

Significacións gramaticais e sentidos colocacionais: ¿máis ca unha simple coincidencia?¹

Xavier Blanco Escoda²

Univesitat Autònoma de Barcelona

O artigo pon de relevo a grande área de intersección semántica que existe entre o conxunto de sentidos vehiculados mediante flexemas e derivatemas, por unha banda, e mediante colocacións léxicas, por outra. Saliéntase a importancia que esta cuestión pode ter respecto ó deseño de sistemas de tradución automática con interlingua gramatical. Emprégase o aparello teórico desenvolvido no marco da Teoría Sentido-Texto (TST), particularmente o concepto de función léxica. Faise unha breve presentación do proxecto InterGram.

Palabras clave: significacións gramaticais, sentidos colocacionais, interlingua, tradución automática, función léxica, Teoría Sentido-Texto, InterGram.

This paper highlights the big area of semantic intersection that exists between the set of meanings expressed by means of flexemes and derivatememes, on the one hand, and by means of lexical collocations, on the other. We underline the importance of this issue regarding the design of machine translation systems with grammatical interlingua. We use the theoretical frame of the Meaning-Text Theory (MTT), especially the concept of Lexical Function. A concise presentation of the InterGram project is done.

Keywords: Grammatical Meanings, Collocational Meanings, Interlingua, Machine Translation, Lexical Function, Meaning-Text Theory, InterGram.

1. Introducción

O obxectivo deste artigo é claro: trátase de poñer de relevo que existe unha grande área de intersección semántica entre os sentidos das colocacións léxicas e as significacións gramaticais (flexivas e derivativas) empiricamente observadas en lingüística xeral.

¹ Esta investigación contou co financiamento do Ministerio de Educación y Ciencia (España) no marco do proxecto I+D+i HUM 2005-08052-C02-01 InterGram (Interlingua Gramatical para Sistemas de Procesamiento Automático del Lenguaje Natural).

² Tradución de Cristina Veiga Novoa.

Tanto é así, que poderíamos pensar que estamos, basicamente, diante dos mesmos sentidos vehiculados a través de distintos medios de expresión. Defenderemos, así mesmo, a hipótese de que estes sentidos son universais e extremadamente especiais no que se refire á súa incardinación lingüística, xa que, sexa cal sexa a súa expresión formal, esta non é representable por medios distribucionais, senón que está lexicamente (no caso das colocacións) ou morfoloxicamente (no caso dos flexemas e derivatemias) condicionada.

Disto derivanse, ó noso xuízo, dúas consecuencias importantes (unha teórica e outra aplicada) para o deseño de ferramentas de recoñecemento automático ou de tradución: dende o punto de vista teórico, é imposible, *stricto sensu*, traducir unha colocación, do mesmo xeito que é imposible traducir un flexema ou un derivatema. Só poden recoñecerse e/ou xerarse, mais non traducirse. Para garantir que o sentido da lingua de partida se mantén na lingua de chegada requírirase unha representación intermedia de carácter puramente metalingüístico. Dende o punto de vista aplicativo, non é posible trata-las colocacións léxicas por transferencia, senón unicamente por interlingua. Ó longo deste artigo presentáremo-las bases conceptuais do noso razoamento e tentaremos xustifica-las conclusións apuntadas.

Antes de comezar precisarémo-lo uso que faremos, neste artigo, de tres termos importantes: *significado*, *sentido* e *significación*. O *significado* é unha das tres compoñentes do signo lingüístico. Por conseguinte, un significado sempre vai intimamente vinculado a un significante e a unha combinatoria. Insistamos, pois, en que o significado é sempre o significado dun significante. Non pode, polo tanto, falarse de significado independentemente dun signo lingüístico determinado nunha lingua dada.

O *sentido*, pola contra, corresponde á variante das paráfrases lingüísticas, pode coincidir ou non cun significado no marco dunha lingua determinada, é translingüístico e é previo á operación de lexicalización entendida como paso dunha representación semántica (semas e semantemas) a unha representación sintáctica profunda, que conta xa con unidades léxicas como nodos.

Finalmente, a noción de *significación* presenta un carácter moito máis instrumental, xa que se trata, basicamente, dunha etiqueta que nos permite reagrupar de maneira cómoda, aínda que só aproximada, conxuntos de sentidos interlingüísticos estreitamente emparentados. Así, por exemplo, a significación “plural” reagrupa sentidos moi semellantes nunha mesma lingua (p. ex. plural cuantitativo *versus* plural sortal) ou oposicións semellantes aínda que non idénticas dende unha perspectiva translingüística (p. ex. o plural do castelán oponse unicamente ó singular, mentres que o plural do árabe se opón ó singular e ó dual).

2. A noción de colocación

Na última década, o estudo das denominadas *colocacións* suscitou un grande interese e deu lugar a un gran número de publicacións en distintos eidos da lingüística aplicada (lingüística de corpus, lingüística computacional, didáctica das linguas estranxeiras...).

Ata o punto de que un panorama bibliográfico, mesmo superficial, destes estudos resultaría unha empresa titánica da que nos manteremos prudentemente á marxe.

A isto cómpre engadir que a propia noción de *colocación* dista moito de recibir unha definición unanimemente aceptada pola comunidade lingüística; mesmo nos atreveríamos a afirmar que existe un certo *cafarnaum* terminolóxico sobre esta noción. Tampouco é o noso propósito entrar neste debate e evitaremos escrupulosamente tanto considera-los múltiples parasinónimos do termo *colocación* como os seus variados significados. Adoptaremos, logo, de maneira axiomática, a definición seguinte (Mel'čuk 1995):

Unha colocación é un frasema semántico AB (S), onde o sentido (S) \supset (A), de xeito que (A) é comunicativamente dominante en (S) e A é seleccionado de maneira regular e non forzada, mentres que B é seleccionado de maneira non regular.

Tomemos un par de exemplos sinxelos: nas secuencias *error garrafal* e *negar rotundamente*, as unidades léxicas *error* e *negar* son seleccionadas de maneira libre, mentres que *garrafal* e *rotundamente* corresponden a colocacións seleccionadas polas súas bases (*error* e *negar* respectivamente) para expresa-lo significado 'intenso'.

