

PROTÓTIPO EM LINGUÍSTICA COGNITIVA: O EXEMPLO DO PROTÓTIPO ESPACIAL¹

HANNA JAKUBOWICZ BATORÉO
(Universidade Aberta - Lisboa)

Procurando responder à pergunta: "Where Does Prototypicality Come From?", Geeraerts (1988) refere quatro hipóteses diferentes para abordar esta questão, especificando-as como a hipótese (i) fisiológica, (ii) referencial, (iii) estatística e (iv) psicológica. Segundo este autor, todas elas foram, em diferentes alturas, propostas e/ou formuladas por E. Rosch (1973), facto que permite defender a sua complementaridade.

O presente estudo, situado no enquadramento teórico da Linguística Cognitiva, foca a elaboração de uma proposta de definição do protótipo espacial do Português Europeu, conforme defendido em Batoréo (1996/2000).

Em primeiro lugar, entende-se que a prototipicidade espacial resulta basicamente da estrutura fisiológica do aparelho perceptual humano. Considerando-se, porém, este critério manifestamente insuficiente para a descrição do Espaço, recorre-se ao critério referencial, que gerou o modelo de *parecença (semelhança) de família*.

Defende-se, por isso, em segundo lugar, que certas instâncias de uma categoria partilham mais atributos com outras instâncias da mesma categoria do que com alguns dos seus membros periféricos.

Por seu lado, e em terceiro lugar, o factor psicológico, faz otimizar a riqueza conceptual pelo destacamento de pormenores estritamente relacionados entre si, assim como pela sua integração no mesmo conceito, tornando-o mais "denso" e, simultaneamente, mais económico.

Os três critérios acima referidos encontram-se implicitamente subjacentes a toda a Linguística Cognitiva e, muito especialmente, à proposta que L. Talmy (1983, 1988) apresenta para a construção de um "esqueleto espacial" de uma

língua particular. No nosso caso, trata-se da apresentação do "esqueleto espacial" do Português Europeu, definido segundo os critérios acima especificados. No esquema de Deslocação da Figura em relação ao Fundo distingue-se entre a Localização e o Movimento, sendo este construído em função dos marcos constituintes do Fundo: a Origem, o Percurso, o Alvo e a Direcção, tendo simultaneamente em consideração a Causa e o Modo em que a Deslocação é efectuada. Consideramos que estes três critérios devem ser conjugados e integrados no quadro explicativo com um factor especificamente linguístico - o estatístico -, na medida em que, deste modo, se traduz a frequência de ocorrência dos itens lexicais para o grau de representatividade prototípica dos conceitos a eles subjacentes. Conjugando o factor frequência² com o de parecença de família, defende-se que o peso de um atributo conceptual não depende apenas do papel que este desempenha dentro de uma família de aplicações, mas também é medido pela frequência relativa com que é experimentado no uso linguístico.

O léxico usado na expressão do Espaço num idioma particular, isto é, os marcadores linguísticos para tal utilizados, pertencem ao seu vocabulário fundamental, comum e frequente. A análise que aqui se efectuará com base nos dados recolhidos do *Português Fundamental* (1987) mostrará como isto se verifica a respeito do Português Europeu. A frequência é um dos dois critérios estatísticos utilizados para aferir o índice de uso de cada unidade do léxico do *corpus* referido³.

O vocábulo mais frequente em Português é o verbo *ser*, com 34760 ocorrências, mas, em geral, são as palavras gramaticais, muitas vezes homógrafas, que apresentam frequências mais altas, de modo análogo ao que se verificou no estabelecimento de outros vocabulários fundamentais⁴. A análise de frequência das palavras gramaticais é importante para o nosso estudo, na medida em que elas abrangem, por exemplo, os marcadores espaciais tais como as preposições locativas⁵. No entanto, o facto de o *Português Fundamental* não analisar os vocábulos gramaticais de alta frequência, nem facultar dados mais pormenorizados acerca das formas homógrafas, inviabiliza uma pesquisa bem fundamentada em relação a três das preposições locativas mais frequentes *a*, *em* e *de*, dado que estas apresentam formas homógrafas com outros vocábulos gramaticais. Com base nos dados disponíveis e tendo consciência de todos os riscos que uma tarefa deste tipo acarreta, procurar-se-á, no entanto, estabelecer para estas três preposições uma ordem aproximativa de valores de frequência - em milhares de ocorrências - com o intuito de as inserir numa proposta de quadro comparativo dos marcadores espaciais atestados no *Português Fundamental*.

No que refere à preposição locativa *a*, o cálculo da sua frequência pode ser avaliado em função de dados referentes a duas formas: *a* e *o*. A forma *a* (preposição, artigo definido, demonstrativo, pronome pessoal) apresenta 21907 ocorrências e a forma *o* (artigo, demonstrativo, pronome pessoal) tem 14949.

Comparando ambas as frequências, estima-se que à preposição *a* corresponderão cerca de sete mil ocorrências, proposta resultante da diferença entre estes dois valores. Ela é superior aos valores verificados para a preposição *em* - 5319 (parcialmente homógrafa⁶) e para a preposição *por* - 6143, mas inferior ao valor da frequência da preposição *para* - 8939. O caso da preposição homógrafa *de* (frequência global de 33160) apresenta-se como o caso ainda mais difícil e de cálculo aproximativo mais arriscado do que as homógrafas citadas acima, dada a sua riqueza polissêmica: neste caso pode tratar-se quer do marcador espacial de Origem, quer de Posse, quer ainda de Modo. Avançar uma aproximação muito vaga - na ordem de 10 a 15 mil ocorrências - parece ser o único tipo de referência possível de estabelecer, neste momento, para fins comparativos.

