

PRESENTACIÓ DEL PROJECTE *APLICACIÓ MULTIMÈDIA PER AL TREBALL AMB UNITATS FRASEOLÒGIQUES VERBALS ESPANYOL, FRANCÉS, ITALIÀ, RUS, POLONÉS, XINÉS I CATALÀ*¹

PEDRO MOGORRÓN HUERTA (pedro.mogorron@ua.es)
M. ISABEL GUARDIOLA I SAVALL (maribel.guardiola@ua.es)
SANDRA MONTSERRAT I BUENDIA (sandra.montserrat@ua.es)
MIGUEL TOLOSA IGUALADA (miguel.tolosa@ua.es)
Universitat d'Alacant. Frasitram

1. INTRODUCCIÓ

L'estudi de les unitats fraseològiques presenta un gran interès en les investigacions desenvolupades per les tendències lingüístiques i traductològiques actuals. Resultat d'aqueixes investigacions és el fet que aquests elements lingüístics es consideren sistematitzables i ja no un calaix de sastre. Així i tot, des del punt de vista traductològic continuen percebent-se com a elements complexos i difícils de traduir; de fet, no hem d'oblidar que es consideren els elements més "genuïns" de cada llengua i de sobra és conegut que qui domina la fraseologia d'una llengua està més prop de conèixer de manera més aprofundida el sistema d'aqueixa llengua.

Aquesta preocupació recent per aquest tipus d'unitats explica l'aparició de diversos productes que tracten d'establir els fraseologismes d'una llengua. No obstant això, aquesta producció és moltes vegades insuficient i insatisfactòria per a cobrir les necessitats dels diferents tipus d'usuaris que hi accedeixen. El projecte que presentem en aquest article naix de l'observació d'algunes dificultats de la tasca diària del traductor –fonamentalment– en relació amb aquests elements i, de retop, de la curiositat manifesta en els lingüistes.

2. EL PROJECTE *APLICACIÓ MULTIMÈDIA PER AL TREBALL AMB UNITATS FRASEOLÒGIQUES VERBALS*

2.1. Objectius

A l'hora d'enfrontar-se a la tasca de trobar equivalents entre les estructures fraseològiques en llengües diverses, traductors, lingüistes i, fins i tot, nadius es troben amb obstacles varis. D'una banda, es pot constatar que és gairebé impossible trobar tractament d'unitats fraseològiques en llengües com ara el romanés, l'àrab, el xinés, etc.². De l'altra, hi ha una mancança evident de diccionaris fraseològics que tracten més d'una llengua i, per tant, el traductor ha de trobar aquests elements en diccionaris bilingües. A l'últim, ateses les característiques de la faena del traductor –dinàmica– i les dels diccionaris –estàtics– que té a l'abast, la consulta dels diccionaris fraseològics o bilingües s'esdevé poc útil per diversos motius: *a)* l'ordenació de la informació és alfabètica i/o es basa en alguna part de la unitat fraseològica, sovint no clara o bé, no especificada en les instruccions del diccionari; *b)* la consulta en paper és poc ràpida; *c)* la informació és escassa, és a dir, hi ha mancança d'exemples o de contextualitzacions d'ús, no solen haver-hi definicions o anotacions enciclopèdiques i hi ha poques informacions d'ús referides a la freqüència, la varietat dialectal o els nivells de llengua. Precisament, en aquestes mancances se situa l'origen del projecte que presentem en aquest article.

El projecte que presentem naix amb l'objectiu de facilitar la tasca del traductor i també del lingüista en la consulta de les unitats fraseològiques. Per aconseguir-ho es treballa sobre el disseny d'una aplicació virtual dinàmica de tractament d'aquestes unitats, una espècie de diccionari electrònic, que: *a)* permetrà als traductors trobar el significat i les variants parasonímiques de les expressions fixes en cadascuna de les llengües que aquesta eina ens ofereix i que sovint no figuren en els diccionaris; *b)* oferirà tota la informació accessible des d'un web, amb accés en funció dels diferents perfils: investigador, traductor, docent, etc., cosa que marcarà el tipus d'informació que s'hi podrà trobar. Així, per exemple, el lingüista

¹ Aquest article s'inscriu dins del projecte d'investigació "Estudio de locuciones verbales y elaboración de diccionarios electrónicos plurilingües" (GV05/163), subvencionat per la Conselleria d'Empresa, Universitat i Ciència de la Generalitat Valenciana.