Fagamos notar, entón, que contrariamente ó que ocorre noutros marcos de análise, de ningún modo consideramos colocacións secuencias AB (S) en que:

- 1) ben o sentido (A) $\not\subset$ (S) e (B) $\not\subset$ (S) (como, por exemplo, en *tocar las narices* ou *ser culo de mal asiento*), ben (B) \subset (S), pero (A) $\not\subset$ (S), sendo (A) a posición comunicativamente dominante³.
- 2) (A) \subset (S) pero ben (B) $\not\subset$ (S) (por exemplo, *llave inglesa*) ou (B) \subset (S) e asemade (C) \subset (S) (por exemplo, *agua con gas*).
- 3) (S) pode descompoñerse nunha relación semántica de tipo argumental (por exemplo, *el barco atraca en el puerto de Nloc* ou simplemente *beber agua*)⁴.

Tampouco consideramos colocacións secuencias de tipo frasemático nas que a fixación non é morfosintáctica senón puramente semántica; é dicir, expresións convencionais para vehicular determinados contidos de maneira estreitamente vinculada a unha situación de comunicación concreta (por exemplo: *recién pintado*, *reservado el derecho de admisión*, etc.).

³ Un sentido (A) é comunicativamente dominante en (S) se e só se (S) pode reducirse a (A) con perda de información pero sen cambio de referencia. Isto é, (S) é unha especie de (A).

⁴ Este último tipo de relación tense considerado, a miúdo, colocacional en tradicións lingüísticas que empregan o termo *colocación* para caracterizar calquera grupo de unidades que coaparecen cunha determinada frecuencia (calculada mediante diversos métodos: *mutual information*, *T-score*, etc.). Ó noso entender, trátase dun uso impropio do termo, posto que estamos perante restricións de selección totalmente descritibles por medios distribucionais clásicos.

Cómpre indicar que, se ben a clasificación de frasemas antes apuntada utiliza os mesmos elementos cá ofrecida no marco da Teoría Sentido-Texto (a través dos conceptos de frasema completo, cuasifrasema e semifrasema), difire da mesma en dous aspectos: por unha parte, non consideramos colocacións, senón unidades léxicas complexas (cuasifrasemas na terminoloxía da TST), casos como *agua con gas* ou *café con leche*, aínda que $(A) \subset (S)$ e $(B) \subset (S)$ porque, ó mesmo tempo, o sentido (S) inclúe unha compoñente ausente de (A) e de (B) á vez (a unidade léxica *agua con gas* non pode aplicarse, en linguaxe ordinaria, a auga que conteña gas freón, por exemplo, senón unicamente a aquela que conteña gas carbónico e ademais sexa potable, etc.)⁵. Por outra parte, tampouco consideramos colocacións, senón cuasifrasemas, unidades léxicas como *llave inglesa*, aínda que $(B) \not\subset (S)$, porque o sentido (B) se aplica só a un número moi restrinxido de unidades léxicas (o significado de *inglesa* respecto do de *llave* non se atopa na coocorrencia de *inglesa* con outras formas (como, por exemplo, *crema inglesa* ou *semana inglesa*); isto é, por unha parte, [*inglesa* | *llave*] \neq [*inglesa* | *crema*], e, por outra, [*inglesa* | *llave*] \neq [*inglesa* | *semana*]).

3. Fixación *versus* idiomaticidade

Sinalemos que o feito de que unha determinada colocación teña unha maior ou menor frecuencia de uso e se combine cun maior ou menor número de bases non está relacionado (como comunmente se pensa) co feito de que a colocación sexa menos estable. Dito doutro xeito, o grao de idiomaticidade é un parámetro independente, dende o punto de vista lóxico, do grao de fixación léxica. Unha secuencia como *llover a cántaros* é fixa e moi idiomática, mentres que unha secuencia como *herir de gravedad* presenta menor idiomaticidade mais, a grandes trazos, a mesma achega semántica e un grao de fixación comparable. En ambos casos, a combinatoria das colocacións coas súas bases é lexicamente restrinxida; é dicir, non pode representarse nun dicionario electrónico mediante restricións de selección de carácter sintáctico-semántico, senón unicamente mediante indicación das unidades léxicas precisas que aceptan a posición indicada; en ambos casos, as colocacións presentan comportamentos idiosincráticos tanto respecto a substitucións parasinonímicas (*llover mucho* pero **herir mucho*; *herir seriamente* pero **herir de importancia*; en cambio *pérdida grave*, *pérdida importante*, *pérdida de importancia* fronte a *pérdida de gravedad*, etc.) como respecto a equivalencias de tradución non controladas (**blesser de gravité*, *to injure seriously*; *schwer verwunden* fronte a **herir pesadamente*, etc.).

Dende o punto de vista teórico, non sería razoable pensar que os sentidos expresables colocacionalmente puidesen ser moi numerosos xa que, ó fin e ó cabo, a expresión colocacional dun sentido dado resulta particularmente antieconómica: empréganse

⁵ De feito, isto lévanos a elimina-lo concepto de *función léxica non estándar* (Mel'čuk 1995), que queda así asimilado a un caso especial de cuasifrasema. Procederemos deste xeito non tanto por razóns teóricas canto por consideracións de orde eminentemente práctica. En efecto, tendo en conta a estrutura adoptada no marco dos dicionarios electrónicos *flexSem* (Blanco 2001), resultáanos moito máis sinxelo introduci-las secuencias correspondentes a funcións léxicas non estándar como lemas do dicionario de n-gramas (cuasifrasemas, en xeral nominais) que tratalas como elementos de microestrutura no dicionario de unigramas.

decenas, centos e ata milleiros de formas distintas para a expresión dun mesmo sentido (se ignoramos algúns matices ou certos valores expresivos e estilísticos). Dende un punto de vista empírico, as investigacións efectuadas no eido das teorías colocacionais corroboraron que, en efecto, a expresión por medios colocacionais afecta unicamente a unha pequena cantidade de sentidos, arredor de sesenta sentidos de base, amais dalgunha das súas posibles combinacións. Alén disto, cómpre salientar un dato da máxima importancia: estes sentidos presentan un carácter universal. É dicir, que os sentidos potencialmente colocacionais son aproximadamente OS MESMOS para calquera lingua⁶.

