

# Para o estudo das relações lexicais entre o Português Europeu e o Português do Brasil: Elementos de sociolexicologia cognitiva e quantitativa do Português

Augusto Soares da Silva  
Universidade Católica Portuguesa, Braga

## 1. Introdução

Na base deste estudo está um projecto de investigação em curso sobre as relações lexicais entre o Português Europeu e o Português do Brasil ao longo dos últimos 50 anos, intitulado “Convergência e divergência no léxico do Português”.<sup>1</sup> Esta investigação apoia-se na concepção geral e nos métodos quantitativos da investigação sociolexicológica cognitiva desenvolvida por Dirk Geeraerts e sua equipa para o Neerlandês (Geeraerts, Grondelaers & Speelman 1999).

Pretendemos explorar o domínio de uma sociolexicologia *cognitiva e quantitativa* da língua portuguesa e apresentar os primeiros resultados do referido projecto de investigação. Num primeiro momento, situaremos a sociolexicologia no contexto de um modelo explicativo das diferentes formas de variação lexical. Procuraremos mostrar as vantagens da onomasiologia *pragmática*, de métodos *quantitativos* e da perspectiva da Linguística Cognitiva para o estudo das relações lexicológicas entre diferentes variedades de uma língua. A seguir, indicaremos os métodos para calcular a *uniformidade* lexical entre variedades de uma língua. Finalmente, apresentaremos os elementos e resultados da investigação em curso.

## 2. Sociolexicologia cognitiva

### 2.1. Mapa conceptual da variação lexical

No estudo das palavras e seus significados, em geral, bem como no estudo da variação lexical, em particular, podemos partir, ora da palavra para os seus significados e referentes (ou campo de aplicação semântica e referencial dessa palavra ou item lexical), ora de um significado ou conceito (ou ainda uma entidade referencial) para as diferentes palavras ou itens lexicais que o designam. Esta distinção foi bem estabelecida

---

<sup>1</sup> O Projecto é financiado pela Fundação para a Ciência e a Tecnologia (Ref<sup>o</sup> POCTI/ LIN/48575/2002) e tem a duração de dois anos (Janeiro 2004 – Janeiro 2006). Integram a equipa de investigação, para além do autor deste texto, José João Dias Almeida, Alberto Manuel Simões, Ana Margarida Abrantes, Ana Margarida Nunes e José Luiz de Lucca.

na tradição continental da semântica estruturalista (mas quase desconhecida na tradição anglo-saxónica), sob as designações de *semasiologia* e *onomasiologia*, respectivamente (cf. Baldinger, 1964). Fundamentalmente, baseia-se esta distinção tradicional na diferença entre dois importantes fenómenos semântico-lexicais: *significação* e *nomeação*. Mais concretamente, enquanto a semasiologia faz a descrição dos vários sentidos de uma palavra ou outra expressão, a onomasiologia analisa as expressões alternativas pelas quais determinado conceito é nomeado.

Tanto para a lexicologia, em geral, como para o estudo da variação lexical, em particular, importa estabelecer uma distinção entre duas modalidades de onomasiologia, tradicionalmente não diferenciadas. Por um lado, o estudo da estrutura semântica do léxico, organizado em campos lexicais, taxionomias hierárquicas, “frames” e por relações como sinonímia, antonímia, hiponímia, etc. Por outro lado, o estudo da escolha real que os falantes têm que fazer de uma expressão particular para designar ou nomear determinado conceito ou referente. Por outras palavras, esta é a distinção, bem estabelecida por Geeraerts (1998; Grondelaers & Geeraerts, 2003), entre uma onomasiologia *estrutural*, que relewa do plano da *langue* ou estrutura – e constitui a modalidade tradicional da onomasiologia –, e uma onomasiologia *pragmática*, que relewa do plano da *parole* ou uso real das expressões linguísticas – e se apresenta como um novo desafio ao lexicólogo. A primeira ocupa-se de conjuntos de expressões inter-relacionadas e procura responder à questão de saber quais são as relações existentes entre as expressões alternativas, ao passo que a segunda ocupa-se das escolhas actuais de entre as alternativas disponíveis e procura responder à questão de saber quais são os factores que determinam a escolha de uma ou outra alternativa.

Uma outra distinção básica é a que deve estabelecer-se entre os aspectos *qualitativos* e os aspectos *quantitativos* das estruturas semântico-lexicais, tanto semasiológicas como onomasiológicas. Semasiologicamente, a questão qualitativa tem a ver com a identificação dos sentidos (e tipos de referentes) de uma palavra e das relações entre esses sentidos/referentes, isto é, envolve o estudo da polissemia (e outros fenómenos afins, como a vaguidade) e, nesta implicado, o estudo de relações semasiológicas como a metáfora, a metonímia, a generalização e a especialização de sentido. Onomasiologicamente, a mesma questão conduz ao estudo das relações entre diferentes itens lexicais e, desta forma, à investigação dos diferentes tipos de estruturação do léxico, bem populares na tradição da semântica estrutural, designadamente campos lexicais, taxionomias hierárquicas, mas também “frames” (no sentido de Fillmore, 1977), e relações como a sinonímia, a antonímia, a hiponímia, a meronímia (sobre estas e outras relações, ver Cruse, 1986), mas também as chamadas metáforas e metonímias *conceptuais* (Lakoff & Johnson, 1980, 1999; Silva, 2003b), na medida em que constituem verdadeiras estruturas onomasiológicas (metafóricas e metonímicas). Pelo contrário, a dimensão quantitativa, introduzida em semântica lexical pela Semântica Cognitiva, sob a forma da teoria do protótipo (Taylor, 1995; Geeraerts, 1997) e da teoria do nível básico, envolve, semasiologicamente, as diferenças de saliência entre os vários sentidos/referentes de uma palavra, isto é, o fenómeno da *prototipicidade* e seus efeitos (estrutura de centro vs. periferia, redes radiais, etc.);

e, onomasiologicamente, as diferenças de saliência ou *ancoragem* conceptual (“entrenchment”; Langacker, 1987) entre diferentes categorias conceptuais e a hipótese do *nível básico* das taxionomias lexicais.