Tendo em consideração o valor estimativo de muitos destes cálculos, pode adiantar-se, no entanto, que, genericamente, as preposições locativas básicas apresentam todas valores superiores a cinco mil ocorrências: *para* (9 mil) *a* (aprox. 7 mil), *por* (6 mil), *em* (aprox. 5 mil) e *de* (de 10 a 15 mil ocorrências), constituindo, assim, um dos subgrupos de unidades linguísticas mais comumente utilizados na Língua Portuguesa. As outras preposições, advérbios ou partes nucleares de locuções prepositivas e/ou adverbiais apresentam índices de frequência muito mais baixos, com a excepção, apenas, da preposição *depois* (3470). Os restantes valores são muito inferiores: *até* (prep.) - 628, *dentro* - 608, *fora* - 489, *cima* - 373, *sobre* - 322, *desde* - 312, *baixo* (adv.) - 267, *antes* - 208, *frente* - 151, *perto* - 149, *longe* - 95 ocorrências, etc.

É óbvio, no entanto, que o vocabulário frequente não é constituído apenas pelo vocabulário gramatical. Os verbos básicos de expressão de Existência, Posse e Localização - tal como estudados por Lyons (1975), E. Clark (1978) e Cláudia de Lemos (1987), para o Português, - têm, até, índices de frequência superiores aos das preposições acima citadas. Assim, no Português Europeu, a frequência global do verbo *ser* é de 34740, *ter* - 12968, *estar* - 8268 e *haver* - 6350. Os verbos deícticos *ir* e *vir* são igualmente frequentes: *ir* com 6724 ocorrências e *vir* com 2385, enquanto a maioria dos verbos lexicalizados (incluindo, no entanto também, os seus empregos gramaticalizados, como, por exemplo, variantes aspectuais nas perífrases verbais) não ultrapassa, em geral, as mil ocorrências (*andar* - 1069, *chegar* - 1051, *passar* - 966, *pôr* - 867, *levar* - 727, *sair* - 804, *entrar* - 443, *meter* - 329, *trazer* - 287, *encontrar* - 279, *tomar* - 234, *cair* - 150, *correr* - 137, *pegar* - 128, *agarrar* - 91, etc.). Um terceiro grupo de altíssima frequência no vocabulário português tem, igualmente, carácter espacial. Trata-se da Deixis com o marcador *lá* de 4408 ocorrências, *aquí* - 3764, *cá* - 1580, *ali* - 1376, *aí* - 1136 e *além* - 276.

A comparação efectuada entre os grupos (1), (2) e (3) (Cf. Quadro Sinóptico em Anexo) demonstra que, à semelhança do que acontece noutras línguas particulares, a expressão básica do Espaço no Português Europeu é

assegurada por um número muito reduzido de itens. Tomando em consideração o vocabulário do *Português Fundamental* - isto é, o que tem a frequência superior a quarenta ocorrências - trata-se de uma centena (cento e quatro) marcadores espaciais nucleares, o que constitui cerca de cinco por cento do vocabulário global do *Português Fundamental*. A marcação dos sucessivos limiares de frequência permite verificar que o número dos marcadores espaciais vai baixando à medida que a frequência de ocorrências em apreciação vai subindo. Surgem, assim, sessenta e nove vocábulos com frequência superior a cem, mas menos de metade deste número (vinte e nove) se for considerado o léxico com frequência alta, isto é, superior a quinhentas ocorrências. Se estimarmos, na globalidade, as ocorrências relativas às locuções prepositivas e adverbiais - e isto tendo em conta que nelas se podem combinar um, dois ou mais elementos nucleares (como em, por exemplo, *por baixo de*) - o total rondará, aproximadamente, cinquenta mil, enquanto no caso dos verbos, o número será de oitenta mil e no caso dos pronomes deícticos, de treze mil ocorrências. Na globalidade, tratar-se-á de mais de cento e quarenta mil ocorrências, o que irá perfazer, aproximadamente, um quinto de todas as palavras atestadas no *Português Fundamental*.

As estimativas globais levam a concluir, por conseguinte, que o léxico espacial que constitui cerca de cinco por cento dos vocábulos listados permite gerar, aproximadamente, vinte por cento do *Corpus* reunido o que implica uma enorme produtividade dos marcadores espaciais básicos. Uma análise pormenorizada demonstra como esta capacidade produtiva se organiza e estrutura (Cf. Batoréo, 1996/2000).

Repare-se que a constatação de que um número muito reduzido de itens linguísticos é responsável por uma parte substancial da expressão linguística de um idioma - e, neste caso específico, da expressão espacial entendida no sentido lato da Teoria Localista - é um argumento que vem ao encontro dos pressupostos talmianos da existência de um esqueleto espacial que estrutura todo o funcionamento da linguagem. Observe-se, no caso presente, que esta estrutura se apoia - no caso de frequências altas ($f > 500$) - quase equitativamente nos verbos espaciais do Grupo 2 ($n = 16$), por um lado, e, por outro, nos seus satélites ($n = 13$), isto é, os itens preposicionais e adverbiais do Grupo 1 ($n = 8$), assim como nos deícticos do Grupo 3 ($n = 5$).

Independentemente do nível de frequência apreciado, o grupo mais numeroso é o dos verbos (Grupo 2), constituído - no *corpus* do *Português Fundamental* - quase exclusivamente por radicais, verificando-se apenas três casos de sufixação, isto é, de satélites verbais, nos verbos *agarrar* (91) *desaparecer* (54) e *atirar* (50). Na formação dos verbos, os satélites surgem, marcadamente, apenas para as frequências abaixo de 40 (isto é, fora do escopo do

Português Fundamental), conforme se pode observar nos exemplos: *deslocar* (39), *substituir* (36), *regressar* (35), *integrar* (32), etc.