4. A “tradución” das colocacións

Estes dous trazos das colocacións (o seu carácter lexicamente restrinxido e o reducido número de sentidos ós que afectan) resultan fundamentais para a tradución, pois a primeira fai impracticable o seu tratamento mediante asignación de equivalentes de tradución (postular regras sensibles ó contexto do tipo *a cántaros* | *llover* = *cats and dogs* resulta inviable⁷, a non ser que esteamos dispostos a acepta-la existencia de decenas de miles de regras *ad hoc* para cada par de linguas) e a segunda dános a chave para o tratamento lexicográfico contrastivo das colocacións: o cálculo do equivalente de tradución.

En efecto, para chegar a *it rains cats and dogs* a partir de *llueve a cántaros* debemos pasar pola codificación de *llueve a cántaros* como **Magn**(*llover*), a tradución de *llover* = *to rain* e o cálculo de **Magn**(*rain*) = *cats and dogs* a partir da consulta do dicionario monolingüe da lingua de chegada. En realidade, non se trata máis que de empregar-lo formalismo das funcións léxicas en lexicografía bilingüe automatizada.

Naturalmente, non tódolos cálculos que implican valores colocacionais son tan sinxelos. Pode suceder, e a miúdo sucede, que o dicionario da lingua de chegada simplemente non presente o valor axeitado da función léxica detectada na lingua de partida en relación coa base correspondente ó equivalente de tradución. Así, por exemplo, *tomber amoureux*, analizado como **IncepCopul**(*amoureux*) = *tomber*, non pode traducirse a partir da equivalencia por transferencia *amoureux* = *enamorado*, xa que nin **caer enamorado* nin ningún outro incoativo parece funcionar correctamente (agás se considerarmos posibilidades claramente menos comúns como *quedar prendado de*). Requirirase, polo tanto, activar unha regra de fisión por verbo soporte para pasar de *tomber amoureux* a *enamorarse*, que é un valor fusionado incoativo de *amar* (como *dormirse* o é de *dormir*⁸).

Con todo, trátase dun problema para o que dispoñemos de diversas vías de solución. De feito, tantas como regras de paráfrase para funcións léxicas, que inclúen substitucións

⁶ Deixando de lado as colocacións non estándar, que deben descubrirse empiricamente para cada lingua, pero que teñen moita menor importancia posto que xeralmente afectan a un número de bases moi reducido.

⁷ Véxase ó respecto Danlos (1994).

⁸ Nótese que **empecé a dormir a las 11* e **empecé a amar a Ana aquel verano* resultan secuencias forzadas en comparación con *me dormí a las 11* e *me enamoré de Ana aquel verano*.

sinonímicas, conversivas, derivativas e metonímicas, así como fisións/fusións de cópula, de negación, de verbo soporte, de hiperónimo, metafóricas e metonímicas. Para o inventario destas regras véxase Mel'čuk (1984).

5. Significados léxicos *versus* significacións gramaticais

No eido das significacións, cómpre facer unha distinción fundamental entre os significados léxicos e as significacións gramaticais. Se ben non se trata en absoluto de dúas categorías estancas, si é certo que ámbolos tipos de sentidos (tanto respecto a unha lingua dada como respecto ó conxunto de linguas naturais) presentan, globalmente, características ben diferenciadas.

De maneira xeral, poderíamos dicir que os significados léxicos constitúen un conxunto aberto e están directamente vinculados á realidade extralingüística (a invención dun novo procedemento ou o descubrimiento dun novo ente poden dar lugar a unha creación léxica dotada dun significado inexistente ata ese momento). Como consecuencia diso, podemos observar que os significados léxicos son moito máis numerosos cás significacións gramaticais e corresponden a un obxecto ou a un predicado determinado (ser humano, animal, vexetal, estado, acción, acontecemento, lugar, unidade temporal...). Normalmente, os significados léxicos non se organizan en sistemas de oposición xa que, aínda que entran a formar parte de redes hiperonímicas, sinonímicas, antonímicas, etc., raramente se exclúen uns ós outros. Organízanse, logo, en clases, non en categorías. Dentro dos significados léxicos cabe establecer unha división fundamental entre argumentos elementais (ou obxectos) e predicados, aínda que non nos deteremos aquí nesta cuestión.

As significacións gramaticais, en cambio, forman un conxunto pechado ou semipechado. Os procesos de gramaticalización son moito máis lentos e progresivos cós procesos de lexicoxénese. De feito, un cambio profundo nas significacións gramaticais dunha lingua dada pode mesmo orixina-la aparición dunha nova lingua (cousa que nunca ocorrerá por moitos significados léxicos que introduzamos en ou eliminemos dunha determinada lingua). Ademais, as significacións gramaticais non están tan directamente vinculadas á realidade extralingüística como os significados léxicos. Se ben non é certo en tódolos casos, de maneira global pode afirmarse que as significacións gramaticais son sempre predicados e, moi a miúdo, predicados de segunda orde, é dicir, predicados que se aplican sobre outros predicados. Salientemos, finalmente, que as significacións gramaticais se organizan en categorías, isto é, en sistemas cuxos valores se exclúen uns ós outros (non é posible, por exemplo, ser á vez 'presente' e 'futuro'; en calquera caso deberemos postular para unha lingua dada a non distinción gramatical entre presente e futuro, que se oporía, eventualmente, ó pasado⁹).