Esta oposição ‘qualitativo/quantitativo’ pode aplicar-se também à onomasiologia pragmática. Mas agora já não se trata tanto da separação entre elementos e relações, de um lado, e diferenças de saliência, do outro, mas sobretudo de diferenças conceptuais, de um lado, e diferenças não-conceptuais entre categorias qualitativamente idênticas, do outro. Assim, a onomasiologia pragmática, em termos qualitativos, ocupa-se da selecção entre categorias conceptualmente diferentes, ora de diferentes níveis taxionómicos ora do mesmo nível taxionómico, ao passo que, quantitativamente, tem a ver com a selecção entre categorias que diferem em termos não-conceptuais, seja a nível emotivo, estilístico, sociolinguístico ou discursivo, isto é, ocupa-se da selecção entre sinónimos conceptuais (denotacionais). Naturalmente que no plano quantitativo há lugar também para diferenças de saliência, na forma de *prevalência sociolinguística* entre esses sinónimos.

Ainda outras distinções relevantes são as que se devem estabelecer entre significado *conceptual* (denotacional ou referencial) e significado *não-conceptual* (emotivo, estilístico, sociolinguístico e discursivo), como acabámos de verificar, e entre sincronia e diacronia.

Estabelecidas estas distinções básicas, chegamos ao próprio campo ou mapa conceptual da semântica lexical, sintetizado na Figura 1.

	QUALIDADE: entidades e relações	QUANTIDADE: diferenças de saliência
SEMASIOLOGIA	sentidos (polissemia) e relações entre sentidos (metáfora, metonímia, etc.)	prototipicidade
ONOMASIOLOGIA ESTRUTURAL	ítems lexicais e suas relações (campos lexicais, taxionomias, <i>frames</i> , sinonímia, antonímia, etc)	“entrenchment” e nível básico
ONOMASIOLOGIA PRAGMÁTICA	selecção entre categorias conceptualmente diferentes	prevalência sociolinguística selecção entre sinónimos conceptuais

Figura 1. Mapa conceptual da semântica lexical

Torna-se agora mais fácil identificar as diferentes formas de variação lexical e, simultaneamente, compreender algumas das suas interacções e condicionamentos recíprocos (ver Gearerets, Grondelaers & Bakema, 1994; para uma síntese, Silva, 1996). É o que se representa na Figura 2. Temos assim as seguintes formas de variação lexical:

- variação *semasiológica*: diferentes sentidos ou (tipos de) referentes de um item lexical e efeitos de prototipicidade; por exemplo, o verbo *deixar*, estudado por Silva (1999, 2003a);

- variação *onomasiológica categorial* (ou *conceptual*): diferentes categorias conceptuais para nomear determinado conceito ou tipo de referentes e diferenças de ancoragem; por exemplo, *jogador* e *atacante*;
- variação *onomasiológica formal* (ou, se quisermos, variação onomasiológica pragmática “não-qualitativa”): diferentes nomes para uma mesma categoria conceptual e diferenças de prevalência sociolinguística; por exemplo, *atacante* e *avançado* ou *guarda-redes* e *goleiro*;
- variação *externa*: variação sociolinguística e estilística, contextual (pragmático-discursiva) e diacrónica.

Os diferentes tipos de variação *externa* podem influenciar qualquer outra variação lexical. Mas há aqui uma diferença a assinalar: a variação externa constitui uma dimensão adicional para as variações semasiológica e onomasiológica categorial, ao passo que é justamente essa variação externa a que define a variação onomasiológica formal, isto é, as diferenças entre os sinónimos denotacionais, de casos como *atacante* vs. *avançado*, *guarda-redes* vs. *goleiro*.

	QUALIDADE: entidades e relações	QUANTIDADE: diferenças de saliência
SEMASIOLOGIA	variação semasiológica	variação semasiológica
ONOMASIOLOGIA ESTRUTURAL	variação onomasiológica categorial	variação onomasiológica categorial
ONOMASIOLOGIA PRAGMÁTICA	variação onomasiológica categorial	variação onomasiológica formal
	variação externa (sociolinguística, contextual, diacrónica)	

Figura 2. Mapa conceptual da variação lexical

## 2.2. Onomasiologia pragmática e sociolexicologia

É precisamente a variação *onomasiológica pragmática formal*<sup>2</sup> ou “não-qualitativa” e, dela fazendo parte integrante, a variação *externa* o objecto mais específico do que se pode designar como *sociolexicologia*; mas também é verdade que este sub-domínio da sociolinguística compreende todas as demais variações lexicais enquanto condicionadas ou influenciadas pelos diferentes tipos de variação *externa*. E é igualmente esta variação onomasiológica pragmática formal, e portanto a questão das escolhas que os falantes têm que fazer entre palavras que diferem apenas no significado não-conceptual, o objecto do presente projecto de investigação.