É curioso observar que entre os verbos, quatro deles - *ir*, *vir*, *levar* e *trazer* - têm carácter deictico, sendo os três primeiros de frequência muito alta. Os verbos de frequência superior - *ser*, *ter*, *estar* e *haver* - são também quatro e constituem um grupo que partilha as características já referidas de Existência-Posse-Localização. Pertence a este grupo também o verbo *ficar* (1575), variante aspectual do verbo *estar*, sendo as mesmas características partilhadas, também, por alguns verbos plenos de frequência muito mais pequena, indicadores de Existência-Localização, como *encontrar-se* (279)⁸, *existir* (222) e *faltar* (126), ou da sua introdução/(eliminação) *aparecer* (390), *surgir* (82) e *desaparecer* (54). Seguem-se os verbos de movimento efectuado pela própria Figura (= Locutor): os verbos básicos de movimentos - *andar* (1069), *seguir* (176), *deslocar* (39) e *viajar* (32) -, de indicação da Origem - *sair* (804), *partir* (204) -, do Alvo - *chegar* (1051), *entrar* (443) e *meter-se* (329)⁹ -, do Percurso - *passar* (966) e *passear* (52) - e do Modo como o movimento é efectuado - *cair* (150), *correr* (137) *saltar* (73), etc. Um outro subgrupo é constituído pelos verbos relativos ao movimento que o Locutor efectua em relação à Figura exterior a ele: *pôr* (867), *tirar* (640), *deixar* (599)¹⁰, *meter* (329), *tomar* (234), *deitar* (214), *levantar* (148), *tocar* (125), *pegar* (128), *agarrar* (91), *sentar* (73), etc., sendo três deles, especificamente, referidos a três tipos básicos de Postura¹¹: *deitar*, *levantar* e *sentar*. O Modo como o movimento é efectuado exprime-se por *puxar* (87), *mexer* (79), *encher* (53), *largar* (41), *arrancar* (40), etc. Um último subgrupo é constituído pelos verbos que exprimem a intenção de descoberta de uma localização desconhecida - *procurar* (229), *buscar* (155) - e a sua conclusão - *encontrar* (279)¹².

Os pouco mais de oitenta verbos aqui agrupados (n = 82) constituem uma base verbal para a expressão do Espaço em Português, preenchendo o esquema de Deslocação proposto por Talmy. Este esquema dispõe, no entanto, de um núcleo que é bem demarcado para as frequências superiores a cem, mas que já se começa a delinear com regularidade e solidez nos verbos frequentes, que são apenas dezasseis, pouco mais de uma terça parte dos que aparecem no limiar mais baixo. Surge, aqui, o grupo dos quatro verbos de Existência-Posse-Localização de maior frequência (*ser*, *ter*, *estar*, *haver*)¹³, a que se juntam os verbos *ficar* e *viver*, três dos quatro verbos deicticos mais frequentes (*ir*, *vir*, *levar*) na expressão da Direcção, o verbo básico de Movimento (*andar*) e um para cada marco importante do Movimento - Origem (*sair*), Percurso (*passar*), Alvo (*chegar*) -, assim como dois verbos de Movimento relativos à deslocação de uma Figura exterior ao Locutor: *tirar* e *pôr*¹⁴. O verbo *deixar* é altamente polissémico¹⁵, apresentando duas acepções espaciais: *deixar1*, no sentido de o Agente (= Figura) abandonar um local (marcação de Origem) e *deixar2* em

relação a uma Figura, colocada num local pelo Agente (≠ Figura), continuando ainda o Agente o seu próprio Movimento. Com estes dezasseis verbos alcança-se um esqueleto espacial *mínimo*, pressupondo-se que cada um dos seus elementos constitua um marcador verbal básico, ou seja, o fundamento que sirva de suporte para a construção da expressão espacial do Português.

Alguma da informação adicional acerca dos verbos espaciais é fornecida pelo *Inquérito Principal de Disponibilidade do Português Fundamental*, no âmbito da área temática dos *Meios de Transporte*¹⁶. Encontram-se aí referidos os verbos espaciais (com as respectivas frequências) utilizados especificamente na área dos transportes, como, por exemplo, *andar* (321) - abrangendo *andar de* (270), assim como várias expressões fixas do tipo: *andar a pé*, *andar a bates*, *andar à boleia*, etc. - , *viajar* (252), *voar* (183), *conduzir* (179), *guiar* (134), *navegar* (113), *ir de/ a* (97), *apanhar* (um meio de transporte) (96), *transportar* (94), *perder* (um meio de transporte) (56), *entrar* (61), *chegar* (57), *sair* (56), *caminhar* (50), *subir* (46), *aterrar* (44), *descer* (42), *cavalgar* (39), *levantar* (37), *ultrapassar* (36), *arrancar* (33), *embarcar* (33), *montar* (31), etc. Esta pequena listagem mostra como é importante analisar algumas das entradas polissémicas apenas numa das suas várias acepções, e como isto se traduz ao nível de frequência verificada ao nível de um inquérito específico. É importante verificar, também, a ocorrência de alguns verbos de Modo que não surgem no *Português Fundamental*, tais como *voar*, *guiar*, *navegar*, *embarcar*, etc.