Resulta fundamental facer notar que unha lingua dada se caracteriza, sobre todo, polo seu sistema específico de significacións gramaticais e polos medios formais que

⁹ É o que ocorre, por exemplo, no xaponés, onde atopamos un pasado que se opón ó non-pasado.

emprega para expresalas. É o que poderíamos chamar o “núcleo duro” dunha lingua. Dentro deste núcleo existe amais un subconxunto privilexiado que corresponde ás significacións gramaticais de tipo flexivo: os flexemas. Un flexema é unha significación que pertence a unha categoría flexiva dunha lingua dada (por exemplo, o plural en español ou o perfectivo en ruso). Para podermos falar dunha determinada categoría flexiva (aplicada a un tipo de signos X) no seo dunha lingua dada é preciso que as significacións desa categoría ($'Si_1'$, $'Si_2'$... $'Si_n'$) cumpran, simultaneamente, cinco condicións: primeira, respecto a todo signo de tipo X, unha e só unha das significacións Si debe expresarse obrigatoriamente; segunda, toda Si_n exprésase polo menos respecto a un signo X; terceira, o resultado da unión de X con Si é calculable de maneira estritamente composicional; cuarta, o número de signos X é moito maior có de significacións Si; quinta e derradeira, case tódalas significacións Si poden expresarse respecto a case tódolos signos X.

O outro tipo de significacións gramaticais son as derivativas. Os derivatemas son moi semellantes ós flexemas canto ós medios de expresión formal (afixos, operacións morfolóxicas, mesmo ás veces supramorfos, é dicir, significantes morfolóxicos de carácter prosódico), pero a súa expresión é moito menos regular (en relación coas condicións arriba enumeradas) e, ademais, non é obrigatoria. Unha diferenza importantísima radica no feito de que as distintas combinacións dun signo X e os flexemas que se lle aplican forman un único lexema (por exemplo, *amarillo, amarilla, amarillos, amarillas*) mentres que as distintas combinacións dun signo X cos derivatemas que se lle aplican dan lugar a distintos lexemas (por exemplo, *amarillo, amarillento, amarillecer*...).

Por outra parte, os derivatemas non teñen por que formar sistemas de oposicións e poden presentar significacións moito máis vinculadas á realidade extralingüística cós gramemas, por exemplo, o derivatema ‘sexo feminino’¹⁰ en francés (*acteur, actrice*) ou ‘golpe de’ en castelán (*mazazo, garrotazo*).

6. Sentidos colocacionais e significacións gramaticais: ¿máis ca unha simple coincidencia?

Unha vez feitas as excepcións precedentes, podemos xa presenta-la hipótese central do noso artigo:

Dende o punto de vista da lingüística xeral, o conxunto de sentidos vehiculables por medio de colocacións léxicas, por unha banda, e o conxunto das significacións correspondentes a flexemas e derivatemas, por outra, presentan unha importante área de intersección semántica.

Dito doutro xeito, os sentidos das colocacións (tal como foron definidas anteriormente) non son arbitrarios, senón que se corresponden de maneira bastante estrita cos sentidos adaptados polas linguas naturais como flexemas e derivatemas.

¹⁰ Acerca da consideración do feminino como un derivatema en francés no canto de como un flexema, como é normalmente tratado, véxase Mel'čuk (2000).

Engadamos que os mencionados valores (sobre todo os correspondentes a flexemas) son sistematizables e, ata certo punto, calculables. Polo tanto, pódese contar cun inventario presumiblemente próximo á exhaustividade.

Trátase dunha hipótese forte que, ata onde temos coñecemento, non foi formulada anteriormente dunha maneira explícita e expresa. Cremos que é certa (*cum grano salis*) e esforzarémonos en demostralo neste artigo. Nela descansa, en certo modo, o eixo central do proxecto de investigación coordinado (Universitat Autònoma de Barcelona e Universidade da Coruña) I+D+i HUM 2005-08052-C02-01 InterGram *Interlingua Gramatical para Sistemas de Procesamiento Automático del Lenguaje Natural: Significaciones Gramaticales y Colocacionales*. Desta hipótese despréndense tamén liñas de investigación particularmente aptas para aplicacións en tradución automática.

Subliñemos que incluímos na nosa hipótese a mención á lingüística xeral porque nos estamos referindo non a unha lingua ou familia de linguas concretas senón a calquera lingua natural. Obviamente, a nosa hipótese resultaría trivialmente falsa de referírmonos a un único sistema lingüístico, pois o número de flexemas e derivatemas de calquera lingua cobre só unha pequena parte¹¹ dos sentidos das colocacións léxicas.

É necesario facer tamén fincapé en que a hipótese expresada sería igualmente falsa se o concepto de colocación utilizado incluíse os casos descartados no apartado 2. De aí que resulte crucial definir ben o concepto de colocación, do contrario é imposible formularmos ningunha afirmación ó respecto. Pode resultar paradoxal mais entre os lingüistas profesionais abundan os malentendidos por mor do uso laxo dos termos da nosa propia disciplina.

Outra observación importante radica na distinción entre as colocacións léxicas e as colocacións que cumpren tódalas condicións especificadas no apartado 2 mais nun nivel inferior ó do lexema¹². Por exemplo, en alemán son frecuentes as colocacións de tipo intensivo como *kristallklar* ou *blitzsauber* que forman un composto adxectivo. A análise destas formas de cara á tradución debe se-la mesma cá das colocacións pluriléxicas: **Magn(klar)**, **Magn(sauber)**. Tamén debemos aplica-la mesma análise, aínda que debido a outro fenómeno (a suplección), en casos como o francés *navet* ('película moi mala'), no que unha forma totalmente idiomática constitúe o valor da colocación **AntiBon(film)**. En ámbolos casos, a tradución menos forzada ó castelán é plurilexemática: *más claro que el agua, limpio como una patena e película malísima*.

A continuación, non nos queda máis que presentar brevemente o sistema de cálculo das significacións gramaticais e, de seguida, poñer en relación un amplo subconxunto de significacións flexivas e derivativas cos sentidos colocacionais, que representaremos mediante o formalismo das funcións léxicas.

¹¹ Aínda que non ten por que ser un subconxunto estrito, xa que podemos atopar facilmente linguas que posúen derivatemas que non constitúen colocacións léxicas, p. ex. *-burger* é actualmente un derivatema do inglés (*cheeseburger, fishburger...*) que non corresponde a ningunha colocación formalizable mediante funcións léxicas estándar.