Importa ainda esclarecer a questão essencial da onomasiologia pragmática: quais os factores que determinam ou, pelo menos, influenciam a escolha de um nome para determinado conceito ou referente; por outras palavras, quais os factores que

<sup>2</sup> Para uma exploração do domínio da onomasiologia, incluindo a onomasiologia diacrónica (campo em foco da presente investigação), ver Silva (no prelo).

condicionam a *escolha onomasiológica* ou o acto efectivo de *nomeação*? A resposta, dada por Geeraerts, Grondelaers & Bakema (1994) com base numa investigação sobre termos de vestuário do neerlandês (variedades holandesa e belga), inclui os seguintes factores: (i) a saliência semasiológica ou grau de prototipicidade do conceito/referente relativamente à estrutura semasiológica da categoria; (ii) a saliência onomasiológica ou ancoragem ("entrenchement") da categoria nomeada pela expressão; e (iii) valores contextuais de natureza emotiva, pragmático-discursiva e, particularmente, sociolinguística (geográfica, social, estilística, histórica), envolvendo a competição de diferentes variedades e/ou variantes de uma língua. Dito de modo mais resumido, os factores determinantes da escolha onomasiológica são (i) a distância semântica entre o alvo e as designações alternativas, (ii) o peso onomasiológico das designações alternativas e (iii) traços sociolinguísticos das expressões alternativas. É esta terceira categoria de factores o objecto específico da investigação sociolinguística, e do presente projecto de investigação.

É momento agora de responder às questões do *porquê* e *como* do nosso objecto de investigação. Porquê a onomasiologia pragmática formal ou não-qualitativa como perspectiva preferencial para o estudo das relações lexicológicas entre duas variedades de uma língua, neste caso, o Português Europeu e o Português do Brasil? Em primeiro lugar, a variação implica escolha, o que significa que é necessário estudar as escolhas efectivas que os falantes têm que fazer entre vários itens lexicais disponíveis para nomear determinado conceito ou referente particular. Daqui a relevância da onomasiologia pragmática. Em segundo lugar, os sinónimos denotacionais (referenciais) são os que melhor revelam diferenças regionais, sociais, estilísticas e históricas, na base das quais se definem a própria existência e a competição de diferentes variedades e variantes de uma língua. Daqui a importância da onomasiologia pragmática formal ou não-qualitativa. Finalmente, o estudo da variação onomasiológica pragmática formal e da consequente variação externa permite compreender os aspectos sociais da variação lexical, incluindo as relações entre diferentes variedades de uma língua.

### 2.3. Observação do uso e perspectiva cognitivista

Uma investigação desta natureza tem que necessariamente optar por uma metodologia *baseada-no-uso*: observação do uso real das expressões lexicais, observação das escolhas onomasiológicas efectivamente realizadas pelos falantes entre alternativas disponíveis. Quer isto dizer que a investigação tem que ter por base empírica um *corpus* representativo. Igualmente necessários serão métodos que permitam *medir* a distância lexical entre as diferentes variedades do Português, a que faremos referência na próxima secção.

A orientação teórica é, como já foi referido, a da Linguística Cognitiva. Porquê ou quais as vantagens da perspectiva cognitiva nesta investigação (sócio)lexicológica, ou numa outra qualquer? Poderemos apontar três ordens de razões. Primeiro, de entre as principais teorias de semântica lexical, é indubitavelmente a semântica cognitiva a que se tem orientado e mais tem contribuído para o estudo dos aspectos *quantitativos* das

estruturas lexicais – diferentes manifestações de efeitos de prototipicidade, do lado semasiológico; e ancoragem conceptual de categorias e níveis básicos, do lado onomasiológico. Ao mesmo tempo, é ela a que tem desenvolvido a onomasiologia pragmática, quase inexistente nas outras tradições teóricas. Pelo contrário, o foco da semântica estrutural foi e é o dos aspectos qualitativos das estruturas onomasiológicas; a tradição pré-estruturalista esteve centrada na dimensão qualitativa das estruturas semasiológicas; e tem sido também para estas últimas, mais concretamente para a polissemia regular, que se tem dirigido a semântica neo-generativista (Pustejovsky, 1995).

Ainda nesta primeira ordem de razões, estando a Linguística Cognitiva, desde o início, empenhada no estudo da categorização, entendida como sendo a função básica da linguagem (Silva, 2004b), é natural que a dimensão onomasiológica esteja no centro das atenções: do ponto de vista do falante, o acto básico de categorização é precisamente a escolha onomasiológica de uma categoria para exprimir determinada ideia.

Em segundo lugar, a Linguística Cognitiva caracteriza-se especificamente por uma perspectiva *recontextualizante* (reintegrando no estudo da linguagem as diferentes formas de *contexto*, excluídas pelos modelos gramaticais autonomistas), *experientialista* (Lakoff & Johnson, 1999; Silva, 2004a), orientada para o significado (nos seus diferentes tipos, incluindo o não-referencial) e baseada no uso. Resulta daí a inevitabilidade de uma sociolexicologia cognitiva, em particular, e de uma sociolinguística cognitiva, em geral, justificada por Geeraerts (2003).

Finalmente, a Linguística Cognitiva, provavelmente mais do que qualquer outra abordagem contemporânea da linguagem, reconhece explicitamente, não só que a capacidade para a linguagem se fundamenta em capacidades cognitivas gerais, como também que todas estas capacidades são cultural e socialmente situadas e definidas (Silva 2004b). Mais concretamente, o estudo do papel da conceptualização da realidade social na constituição dos factos sociolinguísticos é uma das principais tarefas de uma linha de investigação em linguística cognitiva, conhecida como teoria dos *modelos culturais* (Palmer, 1996; e, para o domínio de investigação em causa, Geeraerts, 2004). Ora, a sociologia da linguagem envolve justamente modelos culturais sobre variedades e variantes de uma língua, comportamentos e atitudes para com essa variação.

### 3. Cálculos de uniformidade e (socio)lexicologia quantitativa

Como medir o grau de conformidade onomasiológica – grau de conformidade da escolha lexical – entre dois sub-conjuntos de uma base de dados de uma língua (por exemplo, entre duas variantes regionais)? Como medir a distância lexical entre duas variedades de uma língua (por exemplo, o Português Europeu e o Português do Brasil)? Como medir a convergência ou divergência lexical entre duas variantes ou variedades? A unidade de investigação em lexicologia quantitativa liderada por D. Geeraerts tem desenvolvido métodos quantitativos, fundamentados nas noções de *perfil onomasiológico* e *uniformidade* (ver exposição sintetizada em Geeraerts, 2001).