Enquanto o sistema básico verbal (Grupo 2) assegura a expressão da Deslocação - ou seja, Localização e Movimento -, aos satélites (Grupos 1 e 3) cabe o papel de assegurar as relações espaciais tanto ao nível interno do sistema linguístico (Grupo 1) como em relação ao Universo de Referência (Grupo 3). Os dezanove satélites do Grupo 1 são constituídos por marcadores tradicionalmente classificados como preposições (*a*, *para*, *em*, *de*, *até*, etc.), advérbios (*baixo*, *dentro*, *embora*, etc.) ou nomes (*cima* e *frente*). Destaca-se aqui, claramente, o grupo das cinco primeiras preposições de altíssima frequência (acima das cinco mil ocorrências): *de*, *para*, *a*, *por*, *em* e que correspondem ao esquema básico de Deslocação do enquadramento talmiano: *EM* constitui o suporte prototípico para o padrão de Localização, *DE* é o suporte de Origem, *POR* de Percurso e *PARA* de Alvo, enquanto *A* é prototipicamente considerada a preposição de Direcção. Os restantes satélites deste grupo designam as relações relativas aos eixos Vertical (*cima*, *sobre*, *baixo*, etc.) e Sagital (*frente*, *atrás*, *trás*, etc.). às de relações de Inclusão/Exclusão (*dentro*, *fora*, *embora*), Proximidade/Afastamento (*perto*, *longe*, etc.) e Posterioridade/Anterioridade (*depois*, *antes*) e referem os marcos Limitrofes de Origem (*desde*) e de Alvo (*até*). É interessante verificar que - de um modo análogo ao que aconteceu no caso atrás discutido dos verbos espaciais - também no caso dos satélites, um núcleo básico que define as relações do esqueleto espacial da língua fica bem delineado para as frequências altas,

aparecendo os valores aproximados na designação dos pólos opostos. É o que se verifica na orientação Vertical - *cima* (373), *sobre* (322) e *baixo* (267) -, de Inclusão - *dentro* (608), *fora* (489) - ou de Marcos Limitrofes - *até* (628) e *desde* (312). No entanto, enquanto para os verbos este núcleo emerge nas frequências superiores a quinhentos, para os satélites prepositivos e adverbiais este limiar é relativamente mais baixo e menos homogêneo. Se para a Inclusão ele se situa na ordem de quinhentas ocorrências e, para a Verticalidade, de trezentas, o pólo marcado da orientação Sagital surge no limiar mais baixo - *frente* (151) -, só mais tarde aparecendo as designações do pólo oposto *atrás* (84) e *trás* (78), assim como as outras designações do pólo marcado: *perante* (47) e *diante* (30). No mesmo limiar, fica, também, estabelecida a relação Proximidade/Afastamento, primeiro com o pólo marcado *perto* (149), depois com o pólo oposto *longe* (95) e, a seguir, com as outras designações de Proximidade/ Vizinhaça: *cerca de* (62) e *junto* (adv.) (54). É, também, nesta ordem de valores que se situa um segundo grupo dos termos relativos à Verticalidade: *abaixo* (71), *acima* (70), *debaixo* (42) e *sob* (42). A determinação dos Marcos Limitrofes apresenta características de alguma dispersão: primeiro é estabelecido o marco de Alvo *até* (628), depois o marco de Origem *desde* (312), o marco da posição do Meio *entre* (225) e, posteriormente, o termo de Translimitação *através de* (72). É aparentemente surpreendente a dispersão ocorrente no caso da oposição, considerada como sendo de Posterioridade/ Anterioridade *depois* (3470) e *antes* (208). Estes números parecem reflectir, em grande parte, os usos de *depois* que não pertencem à oposição referida, mas a empregos fáticos de carácter discursivo¹⁷.

É curioso verificar, igualmente, a ausência no Grupo I da marcação do eixo de lateralidade *esquerda/ direita*. No *Português Fundamental* surge apenas o substantivo *direito* (115) e o adjectivo polissémico *direito* (39), mas na lista de locuções¹⁸ não se encontram atestadas as expressões *à/ para a esquerda* ou *à/ para a direita* - como, à partida, seria de esperar - embora ocorram as expressões *a direito* (2) e *a torto e a direito* (1). A ocorrência destas expressões de baixíssima frequência relativas à acepção *direito* = *recto*, *em frente* - na total ausência de expressões de lateralidade - aponta no sentido de o adjectivo citado abranger apenas esta acepção¹⁹.

Numa apreciação geral é interessante verificar que, no que refere os eixos de orientação, os termos marcados positivamente (*dentro*, *cima*, *frente*, *perto*, *até*, etc.) apresentam-se sempre com maior frequência do que os seus pares não-marcados (*fora*, *baixo*, *trás*, *longe*, *desde*, etc.), o que reflecte a sua prioridade na conceptualização espacial. Por outro lado, os termos relativos aos eixos principais realizados, isto é, o Vertical e o Sagital, apresentam uma característica já anteriormente verificada para os verbos, que se traduz em duas vagas de ocorrências dos respectivos termos: um no limiar de frequência muito alta e outro, num marcadamente mais baixo. Assim, para a Verticalidade, surgem, num

primeiro núcleo, *cima* (373), *sobre* (322) e *baixo* (267) e, posteriormente num segundo, *abaixo* (71), *acima* (70), *debaixo* (42) e *sob* (42). Para a orientação Sagital - embora de um modo inicialmente mais polarizado - observa-se sensivelmente o mesmo: primeiro surge *frente* (151), depois *atrás* (84) e *trás* (78) e, num segundo núcleo, *perante* (47) e *diante* (30).