¹² Tampouco debemos esquecer que unha significación gramatical pode expresarse por medios puramente sintácticos, p. ex. o iterativo mediante a estrutura *V e V'* (*golpeó y golpeó*) (~ 'golpeou repetidamente').

7. Presentación comparada de significacións gramaticais (flexivas e derivativas) e sentidos colocacionais

Ben, ¿cales son as significacións gramaticais (flexivas e/ou derivativas) e os sentidos colocacionais máis comúns?

Boa parte das significacións gramaticais poden preverse, de maneira hipotético-deductiva, a partir da combinatoria das familias de categorías gramaticais semánticas (coordenadas espaciais, coordenadas temporais, coordenadas cuantitativas e coordenadas cualitativas) e os elementos básicos dunha situación de comunicación, é dicir, a enunciación, os actantes (ou participantes) da enunciación, o enunciado e os actantes (ou participantes) do enunciado. Así pois, o tempo absoluto (*tense*, *Tempus*) sería, *grosso modo*, a relación entre o tempo do enunciado e o tempo da enunciación; a voz sería a relación entre o enunciado e os seus actantes, sen referencia á enunciación nin ós seus participantes; o modo sería a relación entre o enunciado e os seus actantes con referencia á enunciación e ós seus participantes, etc.

Estes elementos de base e as súas relacións son coñecidísimos e, polo tanto, renunciaremos a presentalos aquí en detalle. Limitarémonos a dar tres referencias, que poderían considerarse xa clásicas, de utilización destas coordenadas (Jakobson 1971, Pottier 1974 e Mel'čuk 1993-2000¹³).

Para ilustrar, se non probar, a nosa hipótese, non nos queda máis que aliña-las significacións flexivas e derivacionais cos sentidos colocacionais para mostrar ata que punto poden apreciarse coincidencias. Representarémolas significacións gramaticais mediante glosas semánticas do tipo 'x' e os sentidos colocacionais mediante o formalismo das funcións léxicas (véxase, entre outros, Mel'čuk 1993 e Wanner 1996¹⁴). É importante que indiquemos que se trata unicamente dunha primeira aproximación á cuestión. Por conseguinte, un número importante de casiñas están aínda en branco. As correspondencias definitivas esixirán observacións moito máis detalladas¹⁵. Subliñemos, tamén, que as funcións léxicas non son unidades semánticas (senón nodos da representación sintáctica profunda, unidades léxicas "profundas", por dicilo así) e que, polo tanto, non se trata de establecer aquí unha ecuación entre significación gramatical e función léxica, senón unicamente de suxerir unha relación entre medio de expresión gramatical e medio de expresión por coocorrencia léxica restrinxida.

¹³ Este último texto, particularmente o volume II (1994), sérvenos de punto de referencia básico para a columna esquerda das seguintes táboas.

¹⁴ Particularmente, Alonso e Tutin (1996) para unha primeira clasificación completa das funcións léxicas.

¹⁵ No marco do proxecto InterGram, estas observacións levaranse a cabo a partir dun cronograma de traballo que se estende ata mediados de 2008.

Dado que non nos referimos a ningunha lingua en particular, a seguinte táboa deberá quedar necesariamente subespecificada (adoptarémolos sistemas de significacións máis comúns para cada categoría). Para a maior parte de significacións presentadas deberían, non obstante, poderse atopar bos exemplos en linguas románicas, xermánicas e eslavas (basicamente, nas linguas denominadas SAE: *Standard Average European*), sen necesidade de recorrer a linguas tipoloxicamente moi afastadas.

Sinalamos, como nota importante, que excluímos da presentación as categorías gramaticais de tipo *shifter*, dado que as funcións léxicas non formalizan ningunha destas categorías. Non incluiremos, pois, nas táboas ofrecidas a continuación a deixé, a orientación *shifter*, o tempo absoluto nin a distancia temporal.

Comezaremos coas categorías vinculadas á expresión do espazo. Son as seguintes (ver Táboa 1):

Localización (non <i>shifter</i>)	‘dentro de’	Loc_{in}
	‘debaixo de’	Loc_{sub}
	‘detrás de’	Loc_{post}
	‘ó carón de’	Loc_{apud}
	‘entre’ / ‘no medio de’	Loc_{inter}
	‘encol de’	Loc_{super}
	‘diante de’	Loc_{ante}
	‘adxacente a’	Loc_{ad}
	‘arredor de’	Loc_{circum}
Orientación non <i>shifter</i>	‘en repouso’	-
	‘a X’	-
	‘cara a X’	-
	‘ata X’	-
	‘a través de X’	-
	‘de X’	Loc_{ab}
	‘a partir de X’	-

Táboa 1

As categorías vinculadas á expresión do tempo son (ver Táboa 2):

Tempo relativo	‘simultáneo’	-
	‘anterior’	-
	‘posterior’	-
Resultatividade	‘resultativo’	Result_i
	‘non resultativo’	-

Táboa 2

As categorías vinculadas á expresión da cuantificación son (ver Táboa 3). Canto ó aspecto e á súa relación coas funcións léxicas, véxase Nakhimovsky (1996).

Número (nominal)	‘singular’	-
	‘dual’	-
	‘plural’	-

Significacións gramaticais e sentidos colocacionais: ¿máis ca unha simple coincidencia?

Aspecto numérico	'multiplicativo'	-
	'semelfactivo'	-
Mesuratividade	'aumentativo'	Magn / Augm
	'diminutivo'	AntiMagn / Dim
Distributividade	'concentrativo'	-
	'distributivo'	-
Colectividade	'singulativo'	Sing
	'colectivo'	Mult
Intensidade	'intensivo'	Magn, Magn^{tiempo} ...
	'atenuativo'	AntiMagn...
Aspecto de distribución	'concentrado'	-
	'distributivo'	-
	'iterativo'	-
	'iterativo-distributivo'	-
Aspecto de duración	'puntual'	puntualmente+ FL
	'durativo'	continuo+ FL
	'habitual'	habitualmente+ FL
Aspecto de desenvolvemento	'progresivo'	-
	'non progresivo'	-
Aspecto de completión	'perfectivo'	Perf
	'imperfectivo'	-

Táboa 3

A continuación témo-las categorías referidas ás relacións entre a enunciación, os participantes da enunciación, o enunciado e os actantes do enunciado. Son bastante máis numerosas cás referidas ás coordenadas espaciais, temporais e cuantitativas. Por conseguinte, non as enumeraremos *in extenso*, senón que reteremos unicamente aquelas que xulgamos mellor representadas en linguas do tipo SAE (ver Táboa 4).