O *perfil onomasiológico* de um conceito/referente numa determinada variedade linguística é o conjunto de sinónimos usados para designar esse conceito/referente nessa variedade linguística, diferenciados pela sua frequência relativa. Relativamente ao método clássico das *palavras-chave* (em que a frequência de um termo num *corpus* é comparada à frequência desse termo noutra *corpus*), o método dos *perfis* tem, para além do mais, a vantagem de neutralizar efeitos temáticos: por exemplo, a frequência absoluta elevada do termo *goleiro* pode ter a ver com a especificidade temática do contexto em que ocorre (texto sobre futebol e texto da variedade brasileira).

A *uniformidade* é a medida da correspondência entre dois conjuntos de dados, definidos em termos de perfis onomasiológicos. Por exemplo, a uniformidade de um determinado conceito entre duas amostras, em que uma contém 6 ocorrências do termo A e 4 do termo B e a outra 3 ocorrências do termo A e 7 do termo B, resulta do número de pares comuns de nomeação desse conceito (7 pares), sendo portanto de 70%. Este mesmo resultado obtém-se somando as frequências relativas mais pequenas de cada termo alternativo: 30% do termo A (3 usos, provenientes da segunda amostra) e 40% do termo B (4 usos, provenientes da primeira amostra).

Em termos técnicos, o cálculo de uniformidade de um conceito pode ser formulado da seguinte forma:

$$U_z(Y_1, Y_2) = \sum_{i=1}^n \min(F_{z,y_1}(x_i), F_{z,y_2}(x_i))$$

Isto é, a uniformidade U para um conceito Z entre duas amostras Y<sub>1</sub> e Y<sub>2</sub> equivale à soma  $\sum$  dos mínimos das frequências relativas F do termo x nos perfis onomasiológicos de Z em Y<sub>1</sub> e Y<sub>2</sub>. O símbolo x<sub>i</sub> representa os diferentes termos x<sub>1</sub> a x<sub>n</sub> usados nas amostras Y para designar o conceito Z.

Quando estão em causa vários conceitos, a uniformidade média é calculada em termos de *média ponderada*, através da seguinte fórmula:

$$U'(Y_1, Y_2) = \sum_{i=1}^n U_z(Y_1, Y_2) \cdot G_z$$

A uniformidade U' para um conjunto de conceitos Z entre duas amostras Y<sub>1</sub> e Y<sub>2</sub> equivale à soma dos valores-U dos Zs ponderados pela frequência relativa G de Z dentro do conjunto total de Zs. Obtém-se o coeficiente de ponderação G<sub>z</sub> ou frequência relativa de um conceito Z, dividindo a soma das ocorrências de Z em Y<sub>1</sub> e Y<sub>2</sub> pelo número total de ocorrências para os conceitos Z<sub>1</sub> a Z<sub>n</sub>.

Resumindo, o cálculo da conformidade onomasiológica entre duas variantes de uma língua faz-se medindo o grau de uniformidade entre perfis onomasiológicos dessas variantes. Diacronicamente, podemos dizer que essas variantes se encontram num processo de convergência quando a medida de uniformidade U aumenta, e num processo de divergência quando U diminui.

#### 4. Convergência/divergência lexical entre o PE e o PB

Vamos agora apresentar os elementos e os primeiros resultados do projecto de investigação em curso sobre a relação lexicológica entre as variedades europeia e brasileira. O objectivo principal do projecto é saber se as duas variedades nacionais da língua portuguesa estão envolvidas num processo de convergência ou divergência lexical. Tal objectivo implica obviamente uma análise diacrónica. O período em análise abrange as últimas cinco décadas e o material coligido pertence aos anos 50, 70 e 90-00. Secundariamente, pretendemos também investigar, em cada uma das variedades nacionais, a distância sincrónica entre padrão e registos intermédios, e, desse modo, verificar se a estratificação lexical das duas variedades é idêntica ou diferente.

A base empírica da investigação consiste em largos milhares de observações do uso de termos alternativos, concretamente sinónimos referenciais, para nomear 40 conceitos de dois domínios ou campos lexicais – 20 conceitos do campo do futebol e 20 conceitos do campo da moda/vestuário – e ainda 10 conceitos isolados de alta e baixa frequência. A base de dados resultante compreende pois as frequências de cada um dos termos dos 50 *perfis onomasiológicos* em textos portugueses e brasileiros. Poderemos assim verificar, por exemplo, a repartição quantitativa dos termos *atacante*, *avançado*, *avante*, *dianteiro*, *forward* e *ponta-de-lança* para designar o conceito de ‘atacante’, ou para o conceito de ‘fora-de-jogo’ saber se há preferência pela expressão *deslocação*, *fora-de-jogo*, *impedimento* ou então *offside*.

Os dados são extraídos de três fontes: (i) jornais e revistas de desporto e moda dos primeiros anos das décadas de 50, 70 e 90-00 – designadamente, e em relação ao futebol, os jornais portugueses *A Bola*, *Record*, *Mundo Desportivo* e *O Jogo* e os jornais brasileiros *Jornal dos Sports* (Rio de Janeiro), *Gazeta Esportiva* (São Paulo), *Estado de São Paulo* e *Lance* –; (ii) linguagem da Internet (tanto a conversação *off-line* de fóruns de discussão, como a conversação *on-line* de IRC ou *chats*); e (iii) etiquetas e catálogos de lojas de vestuário de diferentes cidades tanto de Portugal como do Brasil. Adicionaremos os resultados de um inquérito dirigido a estudantes dos dois países.