Os satélites do Grupo 3 são, por sua vez, constituídos por advérbios deícticos de carácter locativo de divisão trenária (*aquí, aí, ali*) e binária (*cá e lá*), assim como - mas com frequência muito mais baixa do que os outros cinco elementos - pelo deíctico marcado quanto à distância (*além*). Por outro lado, nem *aquém* nem *acolá* ocorrem no *Português Fundamental*.

Uma das características funcionais mais importantes dos satélites de ambos os grupos (Grupo 1 e Grupo 3) é o facto de poderem entrar em combinatórias, agrupando-se em dois ou três e criando locuções prepositivas ou adverbiais, deícticas ou não. O agrupamento dos satélites é sujeito a restrições formais. Os agrupamentos mais antigos do ponto de vista da história da língua, tais como p.ex., *acima* e *adentro*, apresentam a grafia conjunta e a sua classificação formal (tradicional) de uma única unidade (neste caso, o advérbio). Quando há duas preposições numa locução, elas podem ser diferentes uma da outra (como em: *por fora de*) ou iguais (como em: *de dentro de*), sendo a segunda preposição quase sempre *de* ou, pontualmente, *a*, como por exemplo em *até a*. No lugar da primeira (ou única) preposição, surge, predominantemente, uma das cinco preposições mais frequentes, como em *de fora, por dentro, para além de*, etc., podendo, no entanto, este lugar ser também ocupado pelas preposições que apresentam menos ocorrências, como em *perto de, antes de*, etc. Um outro paradigma aceita a combinação de uma ou duas preposições (conforme se trate de uma locução adverbial ou prepositiva) com um advérbio (deíctico ou não), ou então com um nome. Surgem²⁰, assim por exemplo, as locuções criadas com o advérbio *baixo* - *para baixo de, por baixo de, de baixo, por baixo* (com 78 ocorrências na totalidade) - com o advérbio *dentro* - *de dentro de, para dentro de, por dentro de, de dentro, para dentro* e *por dentro* (com o total de 68 ocorrências), ou com *fora* - *de fora de, para fora de, por fora de* e *por (aqui/ aí/ ali) fora* - apresentando 80 ocorrências no total.

Os nomes mais frequentes que fazem parte destas locuções são *frente* e *cima*²¹, incluídos no Grupo (1). De entre estes dois, *cima* utiliza-se, exclusivamente, como elemento de locuções, conforme ilustram os seguintes exemplos: *em cima de, de cima, em cima, para cima, de cima de, para cima de, por cima de* e *por cima* (totalizando 373 ocorrências), enquanto *frente*, além da designação da parte dianteira de uma entidade (como, por exemplo, um carro, um edifício ou um ser vivo), está na base de locuções tais como: *à frente de, em frente a, em frente de, à (minha/ tua/ nossa frente), de frente, em frente e para a frente* (com o total de 129 ocorrências).

Outros nomes²² que entram no mesmo tipo de paradigma são, por exemplo, os vocábulos que funcionam como marcos de Deslocação em função dos quais são definidas as relações espaciais, tais como *lado* (com frequência de 487 ocorrências), *meio* (346), *pé* (330), *volta* (192), *lugar* (160), *base* (154), *fundo* (147), *centro* (121), *entrada* (75), *caminho* (61), *saída* (53), *local* (47), *roda* (38), etc. Como *pé*, fazem também parte deste grupo os nomes comuns que designam partes do corpo, como por exemplo, *cabeça* (1580), *mão* (173), *braço* (54), *costas* (33), etc., estando na origem de expressões²³ espaciais que apresentam algum grau de fixidez do ponto de vista formal. Surgem, assim, por exemplo, as expressões de Postura *estar deitado de costas, de barriga para baixo, de lado* ou *estar virado de costas*, de Localização *estar à mão, estar à cabeça* (= liderar) ou de Modo de Deslocação *estar de braços cruzados, ir de mãos dadas*, etc. Apesar do alto índice de frequência, estes nomes não foram incluídos no Grupo (1), partindo-se do princípio - arbitrariamente adoptado neste caso - de que as frequências destes vocábulos não reflectem tanto os usos espaciais das locuções em que eles entram, mas as diversas acepções que estes vocábulos têm na totalidade. Assim, por exemplo, o nome polissémico *pé* faz, entre outros, parte de algumas locuções prepositivas e adverbiais, de baixa frequência individual, que designam o Modo, a Postura ou a Proximidade: *ao pé, de pé, em pé, de ao pé de, do pé de, para ao pé de* e *para o pé de*, apresentando o total de sessenta ocorrências. Observem-se, igualmente, os casos da expressão de Percurso ou Modo nas locuções constituídas com o nome *caminho*: *a caminho de, de caminho* e *em caminho* (com o total de 15 ocorrências), da expressão de Proximidade ou Lateralidade não-específica, no caso das locuções com *lado*: *ao lado de, do lado de, para o lado de, ao lado, de lado, do lado* e *para o lado* (com o total de 84 ocorrências) ou da expressão do Movimento circulatorio e de Direcção no caso das locuções construídas com o nome *volta*: *de volta de, em volta de, por volta de, à (minha/ nossa/ sua) volta, de volta* e *em volta* (com 55 ocorrências, na totalidade). Note-se a polissemia existente em todos estes casos de combinatória rica, assim como a baixa frequência das locuções apreciadas individualmente.

