, na secção 1, dedicada a uma reflexão sobre a produtividade lexical no PE e PB, e elaborada com base nos estudos da análise dos itens lexicais ‘*botar*’ e ‘*bota(-)fogo*’, demonstrou-se que (i) o significado estudado com base nos exemplos de uso linguístico pode apontar para empregos sincrónicos muito diferenciados e claramente divergentes no PE e PB a nível das respetivas normas *standard*, não reflectindo aparentemente as redes conceptuais construídas, histórica e socialmente determinadas, entretanto caídas parcialmente em desuso. Observámos, assim, que os empregos da linguagem-em-uso contemporâneas privilegiam certas aceções ou padrões de uso como mais frequentes em detrimento de outros, delimitados diatópica e/ou diastraticamente. Demonstrou-se, igualmente, que (ii) o significado de um item lexical não é objeto mental estável, mas resulta da arquitetura de rede(s) de conceptualizações e interpretações projetada em aberto e construída de acordo com critérios precisos e linguisticamente bem definidos no âmbito da Linguística Cognitiva. Esta arquitetura é, porém, diferenciada, em função dos factores específicos de cada uma das culturas nacionais em que se desenvolve, originando, assim, as especificidades a nível do PE e do PB. Demonstrou-se, também, que (iii) este sistema de estruturação conceptual envolve quer as capacidades cognitivas gerais quer os mecanismos do imaginário como metáfora e metonímia conceptuais, quer ainda conhecimentos que temos sobre o mundo, integrados numa semântica enciclopédica, que abrange, por exemplo, criação tanto de nomes comuns como de novos nomes próprios, tais como antropónimos ou topónimos.

Na secção 2, pudemos verificar, muito sumariamente, que a hipótese de divergência entre PE e PB é socioletometricamente confirmada por indicadores tanto

lexicais (termos de vestuário e de futebol) como construcionais (construções preposicionais, completivas e adjetivas), extraídos de um vasto *corpus* de textos dos últimos 60 anos.

Finalmente, na secção 3, apresentámos alguns elementos que mostram que os indicadores atitudinais replicam os indicadores lexicais e construcionais do *corpus* relativamente à tendência divergente entre as duas variedades nacionais e deixámos pistas para a interpretação dos modelos cognitivo-culturais românticos e racionalistas de convergência e divergência entre PE e PB.

REFERÊNCIAS

- Almeida, M.L.L.; R.G. Ferreira; D. Pinheiro; J. de L. Souza e C.A. Gonçalves (org.). 2009. *Linguística Cognitiva em Foco: Morfologia e Semântica*, Rio de Janeiro: Edições Publit.
- Bagno, Marcos. 2001. *Português ou Brasileiro? Um convite à pesquisa*. São Paulo: Parábola Editorial.
- Batoréo, Hanna Jakubowicz. 2004. *Linguística Portuguesa: Abordagem Cognitiva*. Lisboa: Universidade Aberta. CD-ROM.
- Batoréo, Hanna Jakubowicz. 2011a. As raízes de ‘*botafogo*’: construção de sentidos múltiplos em Português Europeu, *Veredas*, Juiz de Fora: Universidade Federal de Juiz de Fora, Brasil, – v. 15, nº 2/ 2011.
- Batoréo, Hanna Jakubowicz . 2011b. What do we know of Portuguese as a pluricentric language? Portuguese language usage in East Timor: The case of the polysemous marker ‘*já*’ and evidence of language contact in East Timorese written L2 texts. In: Silva *et al.* (org.) (2011) pp. 127-136.
- Batoréo, Hanna Jakubowicz e Margarida Casadinho. 2009a. O Português – uma língua pluricêntrica: O Caso de Timor-Leste. *Revista Portuguesa de Humanidades, Estudos Linguísticos*, 13 – 1, Braga: Universidade Católica Portuguesa, 63-79.
- Batoréo, Hanna Jakubowicz e Margarida Casadinho. 2009b. “«Botar ou não botar ... eis a questão» Estudo contrastivo da produtividade lexical do verbo ‘botar’ no Português Europeu e no Português do Brasil na óptica da Linguística Cognitiva”. In: M.L.L. Almeida; R.G. Ferreira; D. Pinheiro; J. de L. Souza e C.A. Gonçalves (org.) (2009). *Linguística Cognitiva em Foco: Morfologia e Semântica*, Rio de Janeiro: Edições Publit, capítulo 2, 53-66.
- Batoréo, Hanna Jakubowicz e Margarida Casadinho. 2010. “«Botar as mãos na massa?» Estudo Cognitivo da produtividade lexical do verbo ‘botar’ no PE e PB”. In: Marçalo, Maria João *et alii* (coord.) *Língua Portuguesa: ultrapassar fronteiras, juntar culturas*, Évora: Universidade de Évora, 37-55.
- Brandão, Sílvia Figueiredo & Mota, Maria Antónia (orgs.). 2003. *Análise Contrastiva de Variedades do Português. Primeiros Estudos*. Rio de Janeiro: In-Fólio.
- Clyne, Michael (ed.). 1992. *Pluricentric Languages. Differing Norms in Different Nations*. Berlin/New York: Mouton de Gruyter.
- Coelho, Eduardo Prado (org.). 2005. *A Língua Portuguesa: Presente e Futuro*. Lisboa: Fundação Calouste Gulbenkian.