² Pel que fa a les llengües eslaves, com ara el rus, volem indicar que, tot i la llarga i reconeguda tradició lexicogràfica en relació amb la fraseologia, tanmateix, no hi ha obres de recopilació d'unitats fraseològiques més enllà dels anys 80.

podrà accedir a informació sintàctica complexa que, en canvi, podria arribar a ser un destorb per al professional de la traducció.

2.2. Fases

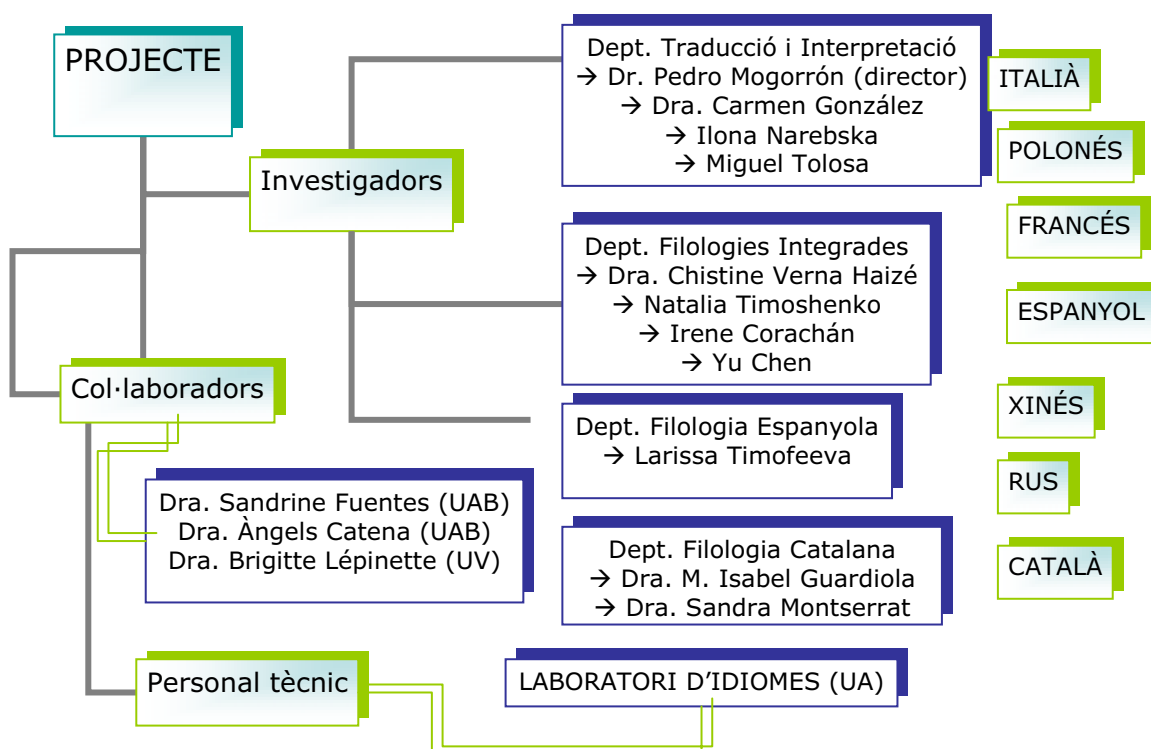
El projecte que presentem en aquest article és molt jove: tot just va començar el 2005 quan es va aconseguir el finançament del programa d'investigació dels "Grups emergents" de la Universitat d'Alacant. En aquesta primera fase només es van traçar els objectius bàsics i es va decidir treballar amb unitats fraseològiques de base verbal. La tasca es va centrar en l'espanyol i el francès, amb la qual cosa es va assolir un nombre important de fraseologismes sobre els quals poder començar a treballar i a crear una base inicial de tractament.