Apreciada na sua globalidade, a grande riqueza potencial das combinatórias formais entre os satélites permite a marcação pormenorizada de Vizinhança (*à volta de, no meio de*, etc.), Proximidade (*ao pé de, ao lado de*, etc.), Percurso (*no caminho, a caminho*, etc.), Postura (*de pé, em pé*, etc.) e Modo de Deslocação (*a pé, à volta*, etc.). Foram precisamente estas características das locuções criadas segundo o paradigma (b) que nos fizemos, à partida, excluir do Grupo (1) os nomes comuns, com excepção de *cima* e *frente*.

O fenómeno de agrupamento que aqui está em discussão não se estende, no entanto, a todos os satélites aqui contemplados. Existem preposições como, por exemplo, *sobre*, ou advérbios, como *embora*, que não entram em

combinatórias com outros satélites. É de salientar, também, que a preposição *de* apresenta uma característica típica de todas as preposições do subgrupo mais frequente (com excepção de *para*²¹) ao aparecer em formas sincréticas com os advérbios deícticos de lugar, dando origem a contracções de tipo *daí, dalém, dali*, etc., tal como acontece com as preposições *a, de, em e por* quando criam formas contraídas²⁵ com os artigos, demonstrativos, indefinidos, pronomes pessoais e relativos interrogativos.

A discussão aqui apresentada, relativa à organização interna do léxico espacial português em função do factor frequência, permite avançar com alguns dos elementos da descrição do esqueleto conceptual pretendido (Cf. Batorêo, 1996/2000).

Recapitulando, o esqueleto espacial do Português Europeu (V. Quadro sinóptico) - na concepção que lhe foi atribuída no enquadramento teórico de Talmy e definido com base nos dados estatísticos - é constituído por vinte e nove marcadores linguísticos considerados frequentes ($f > 500$), dos quais mais de metade é formada por verbos (radicais verbais) e outra parte pelos seus satélites, isto é, predominantemente, preposições e advérbios. Cerca de metade dos satélites e três dos verbos frequentes apresentam carácter deíctico, remetendo para o Universo de Referência exterior à linguagem. Os marcadores permitem definir a Existência, a Localização estática, o Movimento da Figura em relação ao Fundo, destacando-se, para este, os marcos fundamentais de Origem, Percurso, Direcção e Alvo. Distinguem-se as áreas referentes ao Locutor e ao Alocutário, tanto ao nível pontual como ao nível da extensão, podendo os participantes da interacção verbal corresponder à Figura e/ou ao Fundo, dando Origem à Deixis. Ao nível topológico distingue-se, apenas, a marcação de Inclusão, do Marco Limítrofe do Alvo e de Posterioridade.

Quando o escopo de análise muda para o limiar de frequência de cem ocorrências ($f > 100$), o número de marcadores linguísticos do Espaço sobe para o dobro, passando a realizar-se as relações tanto de Posterioridade/ Anterioridade, Inclusão e Superioridade/ Inferioridade no eixo Vertical, como o destaque para a relação de Suporte, assim como o Marco Limítrofe de partida (Origem), além do de chegada (Alvo), já verificado anteriormente. Ficam marcados, igualmente, os pólos positivos da orientação Sagital e de Proximidade, a Deixis, assim como a posição Média em relação a outros marcos já previamente localizados. As combinatórias dos satélites permitem marcar vários graus de Vizinhança, Proximidade, Modo de Deslocação, Postura e diversidade de Percursos efectuados. Não se verifica, no entanto, nenhuma marcação explícita dos pólos no eixo Lateral.

Constatando-se a existência de um leque grande de afastamento dos valores extremos de frequências abordadas no *Português Fundamental*, abordaram-se os valores relativos para a determinação do(s) núcleo(s) do vocabulário espacial

português, onde se verificaram uma regularidade e uma representatividade espacial significativas, principalmente ao nível do núcleo verbal. No conjunto dos satélites surge, igualmente, o mesmo tipo de sistematicidade, mas começa a delinear-se apenas para frequências muito baixas, até, nalguns casos, inferiores às adoptadas como o limiar do *Português Fundamental*.

Quadro sinóptico

Frequência de marcadores espaciais no Português Europeu (com base nos dados recolhidos do *Português Fundamental*, 1987).

GRUPO 1

Satélites (preposições, advérbios, elem. nucleares de locuções)

de (f> 10 mil), *para* (8939), *a* (aprox.7 mil), *por* (6143), *em* (aprox.5 mil), *depois* (3470), *até* (628), *dentro* (608), *fora* (489), *cima* (373), *sobre* (322), *desde* (312), *baixo* (267), *entre* (225), *antes* (208), *durante* (197), *embora* (152), *frente* (151), *perto* (149), *longe* (95), *atrás* (84), *contra* (prep.) (81), *trás* (prep.) (78), *através de* (72), *abaixo* (71), *acima* (70), *dantes* (69), *cerca de* (62), *junto* (adv.) (54), *perante* (47), *debaixo* (42), *sob* (42), *diante* (30).