Fase	'incoativo'	Incep
	'continuativo'	Cont
	'cesativo'	Fin
Modo	'indicativo'	-
	'imperativo'	Imper
	'optativo'	-
	'subxuntivo'	-
	'condicional'	-
	'irreal'	-
Voz	PASIVOS, PERMUTATIVOS, SUPRESIVOS, REFLEXIVOS	Oper₂, Oper₃, Func₀, Func₁, Func₂, Labor₁₂, Labor₂₃, Conv_{ijk...}¹⁶
Avaliatividade	'hipocorístico'	Bon
	'pexorativo'	AntiBon
Modalidade	'desiderativo'	-
	'potencial'	Able₁, Able₂
	'debitativo'	-
Reactividade	'asertivo'	-
	'admirativo'	-
	'probabilitivo'	Qual_i
	'dubitativo'	-

¹⁶ Respecto á relación entre verbos soporte e significacións gramaticais, véxase Blanco e Buvet (2004).

Evidencialidade	‘experencial’	-
	‘citativo’	-
	‘inferencial’	-
	‘presuntivo’	-
Respectividade	‘respectivo’	Pos₁, Pos₂
	‘depreciativo’	AntiPos₁, AntiPos₂
Pertenza/Posesividade		-

Táboa 4

Finalmente, presentaremos algunhas significacións gramaticais de tipo derivacional (ver Táboa 5):

Factitivo	‘causativo’	Caus
	‘permisivo’	Perm
	‘decausativo’	Liqu
Aplicativo	‘benefactivo’	Involv
	‘malefactivo’	Involv
	‘comitativo’	Involv
Modos de acción	‘evolutivo’	PredPlus, PredMinus
	‘delimitativo’	-
	‘perdurativo’	-
	‘exhaustivo’	Stop
	‘completivo’	-
	‘nimifactivo’	Excess
	‘saturativo’	-
	‘negativo-resultativo’	Nocer
	‘cumulativo’	Magn^{cant}
Recíproco	‘R(X,Y) e R(Y,X)’	~ Conv_{ijk}
Opositivo	‘contrario de...’	Anti
Simulativo	‘finxir que R’	~ AntiVer
Transmutativo	‘volverse R ou máis R’	PredPlus, PredMinus
Posibilitivo	‘que pode V doadamente’	Able_i
Identificativo	‘ser un X’	Copul
Habitivo	‘ter X’	-
Produtivo	‘causar que X exista’	-
Usitativo	‘someter ó efecto de X’	Fact_i, Real_i, LabReal_{ij}
Nome de axente	‘o que V’	S₁
Nome de paciente	‘ó que V’	S₂
Nome de lugar	‘onde V’	S_{loc}
Nome de instrumento	‘co que V’	S_{instr}, S_{med}
Nome de maneira	‘como V’	S_{mod}
Nome de resultado	‘como resultado de V’	S_{res}

Táboa 5

Chamémo-la atención sobre o feito de que, aínda que os derivatemases sexan máis difíciles de sistematizar cós flexemas, podemos introducir un principio de clasificación reutilizando as nocións de predicado e actante. Teremos, así, catro tipos de derivatemases segundo nos permitan obter designacións para predicados a partir de designacións doutros predicados, designacións para predicados a partir de designacións de actantes,

designaci3ns de actantes a partir de designaci3ns de predicados e, por 3ltimo, designaci3ns de actantes a partir de designaci3ns doutros actantes.

Unha categor3a *sui generis* no tocante 3 tema do noso artigo ser3a a determinaci3n, porque non corresponde a un valor colocativo canto 3 oposici3n ('definido', 'indefinido', 'partitivo'), mais os artigos deben ser descritos no interior das entradas das funci3ns l3xicas. Deste xeito, por exemplo, se consideramos *tener* (fronte a outras posibilidades como *padeecer*, *sufrir de*, etc.) como valor de **Oper**₁ para as bases *resfriado*, *colera* e *asma*, ter3amos, respectivamente, **Oper**₁(*resfriado*) = *tener* [UN], **Oper**₁(*c3lera*) = *tener* [EL], **Oper**₁(*asma*) = *tener* [Ø].

Antes de darmos por concluido este apartado, deber3amos facer a3nda d3as observaci3ns importantes. En primeiro lugar, antes refer3monos exclusivamente a flexemas e derivatemas sem3nticos. Agora ben, c3mpre referirse tam3n 3s flexemas e derivatemas sint3cticos. Entre estes, os m3is importantes para o tema que nos ocupa son, sen d3bida, as significaci3ns derivacionais sint3cticas, que se dividen en: nominalizadores, p. ex. **S**₀(*admirar*) = *admiraci3n*; verbalizadores, p. ex. **V**₀(*paseo*) = *pasear*; adxectivalizadores, p. ex. **A**₀(*ciencia*) = *cient3fico*, **A**₁(*ira*) = *iracundo*, **A**₂(*odio*) = *odiado* e adverbializadores, p. ex. **Adv**₀(*felicidad*) = *felizmente*, **Adv**₁(*pesar*) = *con pesar*, **Adv**₂(*satisfacci3n*) = *a satisfacci3n de*. T3ñase en conta que, en caso necesario, poden empregarse os super3ndices **FL**^{actual} e **FL**^{usual} para introducir distinc3ns relativas 3 aspecto intr3nseco das unidades l3xicas que constit3en os valores de determinadas funci3ns l3xicas, p. ex. **S**₁^{usual}(*miedo*) = *miedoso* fronte a **S**₁^{actual}(*miedo*) = *asustado*.