Tot just som en la segona fase de treball: el 2006. En aquesta etapa, s'ha aconseguit un finançament major, amb la concessió d'una ajuda per part de la Conselleria d'Empresa, Universitat i Ciència de la Generalitat Valenciana. Aquesta nova mampresa ha implicat la incorporació de més idiomes –i, per tant, de més investigadors– a la faena. Concretament, s'hi han incorporat el català, l'italià i el rus i s'ha assajat la base de tractament a un camp semàntic concret, el de la <comunicació>, per a totes aquestes llengües.

A les acaballes del 2006, on som ara justament, s'han incorporat altres llengües, el polonès i el xinès, i s'han encetat els camps semàntics de <descripció física> i <descripció psíquica (caràcter)>. Actualment, hi ha el nombre d'entrades següents distribuïdes per llengües: espanyol (13.500), francès (4.000), català (2.500), rus (3.000) i italià (1.000). El 2007 s'incorporarà al projecte l'anglès i es preveu que, en trobar investigadors interessats, també s'hi afegiran l'alemany, l'àrab i el portuguès. A més, es realitzarà l'anàlisi sintàctica de les unitats i s'elaborarà també prompte la interfície de consulta per poder penjar-ne una demo a internet.

2.3. Equip, estructura i formació dels investigadors

Aquest projecte presenta un equip multidisciplinari de diversos departaments i de diverses universitats com podeu veure en el quadre resum que us presentem tot seguit, encara que hem d'especificar que el pes principal del projecte el porten els investigadors adscrits a la Universitat d'Alacant³.



Quadre 1. Equip i estructura del projecte

La direcció recau en un dels investigadors de l'equip francès, els membres del qual són especialistes en traducció i en fraseologia bilingüe. La investigadora que treballa actualment amb l'italià és especialista en

³ El personal investigador, col·laborador i tècnic d'aquest quadre correspon al dia de la presentació de la comunicació en el si del II Congrés Internacional de Lexicografia Hispànica.

vocabularis trilingües (espanyol-català-italià) i en marcadors discursius italià-espanyol. D'altra banda, els membres de l'equip de les llengües eslaves, rus-espanyol i polonès-espanyol, són especialistes en l'ensenyament de la fraseologia i la traducció de la fraseologia polonès i rus-espanyol, de la mateixa manera que ho és la persona que treballa amb el xinès. I, finalment, l'equip que ens dediquem als fraseologismes del català hem treballat en la teoria i pràctica lexicografia actual i antiga i també amb estructures perifràstiques. No hem d'oblidar tampoc la faena que desenvolupa el personal tècnic del laboratori d'idiomes, que s'ocupa del suport informàtic del projecte. Pel que fa als col·laboradors externs, aquests tracten d'establir els camps semàntics en què distribuïm els fraseologismes tractats.

Cal apuntar que bona part de l'equip d'investigació que tot just hem presentat viuen també de la traducció professional, cosa que ens permet obtenir de primera mà informacions sobre les necessitats reals dels traductors, no tant com a activitat teòrica simplement, sinó més aïna com a pràctica.

2.4. Metodologia

2.4.1. Metodologia: unitat fraseològica

Com bé indica i demostra Ruiz Gurillo (1997: 71) en el seu treball sobre fraseologia de l'espanyol:

Por lo que a la fraseologia se refiere, parece estar estructurada como un continuum difícil de segmentar. El comportamiento de sus miembros no es homogéneo y fluctúan en un espacio de fronteras poco precisas.

No és objectiu del projecte entrar en la discussió sobre el concepte d'unitat fraseològica, ni pretenem destriar quines unitats en formen part d'una manera prototípica o perifèrica. Així doncs, entenem, com indica Pérez Saldanya (1998), entre d'altres (cfr. Alcaraz y Martínez Linares, 1997: s.v. *frase idiomática*), una unitat fraseològica com una:

Construcció sintagmàtica d'una llengua que es caracteritza per tenir un alt grau de fixació i un significat no componencial, això és, un significat diferent del que resultaria de la suma del significat de les unitats que el conformen⁴.