O *corpus* CONDIVport é, assim, estruturado na base de três variáveis:

- geográfica: Portugal vs. Brasil
- diacrónica: 1950, 1970, 1990-2000
- estilística: jornais e revistas de qualidade > jornais e revistas populares > Net-*off-line* > etiquetas e catálogos + Net-*on-line* (*chats*).

Actualmente com 2,7 milhões de palavras, provenientes dos jornais de desporto referidos, espera-se que chegue aos 5 milhões. Será disponibilizado no sítio da Linguateca, em [www.linguateca.pt](http://www.linguateca.pt).

Apresentamos a seguir a lista dos 21 conceitos nominais (substantivos) de futebol e respectivos sinónimos referenciais,<sup>3</sup> em estudo na fase actual da investigação:

<sup>3</sup> Esta lista foi elaborada a partir dos dicionários de referência da língua portuguesa (Academia das Ciências de Lisboa, Houaiss, Aurélio, Porto Editora) e de alguns dicionários especializados de futebol, tanto brasileiros como portugueses, e ainda de estudos linguísticos sobre o futebol, como o de Feijó (1998). Os estrangeirismos que conservam a sua forma original estão indicados entre aspas.



- ÁRBITRO: *apitador, árbitro, director da partida, juiz, juiz de campo, "ref(eree)", referi, refre.*
- ÁRBITRO AUXILIAR: *árbitro auxiliar, árbitro assistente, auxiliar, 2º/3º/4º árbitro, bandeirinha, fiscal de linha, juiz de linha, "liner".*
- ATACANTE: *atacante, avançado, avante, dianteiro, "forward", ponta-de-lança.*
- BALIZA: *arco, baliza, cidadela, "goal", gol(o), malhas, marco, meta, rede, redes, vala.*
- BOLA: *balão, bola, couro(inho), esfera, esférico, pelota, redondinha.*
- DEFESA: *"(full-)back", beque, bequeira, defensor, defesa, lateral, libero, zagueiro.*
- EQUIPA: *conjunto, formação, eleven, equipa/e, escrete, esquadra, esquadrão, grupo, "match", onze, onzena, plantel, quadro, "scratch(men)", "team", time, turma.*
- EXTREMO: *ala, extremo, ponta, ponteiro.*
- FALTA: *carga, falta, "foul", golpe (baixo, irregular), infra(c)ção, obstru(c)ção, transgressão, violação (das regras).*
- FINTA: *baile, corte, drible(ing), engano, "feint", finta, firula, ginga, lesa, manobra enganadora, simulação.*
- FORA-DE-JOGO: *adiantamento, banheira, deslocação, fora-de-jogo (fora de jogo), impedimento, "offside" ("off-side"), posição irregular.*
- GOLO: *bola, "goal", gol (gól), golo, ponto, tento.*
- GRANDE PENALIDADE: *castigo máximo, castigo-mor, falta máxima, grande penalidade, penalidade, penalidade máxima, penákti (pênalti, pénalti), "penalty".*
- GUARDA-REDES: *arquero, "goal-keeper", goleiro, golquiper, guarda-meta, guarda-rede, guarda-redes, guarda-vala, guarda-valas, guardião, "keeper", porteiro, quiper, vigia.*
- JOGADA: *combinação, jogada, lance.*
- JOGO: *batalha, choque, combate, competição, confronto, desafio, disputa, duelo, embate, encontro, jogo, justa, luta, "match", partida, peladinha, peleja, prélio, prova, pugna.*
- MÉDIO: *alfe, armador, cabeça de área, camisa (camisola) 10, central, centro-campista, centro-médio, distribuidor (de jogo), half, interior, médio, meia, meio-campista, meio-campo, "midfield", motor, número 10, patrão, trinco, volante.*
- PONTAPÉ: *balázio, biqueira, biqueirada, chute, chuto, "kick(-off)", panázio, pelotada, pontapé, quique, "shoot", tiro.*
- PONTAPÉ DE CANTO: *canto, chute de canto, córner (corner), escanteio, esquinado, pontapé de canto, tiro de canto.*
- PONTAPÉ LIVRE: *chute (in)directo, falta, "free(-kick)", livre (directo, indirecto), pontapé livre, tiro dire(c)to, tiro livre (direto, indireto)*
- TREINADOR: *mister, professor, técnico (orientador), treinador.*

Temos resultados dos 21 perfis onomasiológicos relativamente a um *sub-corpus* extraído dos jornais *A Bola* (Portugal) e *Jornal dos Sports* (Brasil), com uma extensão de cerca de 1,7 milhões de palavras. A respectiva base de dados é constituída por 58.313 observações do uso dos referidos termos designativos destes conceitos. A Figura 3 apresenta os valores de uniformidade de cada conceito (U) e de uniformidade média (U') dos 21 conceitos de futebol do português de Portugal (P) e do Brasil (B), das décadas de 50 (P50/B50), 70 (P70/B70) e 90-00 (P00/B00). Em anexo, apresentamos as frequências dos respectivos itens lexicais observados.