En segundo lugar, c3mpre destacar que, nas t3boas presentadas, partimos das significaci3ns flexivas e derivacionais para po3nelas en relaci3n coas funci3ns l3xicas correspondentes. Non obstante, poder3a facerse xusto o contrario; isto 3, partir das funci3ns l3xicas inventariadas para po3nelas en relaci3n coas significaci3ns gramaticais correspondentes. As3, do mesmo xeito que podemos constata-la existencia de significaci3ns aparentemente non relacionadas con funci3ns l3xicas, podemos detectar funci3ns l3xicas sen v3nculo aparente con significaci3ns flexivas e/ou derivacionais. Cremos oportuno, xa que logo, ofrecer aqu3 a lista de funci3ns l3xicas para as que non detectamos polo de agora ningunha significaci3n paralela. Tr3tase de:

- **Cap** ('xefe de...'), p. ex. *el capit3n del barco*.
- **Centr** ('o centro de...', 'o medio de...'), p. ex. *el meollo de la cuesti3n*.
- **Contr** (termo contrastivo), p. ex. *coraz3n respecto de cabeza*.
- **Culm** ('punto culminante de...'), p. ex. *el cl3max de la pel3cula*.
- **Degrad** ('estragarse'), p. ex. *la leche se ha cortado*.
- **Epit** (ep3teto est3ndar desprovisto de sentido), p. ex. *el proceloso oc3ano, el crudo invierno, el cielo azul, el afortunado ganador*.
- **Equip** ('equipo de...'), p. ex. *la tripulaci3n de la aeronave*.
- **Excess** ('funcionar excesivamente'), p. ex. *el caballo se desboc3*.
- **Figur** (met3fora estereotipada), p. ex. *una cortina de humo*.
- **Gener** (termo xen3rico), p. ex. *veh3culo respecto de motocicleta*.
- **Germ** ('o xerme de...'), p. ex. *la chispa de la vida*.

- **Manif** ('manifestarse en...'), p. ex. *una leve sonrisa se dibujó en su rostro*.
- **Obstr** ('funcionar con dificultade'), p. ex. *se le nubla la vista*.
- **Propt** (preposición co significado 'a causa de, debido a...'), p. ex. *bajo los efectos de estupefacientes*.
- **Prox** ('estar a piques de...'), p. ex. *al borde de un ataque de nervios*.
- **Son** ('emiti-lo son característico'), p. ex. *el elefante barrita*.
- **Sympt_{ijk}** ('ser un síntoma físico de'), función moi útil para formalizar unha gran variedade de expresións sobre sentimentos, combinada sempre con **Obstr**, **Excess** ou **Stop**, p. ex. **Excess**(pelos) - **Sympt₁₃**(miedo) = *se le pusieron los pelos de punta [de miedo]*.
- **Syn** (sinónimos e cuasisinónimos da base).

Cabería aínda considera-lo conxunto das funcións léxicas complexas e as configuracións de funcións léxicas. Tampouco hai que esquecer-la posibilidade de marcar relacións semánticas aproximadas mediante os subíndices \subset ($X \subset Y$ implica que o significado de X está incluído en Y e, polo tanto, que X presenta unha maior extensión semántica ca Y), \supset ($X \supset Y$ implica que o significado X inclúe o significado Y e, polo tanto, que X presenta unha maior intensión que Y) e \sqcap ($X \sqcap Y$ implica que existe unha intersección entre os significados de X e Y). Ademais, temos superíndices como os empregados máis arriba: **FL^{cant}**, **FL^{tempo}** (p. ex. con **Magn**), **FL^{actual}**, **FL^{usual}** (p. ex. con **S₁**) e outros como **FL^{cor}**, **FL^{dim}** (dimensión), **FL^{motor}** (movemento), **FL^t** (temperatura), (p. ex. con **Degrad** ou **Excess**), etc.

No marco do proxecto InterGram, trataremos de poñer en relación as funcións léxicas antes enumeradas con derivatemais de distintas linguas naturais. Polo de agora, non podemos ofrecer máis ca este breve esbozo.

8. Conclusións

Resulta tentador preguntarse por que xustamente o conxunto de sentidos ó que nos referimos presenta este funcionamento tan particular. Trátase, sen dúbida, de sentidos moi "importantes". A verdade é que, dende a nosa especialidade (a lingüística aplicada) o "porqué" parécenos escuro e non dubidamos de que outros especialistas (probablemente os expertos en diacronía ou en psicolingüística) estarán en mellores condicións ca nós para tentar responder a esta pregunta. Unha vez establecidos un "que" (e, na medida do posible, un "como" para cada lingua particular), a nosa orientación profesional e científica lévanos inmediatamente a preguntarnos o "para que". As consecuencias, máis cáas causas, lévanos ás aplicacións.

Dende o noso punto de vista, o "para que" consiste na delimitación precisa dos sentidos que deben tratarse por interlingua (tódolos sentidos gramaticais e colocacionais¹⁷)

¹⁷ Agás algúns sentidos demasiado específicos para seren incluídos na interlingua e que, previsiblemente, aparecerán só en funcións léxicas non estándar dalgunhas linguas en particular e nalgúns derivatemais tamén moi específicos de determinadas linguas, especialmente aqueles que forman designacións de participantes a partir doutras designacións de participantes. Así, por exemplo, o significado 'restaurante no que se serve a comida X' (Mel'čuk 1993-2000, vol. II: 381), que constitúe un derivatema en ruso, non

fronte a aqueles que deben tratarse por transferencia. Cumpriríámolo obxectivo deste artigo se puidemos suxerir a existencia dun contínuum semántico entre significación gramatical e sentido colocacional e mostrar, así, que se se desexa obter unha tradución automática de calidade obrigatoriamente haberá que dotar ós sistemas da capacidade de “traducir” un gramema por unha ou varias unidades léxicas e viceversa. Dado que é difícil imaxinar un módulo bilingüe capaz de manexar sen erros nin falsas xeneralizacións equivalencias do tipo *re-V = V de nuevo*, o paso por unha representación abstracta (mediante o formalismo das funcións léxicas ampliadas, se é preciso¹⁸, a tódalas significacións gramaticais) revélase indispensable.