Així, en el diccionari de fraseologismes que presentem hi haurà unitats que van des de la col·locació (cfr. Firth, 1957: 181; Dubois, 1979; Corpas Pastor, 1997; Solà, 1994; entre d'altres), com ara *agafar el telèfon*, a sentències més fixes del tipus, *jurar de dir la veritat, només la veritat i res més que la veritat*, passant per unitats multimot com ara *dir-ne de verdes i de madures (a algú)*, *jno me vengas con cuentos!*, *être bavard comme une pie*, *fer més soroll que un sac de nous dolentes* o *dir una mentida per estirar/traure una veritat*.

2.4.2. Metodologia: criteris d'elaboració

Pel que fa als criteris d'elaboració del projecte, els investigadors treballem:

- a) sistemàticament: a partir d'una base de dades elaborada pel Laboratori d'Idiomes de la Universitat d'Alacant;
- b) en grup: es procuren dos o més especialistes de cada llengua;
- c) conceptualment: es parteix de l'espanyol per a proposar un seguit de camps semàntics com a criteri organitzador de l'aplicació i després cada idioma realitza una investigació individual aïllada per a agrupar les unitats fraseològiques en els seus camps semàntics, a partir dels quals s'amplia el repertori. Podem assenyalar que s'han establert els camps semàntics a partir d'obres lexicogràfiques franceses, fonamentalment, i l'òptica de les classes d'objecte del LLI (Laboratoire de Linguistique Informatique, París 13 <<http://-lli.univ-paris13.fr>>), tot i que, com se sol observar generalment en treballs d'aquest tipus, hi ha una bona dosi d'introspecció. Els camps semàntics estan estructurats en un node fonamental –els anomenem “pare”– i diversos nodes secundaris –o “fills”–, que varien depenent del tipus de camp semàntic;
- d) a partir de fonts lexicogràfiques: l'investigador treballa sobre la informació present en diccionaris bilingües, monolingües i fraseològics, contrastada –i si és menester retocada– a partir del coneixement propi. D'aquestes, n'obté: la forma de les expressions, la definició i, si n'hi ha, notes d'ús i notes dialectals (p. ex.: *correr la bola* 'divulgar notícies inquietantes, propagar rumors' dialectal de Guatemala i Perú);

⁴ N'hi ha d'altres, però, que les anomenen “unitats sintagmàtiques pluriverbals”.

- e) a partir de corpus lingüístics: l'investigador treballa sobre exemples reals obtinguts a partir de les llengües actuals, si aquestes disposen de corpus electrònics⁵. En cas que no n'hi haja o bé que el fraseologisme no s'hi documente, provarem de trobar altres fonts contextuals, com ara premsa o literatura i n'especificarem l'origen.

Per a cada fraseologisme de cada llengua, els investigadors especifiquen aquesta informació:

- expressió: introduïda per la forma verbal en infinitiu o en la forma verbal corresponent si és molt fixa (ex. *no em vinguis amb romanços!*);
- definició de l'expressió: s'eviten les definicions basades en construccions;
- font de l'expressió: s'indica el diccionari o diccionaris a partir dels quals s'ha obtingut l'expressió;
- context de l'expressió: els exemples (un màxim de tres) es proporcionen atenent a la forma, l'orde més habitual i l'ús que es marca;
- informació sintàctica: es parteix del mètode Lèxico-gramàtica ("Lexique-grammaire") de Maurice Gross en el LADL (Laboratoire d'Automatique Documentaire et Linguistique: <<http://www-igm.univ-mlv.fr>> a França), utilitzat en altres diccionaris (p. ex.: *demander la parole (à quelqu'un)* [C1DPN: argument constant + det. def. + prep. + nom] o *canviar de cara* [CP1Z: argument constant + compl. prep. + 0 det. def.]). A més, també s'intenta marcar si hi ha possibilitat de variants paradigmàtiques sintàctiques, com, entre altres, la possibilitat d'introduir indefinits, possessius, etc. (p. ex.: [*abrir, destapar*] *la caja de pandora*; *abrirse [camino, paso] a codazos*; *abrir [ojos, los ojos, unos ojos] como platos*) Cal destacar que aquesta informació és més interessant per al lingüista que per al traductor i que per aquest motiu entra en una segona fase de desenvolupament del projecte;
- frequència d'ús: s'estableix si és "freqüent" o "rar". Es parteix de les indicacions en els diccionaris de referència, els corpus i la introspecció;
- camp semàntic de nivell 1: s'indica el camp semàntic principal o "pare";
- camp semàntic de nivell 2, 3, 4...: s'indiquen els camps semàntics secundaris o "fills";
- comentaris: ací es poden ampliar diversos tipus d'informació, com ara la variació de gènere (ex. *ser (algú) més dolç/a que la mel*), la sinonímia (ex. *posar-se pels núvols* i també *posar-se als núvols*), tot i que en aquest cas tindran entrada els dos fraseologismes, i també explicacions enciclopèdiques que poden aclarir alguna qüestió contextual (ex. *jurar dir la veritat, només la veritat i res més que la veritat*: en els judicis, jurament que pronuncia el testimoni abans de parlar).