	U	U'	U	U'	U	U'	total n°
	P50/B50	P50/B50	P70/B70	P70/B70	P00/B00	P00/B00	
ARBITRO	46,3	1,86	15,5	0,59	84,0	2,80	2189
ARBITRO AUXILIAR	21,8	0,12	40,4	0,30	48,5	0,27	356
ATACANTE	20,7	0,74	31,1	0,92	10,4	0,36	1959
BALIZA	25,5	1,20	12,3	0,46	5,6	0,15	2235
BOLA	64,1	5,61	82,9	7,82	93,7	5,39	4722
DEFESA	16,7	0,43	15,1	0,54	32,1	1,27	1903
EQUIPA	29,7	5,55	24,8	5,01	38,3	6,58	10914
EXTREMO	7,1	0,16	2,7	0,04	45,5	0,15	815
FALTA	41,7	0,48	85,4	0,60	89,3	0,97	582
FINTA	62,3	0,21	57,4	0,23	64,3	0,15	192
FORA-DE-JOGO	7,7	0,04	0,0	0,00	0,0	0,00	286
GOLO	16,3	1,87	0,3	0,04	0,0	0,00	7559
GRANDE PENALIDADE	33,2	0,58	0,5	0,01	0,0	0,00	1042
GUARDA-REDES	6,9	0,19	1,3	0,03	0,0	0,00	1561
JOGADA	97,0	4,27	67,8	4,26	49,1	2,37	2971
JOGO	53,1	13,11	74,9	16,31	71,8	15,59	13381
MÉDIO	13,6	0,29	29,7	0,18	0,0	0,00	1024
PONTAPÉ	34,7	0,73	7,9	0,11	3,1	0,03	896
PONTAPÉ DE CANTO	0,4	0,01	11,5	0,08	0,0	0,00	564
PONTAPÉ LIVRE	0,0	0,00	0,0	0,00	5,6	0,06	469
TREINADOR	56,9	0,80	74,6	2,92	62,0	6,14	2693
total		38,24		40,42		42,30	58313

Figura 3. Valores de U e U' de 21 conceitos de futebol

Os totais de U' apontam para uma ligeira convergência:

$$\begin{array}{rcl}
 U' (P50,B50) & < & U' (P70,B70) & < & U' (P00,B00) \\
 38,24 & < & 40,42 & < & 42,30
 \end{array}$$

Na amostra da Figura 3, há perfis onomasiológicos de elevada frequência absoluta e com valores relativamente elevados de U'. São eles JOGO (13.381 ocorrências) EQUIPA (10.914) e BOLA (4.722). Apesar de o cálculo de U' ter já em conta estas assimetrias quantitativas, se eliminarmos estes três perfis, obtemos os seguintes resultados:

U' (P50,B50)	>	U' (P70,B70)	<	U' (P00,B00)
29,19	>	23,22	<	26,62

Teríamos então divergência da década de 50 para a de 70 e uma certa reconvergência de 70 até à actualidade. Este resultado confirmaria a hipótese sociolexicológica que parece colher mais consenso: divergência entre as duas variedades nacionais do Português, mais acentuada até aos anos 70, e menor (ou uma certa reconvergência), a partir de então, por razões sociais, políticas, culturais, económicas bem conhecidas.

## 5. Conclusão

No plano teórico-metodológico, esperamos ter mostrado (i) a relevância da onomasiologia *pragmática* (estudo dos factores que determinam as escolhas entre expressões alternativas) e *formal* (estudo das escolhas entre sinónimos referenciais) para a lexicologia variacional e a sociolexicologia; (ii) a necessidade de métodos *quantitativos*; e (iii) os contributos da Linguística Cognitiva e, assim, as vantagens de uma *sociolexicologia cognitiva*.

Com o presente projecto de sociolexicologia cognitiva do Português, esperamos poder contribuir para (i) o conhecimento das relações lexicológicas actuais entre o Português Europeu e o Português do Brasil, nomeadamente, saber se há convergência ou divergência entre as duas variedades nacionais da língua portuguesa; (ii) o desenvolvimento da sociolexicologia do Português (os estudos sociolinguísticos existentes privilegiam, geralmente, as diferenças fonéticas e sintácticas); (iii) a concepção e implementação de políticas da língua portuguesa, como língua transnacional e transcontinental que é; e (iv) o estudo da função cognitiva básica da linguagem – a categorização –, designadamente no que se refere ao acto de escolha onomasiológica de uma categoria para exprimir determinada ideia.

## Referências Bibliográficas

- BALDINGER, Kurt (1964) Sémasiologie et onomasiologie. *Revue de Linguistique Romane* 28, pp. 249-272.
- CRUSE, D. Alan (1986) *Lexical Semantics*. Cambridge: Cambridge University Press.
- FEIJÓ, Luiz Cesar Saraiva (1998) *Brasil x Portugal: Um derby linguístico*. Rio de Janeiro: Sociedade Brasileira de Língua e Literatura.
- FILLMORE, Charles (1977) Scenes-and-frames semantics. In A. Zampolli (ed.) *Linguistic Structures Processing*. Amsterdam: North Holland, pp. 55-81.