Non obstante, conservamos unha tradución puramente por transferencia para os significados léxicos. Logo, postularemos un sistema mixto transferencia-interlingua para a arquitectura do sistema InterGram (Blanco, en preparación), dentro do sistema de dicionarios electrónicos de *fLexSem*. Esta é a principal novidade do noso enfoque xa que o papel fundamental das significacións gramaticais nunha interlingua xa foi posto de relevo por Igor Mel’čuk dunha maneira diáfana en abril de... ¡1959! (v. Mel’čuk 1976), e sen ter acceso a Jakobson (1963). O artigo que dá conta daquela comunicación foi publicado en 1976 (v. “Bibliografía”) e, ó noso modesto entender, recibiu moita menos atención da que merecía. En calquera caso, estamos convencidos de que o mencionado texto segue de actualidade, especialmente no tocante á concepción da gramática como *ars obligatoria*.

Obviamente, debido á presentación panorámica que adoptamos, unha enorme cantidade de cuestións, algunhas de grande importancia, quedaron sen tratar nestas poucas páxinas. Por non citar máis ca unha, non dixemos nada en relación coa diferenza fundamental entre causación directa e causación indirecta, por conseguinte, é posible que o ‘causativo’ gramatical dun verbo dado sexa directo e o **Caus** aplicado ó **S₀** dese verbo sexa indirecto, o que implica diferenzas importantes, como a non presuposición da unidade de tempo e de espazo no último caso (v. Wierzbicka 1988). E aínda quedaría por considera-lo sema ‘causar’ que está presente no significado de moitos verbos transitivos. É evidente que o noso estudo foi extremadamente superficial e aproximativo, mais non tiñamos moitas alternativas tendo en conta a grande extensión do tema abordado. Así pois, esperamos ter ocasión de precisar moitos destes puntos nun futuro próximo.

9. Bibliografía

quedarán incluído na interlingua senón que se tratarán tódalas unidades léxicas formadas deste xeito na macroestrutura do dicionario electrónico, sen descomposición de ningún tipo (*blinnaja, pirožkovaja, sosisočnaja*, etc.) e traduciranse por transferencia.

18 Mais só se é realmente preciso, xa que o dito formalismo ofrece un conxunto de posibilidades que van moi alén das funcións estándar e que inclúe, como xa sinalamos, a utilización de funcións complexas, configuracións de funcións e funcións mixtas con glosas abertas. Con todo, queda a posibilidade de introducir novas funcións léxicas estándar. Así, por exemplo, o castelán ten un derivatema para ‘femia de X-animal’ (*perra, gata, burra, cerda, leona, tigresa...*) e, asemade, valores supletivos que poden formalizarse mediante funcións léxicas: **Femia**(caballo) = *yegua*, **Femia**(toro) = *vaca*, **Macho**(tortuga) = *tortuga macho*, etc.

- ALONSO, M. e TUTIN, A. (1996): "A Classification and Description of the Lexical Functions of the Explanatory Combinatorial Dictionary for the Treatment of LFs Combinations" en Wanner, L. (ed.): *Lexical Functions in Natural Language Processing and Lexicography*. John Benjamins, Amsterdam/Philadelphia.
- BLANCO, X. (2001): "Dictionnaires électroniques et traduction automatique espagnol-français" en *Langages* 143,2001.
- BLANCO, X. e BUVET, P.-A. (2004): "Verbes support et significations grammaticales. Implications pour la traduction espagnol-français" en *Linguisticae Investigationes* 27,2004,2. John Benjamins, Amsterdam/Philadelphia.
- BLANCO, X. (dir.) (en preparación): *InterGram. Una interlingua gramatical para sistemas de procesamiento automático del lenguaje natural*.
- DANLOS, L. (1994): "Coder des informations monolingues sur les noms pour éviter des règles bilingues sensibles au contexte" en *Langages* 116,1994.
- MEL'ČUK, I. (1976): "Grammatical Meanings in Interlinguas for Automatic Translation and the Concept of Grammatical Meanings" en *Das Wort*. Wilhelm Fink Verlag, München.
- (1993): "Fonctions lexicales dans le traitement du langage naturel" en Clas, A. e Bouillon, P. (eds): *TA-TAO: Recherches de pointe et applications immédiates*. Aupelf-Uref, Montreal.
- (1993-2000): *Cours de Morphologie Générale*, 5 volumes. Les Presses de l'Université de Montréal / CNRS Éditions, Montreal.
- (1995): "Phrasemes in Language and Phraseology in Linguistics" en Everaert, M. e van der Linden, E.-J. (eds): *Idioms: Structural and Psychological Perspectives*. Erlbaum, Hillsdale.
- (2000): "Un fou/une folle : un lexème ou deux ?" en *Bulag* numéro hors série. Les Presses de l'Université de Franche-Comté, Besançon.
- NAKHIMOVSKY, A. (1996): "A Case of Aspectual Polysemy, with Implications for Lexical Functions" en Wanner, L. (ed.): *Lexical Functions in Natural Language Processing and Lexicography*. John Benjamins, Amsterdam/Philadelphia.
- JAKOBSON, R. (1963): "La notion de signification grammaticale selon Boas" en *Essais de linguistique générale*. Les Éditions de Minuit, Paris.
- (1971): "Shifters, Verbal Categories, and the Russian Verb" en *Selected Writings II, Word and Language*. Mouton, The Hague/Paris.
- POTTIER, B. (1974): *Linguistique générale: théorie et description*. Klincksieck, Paris.
- WANNER, L. (ed.) (1996): *Lexical Functions in Natural Language Processing and Lexicography*. John Benjamins, Amsterdam/Philadelphia.
- WIERZBICKA, A. (1988): *The Semantics of Grammar*. John Benjamins, Amsterdam/Philadelphia.