Faraco, Carlos Alberto (org.). 2001. *Estrangeirismos. Guerras em torno da língua*. São Paulo: Parábola Editorial.

Gama, José (org.). 2007. *A Política da Língua Portuguesa*. Braga: Universidade do Minho.

Geeraerts, Dirk. 1997. *Diachronic Prototype Semantics. A Contribution to Historical Lexicology*, Oxford, Oxford University Press.

Geeraerts, Dirk. 2003. Cultural models of linguistic standardization. In R. Dirven, R. Frank & M. Pütz (eds.), *Cognitive Models in Language and Thought*. Berlin/New York: Mouton de Gruyter, 25-68.

Geeraerts, Dirk, Grondelaers, Stefan & Speelman, Dirk. 1999. *Convergentie en divergentie in de Nederlandse woordenschat*. Amsterdam: Meertens Instituut.

Geeraerts, Dirk & Cuyckens, Hubert (eds.). 2007. *The Oxford Handbook of Cognitive Linguistics*. Oxford/New York: Oxford University Press.

Geeraerts, Dirk, Kristiansen, Gitte & Peirsman, Yves (eds.). 2010. *Advances in Cognitive Sociolinguistics*. Berlin/New York: Mouton de Gruyter.

Kristiansen, Gitte & Dirven, René (eds.). 2008. *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social Systems*. Berlin/New York: Mouton de Gruyter.

Linguatca [em linha] disponível a partir de → <http://www.linguatca.pt/ACDC/>

Peres, João Andrade & Kato, Mary Aizawa (orgs.). 2004. Studies in the Comparative Semantics of European and Brazilian Portuguese. Issue of *Journal of Portuguese Linguistics* 3.

Silva, Augusto Soares da. 2005. Para o estudo das relações lexicais entre o Português Europeu e o Português do Brasil: Elementos de sociolexicologia cognitiva e quantitativa do Português. In: Inês Duarte & Isabel Leiria (eds.), *Actas do XX Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: Associação Portuguesa de Linguística, 211-226.

Silva, Augusto Soares da. 2006. Convergência e divergência no léxico do Português Europeu e do Português Brasileiro: resultados do estudo sobre termos de futebol e de moda. In Fátima Oliveira e Joaquim Barbosa (orgs.), *Textos Selecionados do XXI Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: APL, 633-646.

Silva, Augusto Soares da. 2008a. Integrando a variação social e métodos quantitativos na investigação sobre linguagem e cognição: para uma sociolinguística cognitiva do português europeu e brasileiro. *Revista de Estudos da Linguagem* 16 (1): 49-81.

Silva, Augusto Soares da. 2008b. O corpus CONDIV e o estudo da convergência e divergência entre variedades do português. In: Luís Costa, Diana Santos & Nuno Cardoso (orgs.), *Perspectivas sobre a Linguatca / Actas do Encontro Linguatca: 10 anos*. <http://www.linguatca.pt/LivroL10/>

Silva, Augusto Soares da. 2009. A Sociolinguística Cognitiva: razões e objecto de uma nova área de investigação linguística. *Revista Portuguesa de Humanidades – Estudos Linguísticos* 13 (1): 191-212.

Silva, Augusto Soares da. 2010. Measuring and parameterizing lexical convergence and divergence between European and Brazilian Portuguese. In: Dirk Geeraerts, Gitte Kristiansen & Yves Peirsman (eds.), *Advances in Cognitive Sociolinguistics*. Berlin/New York: Mouton de Gruyter, 41-83.

Silva, Augusto Soares da. 2011a. Introdução. Línguas pluricêntricas: variação linguística e dimensões sociocognitivas. In: Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves (orgs.), *Línguas Pluricêntricas: Variação Linguística e Dimensões Sociocognitivas. Pluricentric Languages: Linguistic Variation and Sociocognitive Dimensions*. Braga: Aletheia, Publicações da Faculdade de Filosofia da Universidade Católica Portuguesa, 13-25.

Silva, Augusto Soares da. 2011b. Para a abordagem socioletométrica do pluricentrismo do português europeu e brasileiro: dos indicadores lexicais aos construcionais e atitudinais. In Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves (orgs.), *Línguas Pluricêntricas: Variação Linguística e Dimensões Sociocognitivas. Pluricentric Languages: Linguistic Variation and Sociocognitive Dimensions*. Braga: Aletheia, Publicações da Faculdade de Filosofia da Universidade Católica Portuguesa.

Silva, Augusto Soares da (no prelo). Comparing subjective and objective linguistic distances between European and Brazilian Portuguese. In: Martin Pütz, Monika Reif & Justyna Robinson (eds.), *LAUD 2010 Proceedings – Cognitive Sociolinguistics*. Frankfurt: Peter Lang.

Silva, Augusto Soares da e Hanna Jakubowicz Batoréo. 2010. Gramática Cognitiva: estruturação conceptual, arquitectura e aplicações. In: Ana Maria Brito (org.) *Gramática: História, Teorias, Aplicações*, Porto: Faculdade de Letras da Universidade do Porto, 229-251.