- GEERAERTS, Dirk (1997) *Diachronic Prototype Semantics. A Contribution to Historical Lexicology*. Oxford: Clarendon Press.
- \_\_\_\_\_ (1998) Sémantique cognitive et onomasiologie. *Revista Portuguesa de Filologia* 22, pp. 329-339.
- \_\_\_\_\_ (2001) On measuring lexical variation. In Augusto Soares da Silva (org.) *Linguagem e Cognição: A Perspectiva da Linguística Cognitiva*. Braga: APL e Universidade Católica Portuguesa, pp. 51-61.
- \_\_\_\_\_ (2003) 'Usage-based' implies 'variational'. On the inevitability of Cognitive Sociolinguistics. Conferência plenária apresentada no 8<sup>th</sup> *International Cognitive Linguistics Conference*, Logroño, Universidade de La Rioja, 20-25 Julho 2003.
- \_\_\_\_\_ (2004) Cultural models of linguistic standardization. In Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves (orgs.) *Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva*. Vol. I. Coimbra: Almedina, pp. 47-84.
- GEERAERTS, Dirk, Stefan Grondelaers & Peter Bakema (1994) *The Structure of Lexical Variation. Meaning, Naming, and Context*. Berlin: Mouton de Gruyter.
- GEERAERTS, Dirk, Stefan Grondelaers & Dirk Speelman (1999) *Convergentie en divergentie in de Nederlandse woordenschat*. Amsterdam: Meertens Instituut.
- GRONDELAERS, Stefan & Dirk Geeraerts (2003) Towards a pragmatic model of cognitive onomasiology. In Hubert Cuyckens, René Dirven & John Taylor (eds.) *Cognitive Approaches to Lexical Semantics*. Berlin: Mouton de Gruyter, pp. 67-92.
- LAKOFF, George & Mark Johnson (1980) *Metaphors We Live By*. Chicago: The University of Chicago Press.
- \_\_\_\_\_ (1999) *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.
- LANGACKER, Ronald W. (1987) *Foundations of Cognitive Grammar*. Vol. I. *Theoretical Prerequisites*. Stanford: Stanford University Press.
- PALMER, Gary B. (1996) *Toward a Theory of Cultural Linguistics*. Texas: University of Texas Press.
- PUSTEJOVSKY, James (1995) *The Generative Lexicon: A Theory of Computational Lexical Semantics*. Cambridge, MA: MIT Press.
- SILVA, Augusto Soares da (1996) Sobre a estrutura da variação lexical. Elementos de lexicologia cognitiva. *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística*. Vol. III. Lisboa: Colibri, pp. 413-423.
- \_\_\_\_\_ (1999) *A Semântica de DEIXAR: Uma Contribuição para a Abordagem Cognitiva em Semântica Lexical*. Lisboa: Fundação Calouste Gulbenkian e Fundação para a Ciência e a Tecnologia.
- \_\_\_\_\_ (2003a) Image schemas and category coherence: The case of the Portuguese verb *deixar*. In Hubert Cuyckens, René Dirven & John Taylor (eds.) *Cognitive Approaches to Lexical Semantics*. Berlin: Mouton de Gruyter, pp. 281-322.
- \_\_\_\_\_ (2003b) O poder cognitivo da metáfora e da metonímia. *Revista Portuguesa de Humanidades* 7, pp. 13-75.
- \_\_\_\_\_ (2004a) Protótipos, imagens e metáforas, ou o experiencialismo da linguagem e do pensamento. In Alfredo Dinis & José M. Curado (orgs.) *Consciência e Cognição*. Braga: Publicações da Faculdade de Filosofia da UCP, pp. 79-96.

\_\_\_\_ (2004b) Introdução: linguagem, cultura e cognição, ou a Linguística Cognitiva. In Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves (orgs.) *Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva*. Vol. I. Coimbra: Almedina, pp. 1-18.

\_\_\_\_ (no prelo) Palavras e conceitos no tempo: para uma onomasiologia diacrónica e cognitiva do Português. In Fátima Silva *et al.* (orgs.), *Dar a Palavra à Língua. Homenagem a Mário Vilela*. Porto: Faculdade de Letras do Porto.

TAYLOR, John R. (1995) *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Clarendon Press.

### Anexo

	P50	B50	P70	B70	P90	B00
<b>ÁRBITRO</b>						
árbitro	615	91	414	18	189	231
juiz	56	147	33	200	34	105
juiz de campo	46	0	9	0	0	0
"ref(eree)"	0	1	0	0	0	0
referi	0	0	0	0	0	0
refre	0	0	0	0	0	0
<b>ÁRBITRO AUXILIAR</b>						
árbitro auxiliar	0	0	0	0	2	0
árbitro assistente	0	0	0	0	0	0
assistente	0	0	0	0	2	0
auxiliar	26	12	20	32	33	27
2º/3º/4º árbitro	0	0	0	0	0	0
bandeirinha	0	1	16	8	0	0
fiscal de linha	43	0	38	0	27	0
juiz de linha	50	0	14	0	2	0
liner	0	0	1	0	2	0
<b>ATACANTE</b>						
atacante	80	66	42	169	21	383
avanzado	507	2	147	0	103	0
avante	0	51	0	2	0	0
dianteiro	111	11	64	0	21	0
"forward"	1	17	0	0	0	0
ponta-de-lança	3	1	63	37	57	0
<b>BALIZA</b>						
arco	0	79	0	1	0	0
baliza	630	0	406	1	256	0
cidadela	0	0	0	0	0	0
"goal"	0	41	0	0	0	0
gol(o)	1	0	0	141	0	115
malhas	13	0	15	0	10	0
marco	0	0	0	0	0	0
meta	11	30	0	7	0	8
rede	50	3	3	2	0	45
redes	209	52	60	19	17	10
vaia	0	0	0	0	0	0
<b>BOLA</b>						
balão	0	7	0	0	0	0
bola	1495	266	865	625	522	408
couro (inho)	0	89	0	0	0	0
esfera	0	6	0	0	0	0
esférico	178	0	179	0	35	0
pelota	0	47	0	0	0	0

ACTAS DO XX ENCONTRO NACIONAL DA APL

	P50	B50	P70	B70	P00	B00
<b>DEFESA</b>						
“(full-)back”	0	38	2	0	0	0
beque	0	0	0	6	0	2
bequeira	0	0	0	0	0	0
defensor	111	25	53	10	12	11
defesa	342	1	207	1	116	0
lateral	0	0	19	57	61	161
libero	0	0	16	39	23	4
zagueiro	0	92	0	218	0	276
<b>EQUIPA</b>						
conjunto	210	117	157	22	70	9
eleven	0	0	0	0	0	0
equipa/e	2086	190	1548	303	898	631
escrete	0	0	0	24	0	0
esquadra	1	2	0	1	0	0
esquadrão	0	13	0	1	0	0
formação	77	9	68	3	53	15
grupo	474	3	86	0	14	100
onze	195	50	96	0	23	0
plantel	3	32	3	2	14	0
quadro	7	400	9	11	3	0
team	24	376	32	0	0	0
time	2	3	0	1013	0	1013
turna	168	15	184	9	42	0
<b>EXTREMO</b>						
ala	0	1	1	0	10	12
extremo	413	3	106	0	10	0
ponta	15	4	3	81	2	15
ponteiro	3	93	0	37	0	6
<b>FALTA</b>						
falta	163	32	70	42	75	98
“foul”	0	44	0	0	0	0
golpe (baixo, irregular)	0	0	0	0	0	0
infra(c)ção	10	14	8	0	8	0
obstrução	12	1	4	0	1	0
transgressão	0	0	0	0	0	0
violação (das regras)	0	0	0	0	0	0
<b>FINTA</b>						
corte	0	0	0	1	0	1
drible(ing)	42	6	27	21	9	25
engano	0	0	0	0	0	0
“feint”	0	0	0	0	0	0
finta	20	14	14	0	2	0
ginga	0	0	0	1	0	0
lesa	0	0	0	0	0	0
manobra enganadora	0	0	0	0	0	0
simulação	0	0	6	0	3	0
<b>FORA-DE-JOGO</b>						
banheira	0	0	0	0	0	0
deslocação	31	0	15	0	4	0
fora-de-jogo	28	0	28	0	51	0
impedimento	0	12	0	16	0	9
“offside”	0	0	0	0	0	0
“off-side”	52	1	35	0	4	0
<b>GOLO</b>						
bola	52	0	2	0	2	0
“goal”	1	527	0	1	0	0