GRUPO 2

Verbos espaciais

ser (34740), *ter* (12968), *estar* (8268), *ir* (6724), *haver* (6350), *vir* (2385), *ficar* (1575), *andar* (1069), *chegar* (1051), *passar* (966), *pôr* (867), *sair* (804), *levar* (727), *tirar* (640), *deixar* (599), *viver* (581), *entrar* (443), *aparecer* (390), *meter* (329), *apanhar* (328), *trazer* (287), *mandar* (285), *encontrar* (279), *tomar* (234), *procurar* (229), *existir* (222), *deitar* (214), *partir* (204), *voltar* (177), *seguir* (176), *abrir* (165), *buscar* (155), *cortar* (151), *cair* (150), *escolher* (150), *levantar* (148), *correr* (137), *pegar* (128), *faltar* (126), *fugir* (125), *tocar* (125), *ligar* (119), *mudar* (118), *morar* (105), *parar* (95), *agarrar* (91), *puxar* (87), *surgir* (82), *fechar* (81), *mexer* (79), *subir* (78), *juntar* (76), *aumentar* (74), *sentar* (73), *montar* (67), *descer* (59), *guardar* (55), *desaparecer* (54), *encher* (53), *passar* (52), *atirar* (50), *botar* (46), *saltar* (44), *virar* (43), *largar* (41), *arrancar* (40), *deslocar* (39), *dividir* (38), *estender* (38), *afastar* (37), *crescer* (36), *reunir* (36), *substituir* (36), *tapar* (36), *regressar* (35), *limitar* (34), *carregar* (33), *abalhar* (32), *colher* (32), *fixar* (32), *integrar* (32), *misturar* (32), *viajar* (32), *prender* (30), *transmitir* (30).

GRUPO 3

Satélites (advérbios deícticos de lugar)

lá (4408), *aqui* (3764), *cá* (1580), *ali* (1376), *aí* (1136), *além* (279).

Notas

- 1 Ver o estudo mais aprofundado em Batoréo 1996/2000.
- 2 "Statistical explanations of prototypicality state that the most frequently experienced member of a category is the prototype. At least, this is the simple form of the frequency model. It can also be combined with the family resemblance model. The weight of an attribute within a concept is then not only determined by its role within a family of applications constituting the category, but also by the relative frequency with which it is experienced (Rosch 1975)." (Geeraerts, 1988a, 208).
- 3 "Para aferir o índice de uso de cada unidade do léxico, o método utilizado prevê dois dados estatísticos: a frequência (o número de vezes que foi utilizada) e a repartição. O cálculo de repartição consiste em determinar o número de falantes diferentes que o utilizou para se referir a situações diferentes" (*Português Fundamental*, Volume II, Tomo I, 1987, 412). O limiar de frequência adoptado é de 40 ocorrências por vocábulo Cf. "Para o estabelecimento do limiar mínimo de frequência vários factores foram ponderados, nomeadamente a extensão do corpus, a percentagem de formas homógrafas a analisar, o número aproximado de vocábulos que o vocabulário deveria conter. Assim, e de acordo com a experiência de estabelecimento dos outros vocabulários fundamentais românicos, o limiar foi fixado em 40 ocorrências por vocábulo, depois de se terem feito experiências também com os limiares de 50 e de 30 ocorrências". (*Português Fundamental*, Volume II, Tomo I, 1987, 360). 500 ocorrências constituem o limiar considerado de alta frequência. (P.F. Vol. II, Tomo I, 1987: 317-318). Com base no primeiro destes dois critérios, foram destacados 2217 vocábulos considerados "fundamentais" do Português Europeu num *corpus* de cerca de setecentas mil ocorrências. Quanto ao vocabulário referido ao Espaço, é curioso verificar que tanto o nome *espaço* (40 ocorrências) como o verbo *deslocar* (39 ocorrências) se situam no limiar do *Português Fundamental*.
- 4 Cf. *Português Fundamental*, Volume II, Tomo I, 1987, 403. Conforme se observou na nota anterior, as formas gramaticais com mais de 500 ocorrências não foram analisadas no *Português Fundamental*. O Termo "forma" corresponde, no PF, à "palavra" (*Português Fundamental*, Volume II, Tomo I, 1987, 317-318).
- 5 Embora nem todas as acepções destas preposições sejam espaciais *stricto sensu* - registando-se casos idiossincráticos ou convencionalizados que uma análise de carácter espacial não explicita - consideramo-las como tais na óptica da Teoria Localista.
- 6 *Português Fundamental*, Volume II, Tomo I, 1987, 427-428. Veja-se o caso de homografia entre *nos* = "em + os" e *nos* = "a nós", facto que não permite interpretar o número de ocorrências atribuído, sem análise linguística mais pormenorizada.
- 7 Sublinhamos, mais uma, vez o carácter aproximativo destes cálculos na medida em que o *Português Fundamental* não nos fornece dados mais pormenorizados, nem análises efectuadas para os vocábulos gramaticais homógrafos com frequência superior a quinhentas ocorrências.
- 8 Apesar da existência das listas lematizadas no *Português Fundamental*, é praticamente impossível estimar o valor de ocorrências correspondente a *encontrar-se*.

sabendo que o total para o verbo *encontrar* - isto é, abrangendo, também, a forma com o *se* integrado - é de 279. Este tipo de informação pode ser, no entanto, fornecido pelo Corpus de Referência.

9 Cf. a nota anterior.

10 Neste caso, tão-pouco nos é possível discernir as ocorrências relativas à acepção locativa e não locativa do verbo *deixar*.

11 Conforme explicitam L. Seliar-Cabral e A. M. Roncada, a "*Postura* é a atitude reflexa do corpo em relação ao espaço", enquanto a "*Posição* é a postura característica da espécie" (em: "Ordem de Aquisição de Verticalidade, Prespectividade e Lateralidade em Crianças Paulistas" Comunicação apresentada no 3º ENI, Rio de Janeiro, (1978, 1)).

12 Cf. as notas anteriores referentes aos verbos *encontrar* e *encontrar-se*.

13 Cf. os recentes trabalhos de Batoréo e Batoréo & Duarte.

14 Numa análise um tanto alargada poderia propor-se a inclusão de mais dois verbos frequentes que podem ser analisados como espaciais, em função de troca de "deslocações": *comprar* (freq. de 605) e *vender* (freq. de 556).