PARA O ESTUDO DAS RELAÇÕES LEXICAIS ENTRE O PORTUGUÊS EUROPEU

	P50	B50	P70	B70	P00	B00
gol	0	7	0	956	0	1326
gól	0	0	0	0	0	0
golo	1535	0	1303	0	1040	0
ponto	105	1	26	0	4	0
tento	324	187	107	3	50	0
<b>GRANDE PENALIDADE</b>						
castigo máximo	23	0	4	0	8	0
castigo-mor	0	0	0	0	0	0
falta máxima	0	1	0	0	0	0
grande penalidade	178	0	105	0	107	0
penalidade	14	9	3	0	0	1
penalidade máxima	7	12	1	2	1	0
penalti	0	0	0	0	0	0
pênalti	0	0	0	51	0	167
"penalty"	79	96	82	0	91	0
<b>GUARDA-REDES</b>						
arqueiro	0	105	0	0	0	0
"goal-keeper"	0	1	0	0	0	0
goleiro	0	17	3	212	0	290
guarda-meta	0	0	0	0	0	0
guarda-rede	0	0	0	0	0	0
guarda-redes	427	0	171	0	133	0
guarda-vala	0	0	0	0	0	0
guarda-valas	0	0	0	0	0	0
guardião	59	9	39	0	30	0
"keeper"	8	37	9	0	0	0
porteiro	9	0	2	0	0	0
vigia	0	0	0	0	0	0
<b>JOGADA</b>						
jogada	342	63	234	296	93	313
lance	554	90	437	145	298	106
<b>JOGO</b>						
batalha	3	45	4	7	1	11
choque	0	10	0	2	0	0
combate	0	1	1	3	1	0
competição	50	20	23	11	18	11
confronto	7	8	12	1	27	23
desafio	375	4	191	0	57	0
disputa	10	25	0	21	0	6
duelo	5	3	7	2	7	3
embate	8	11	15	0	6	0
encontro	649	76	312	8	129	3
jogo	2198	486	1355	1025	1178	1147
justa	0	1	0	0	0	0
luta	258	146	79	82	20	19
"match"	4	454	2	0	0	0
partida	542	153	275	352	179	756
peladinha	0	0	0	0	0	0
peleja	8	171	3	0	1	0
prélio	12	6	9	0	8	0
prova	108	4	47	1	32	1
pugna	12	10	2	0	2	0
<b>MÉDIO</b>						
alfe	0	0	0	0	0	0
armador	0	0	2	5	0	3
cabeça de área	0	0	0	0	0	50
central	0	0	7	0	86	0
centro-campista	0	0	16	0	2	0
centro-médio	1	6	0	0	0	0

ACTAS DO XX ENCONTRO NACIONAL DA APL

	P50	B50	P70	B70	P00	B00
distribuidor (de jogo)	0	0	0	1	1	0
médio	401	7	57	6	68	0
meia	7	54	0	9	0	115
meio-campista	0	0	1	0	0	3
meio-campo	0	0	0	1	0	13
"midfield"	0	0	0	0	0	0
trinco	0	0	0	0	29	0
volante	38	1	0	0	0	34
<b>PONTAPÉ</b>						
chute	16	3	0	117	0	94
chuto	44	0	5	0	0	0
"kick(-off)"	0	1	0	0	0	0
panázio	0	0	0	0	0	0
pelotada	0	0	0	0	0	0
pontapé	219	0	76	0	47	0
quique	0	0	0	0	0	0
"shoot"	0	33	0	0	0	0
tiro	128	54	34	10	12	3
<b>PONTAPÉ DE CANTO</b>						
canto	255	0	67	0	31	0
chute de canto	0	0	0	0	0	0
comer (cómer)	1	49	10	28	0	4
escanteio	0	32	0	2	0	42
esquinado	0	0	0	0	0	0
pontapé de canto	7	0	10	0	25	0
tiro de canto	0	1	0	0	0	0
<b>PONTAPÉ LIVRE</b>						
chute (in)direto/livre	0	0	0	0	0	0
falta	0	18	0	31	5	104
"free(-kick)"	0	0	0	0	0	0
livre (directo, indirecto)	135	0	81	0	83	0
pontapé livre	6	0	2	0	1	0
tiro direto	0	0	0	0	0	0
tiro livre (direto, indirecto)	0	2	0	1	0	0
<b>TREINADOR</b>						
mister	0	0	1	0	0	0
professor	0	0	0	0	0	2
técnico	96	114	81	470	84	1135
treinador	113	14	55	86	125	317
<b>Total</b>	<b>18047</b>	<b>5796</b>	<b>10556</b>	<b>7129</b>	<b>6963</b>	<b>9822</b> <b>58313</b>