15 Para o Português Europeu ver os estudos efectuados por Augusto Soares Silva (1997/1999).

16 *Inquérito Principal de Disponibilidade*, in *Português Fundamental*, Tomo 2, cap. 9, p. 179 e ss.

17 Verifica-se, mais uma vez, a insuficiência dos dados fornecidos pelo *Português Fundamental* para uma análise rigorosa que aqui se pretende e a necessidade para o efeito de consulta do *Corpus de Referência do Português Contemporâneo (CRPC)*.

18 *Português Fundamental*, 1987, vol I, Tomo I, 728 e ss.

19 Observe-se, a propósito, que os termos *direita* e *esquerda* surgem recuperados pelo *Inquérito Complementar de Disponibilidade do Português Fundamental*, na área temática da Vida Política (*Português Fundamental*, Tomo II, Cap. 1, p. 423).

20 *Português Fundamental*, Volume II, Tomo I, lista 4, 1987, 728-751.

21 Nem todos os dicionários classificam *címa* como nome. O *Dicionário do Português Básico*, por exemplo, dá-lhe uma classificação de "loc." (= locução). (Vilela, 1990, 164). O *Português Fundamental* classifica-o como "elemento adverbial de locução" (1987, 694).

22 Ver alguns dos exemplos recolhidos no *Português Fundamental*, Volume II, Tomo I, lista 4, 1987, 728-751.

23 Deste grupo, apenas algumas das expressões relativas a *mão* e a *pé* estão atestadas no *Português Fundamental*.

24 Ao nível coloquial esta excepção não se mantém, surgindo formas não aceites (ainda?) pela norma como, por exemplo, *prá frente!* = *para a frente!*

25 *Português Fundamental*, Volume II, Tomo I, 1987, 339-342.

Referências

- BACELAR DO NASCIMENTO, M.F.; M.L. GARCIA MARQUES & M. L. SEGURA DA CRUZ (eds.). (1984/1987). *Português Fundamental*. Vol. I e II, INIC/CLUL. Lisboa.
- BATORÉO, H. J. (1996/2000). *Contribuição para a Caracterização da Interface Expressão Linguística – Cognição Espacial no Português Europeu: Abordagem Psicolinguística da Expressão do Espaço em Narrativas Provocadas*. 2 vols., Dissertação de Doutoramento, FLUL, Lisboa, 1996. Publicada como *Expressão do Espaço no Português Europeu. Contributo Psicolinguístico para o Estudo da Linguagem e Cognição*, Textos Universitários de Ciências Sociais e Humanas, Fundação Calouste Gulbenkian, Fundação para a Ciência e a Tecnologia, Ministério da Ciência e da Tecnologia, Lisboa, 2000.
- BATORÉO, H. J. (1998) "Language Typology and Semantic Primitive of Space: Evidence from European Portuguese", *Actas do 1º Encontro de Linguística Cognitiva*. FLUP, Porto, 33-49.
- BATORÉO, H. J. & I. DUARTE (1999) "Aberturas de Narrativas: Incidência no Ensino de Português como Língua Materna e como Língua Estrangeira", *Palavras*, n° 15, Primavera 1999, APP, Lisboa, 39-58.
- CLARK, E. V. (1978). "Locational: Existential, Locative, and Possessive Constructions", in Greenberg, J. H., *Universals of Human Language*, Vol.4. *Syntax*, Stanford, California, 82-125.
- DE LEMOS, C. T. G. (1987). *'Ser' and 'Estar' in Brazilian Portuguese: with Particular Reference to Child Language Acquisition*. Tübingen, Gunter Narr Verlag.
- GEERAERTS, D. (1988). "Where Does Prototypicality Come From?", in: B. Rudzka-Ostyn (ed.) (1988) *Topics in Cognitive Linguistics*. Amsterdam /Philadelphia. John Benjamins Publishing Company, 1988. 207-230
- LYONS, J. (1975). "Deixis as the Source of Reference", in Keenen, E. L. (ed.), (1975), *Formal Semantics of Natural Language*. London & New York, C.U.P.
- LYONS, J. (1977). *Semantics*, Vol I & II, C.U.P., 1978.
- ROSCH, E. (1973). "Natural Categories", *Cognitive Psychology*, 328-350.
- SCLIAR CABRAL, L. & RONCADA, A. M. de G. (1978). "Ordem de Aquisição dos Eixos de Verticalidade, Perspectividade e Lateralidade em Crianças Paulistas", Comunicação apresentada no 3º ENL, Rio de Janeiro. Outubro de 1978, texto policopiado.
- SILVA, A. Soares (1997/1999). *A Semântica de Deixar: uma Contribuição para a Abordagem Cognitiva em Semântica Lexical*. Dissertação de Doutoramento, Universidade Católica de Braga, Faculdade de Filosofia, Braga, 1997. Publicada em: Textos Universitários de Ciências Sociais e Humanas, Fundação Calouste Gulbenkian, Fundação para a Ciência e a Tecnologia, Ministério da Ciência e da Tecnologia, Lisboa, 1999.

- TALMY, L. (1983). "How Language Structures Space", in: Pick & Acredolo (eds.) (1982), *Spatial Orientation: Theory, Research and Application*. New York & London. Plenum Press, 1983, 225-282.
- TALMY, L. (1988) "The Relation of Grammar to Cognition" in B. Rudzka-Ostyn (ed.) (ed.) (1988) *Topics in Cognitive Linguistics*, Amsterdam/ Philadelphia, John Benjamins Publishing Company, 1988,165-206.
- TAYLOR, J. R. (1989). *Linguistic Categorization: Prototypes in Linguistic Theory*, Clarendon Press, Oxford, 1991.