2.4.3. Metodologia: interfície

A hores d'ara, l'equip informàtic del Laboratori d'Idiomes de la Universitat d'Alacant està treballant en el disseny de la interfície de consulta que volem:

- que siga ràpida i fàcil d'usar;
- que permeta la revisió constant a partir de la faena dels investigadors i els usuaris;
- que permeta la consulta en qualsevol de les llengües establides;
- que permeta la consulta per: mot clau, concepte i frase feta.

2.5. Problemàtica

El treball amb els fraseologismes amb tantes llengües i tan diverses planteja diversos problemes que provarem de resoldre a mesura que avança el projecte.

2.5.1. Problemàtica general

De manera general, podem indicar que els principals problemes comuns deriven del treball amb els camps semàntics. Així doncs, com que no són compartiments discrets, hi ha expressions que poden correspondre a un camp o a un altre (ex. *ser home de paraula* <comunicació>/<caràcter-forma de ser>) i, de la mateixa forma, una expressió polisèmica pot anar en diferents "pares" (ex. *sacar las uñas* <comunicació>/<comportament-actitud>). A més a més, també hem observat que hi ha unitats difícils d'encabir en un camp o altre (ex. *caure de la memòria* ('ser oblidat') (?) o *besar la terra que xafà* ('reverenciar') (?)).

⁵ Partim, per a les llengües romàniques, de les informacions sobre corpus electrònics de Pusch i Raible (2002).

De més a més, hem trobat dificultats a l'hora de treballar amb les expressions que indiquen gestualitat (ex. *fer cara de circumstàncies* o *dar con el codo [a algúen]*), ja que ens sembla imprescindible incorporar a la base aquest tret definitori.

Finalment, volem observar la mancança de materials en general referents als dialectes geogràfics de les llengües amb què treballem.

2.5.2. Problemàtica específica

El treball amb llengües romàniques i la traducció de l'espanyol al català, el francès i l'italià de les unitats fraseològiques planteja, a més, uns altres problemes:

- a) en els diccionaris bilingües de referència sovint no hi ha marques d'ús per a unitats que pertanyen a registres diferents (ex. *no venir al caso – no venir a tomb*);
- b) a vegades és difícil trobar correspondències en relació amb el nivell d'ús (ex. esp. *mear muy alto* - fr. *péter plus haut que son cul*);
- c) en algunes ocasions la traducció d'una unitat fraseològica no es correspon amb una unitat fraseològica (ex. esp. *hacerse un ovillo* - it. *raggomitolarsi / aggrovigliarsi*)⁶;
- d) finalment, en relació amb el català observem confusions derivades de la pèrdua de la intuïció lingüística (Solà, 1994: 287) (ex. en el *Diccionari castellà-català* de l'Enciclopèdia i el diccionari d'Espinal hi ha el fraseologisme *dir al pa pa i al vi vi*, equivalent del castellà *llamar al pan, pan y al vino, vino*. Aquesta traducció, al nostre parer, no és més que una traslació literal de l'original castellà i sembla, doncs, poc genuïna, com de fet també ho pot ratificar que no aparega tampoc en el *DCVB*. En aquest cas, la correspondència més apropiada sembla que és *dir les coses pel seu nom*, que també apareix en aquests dos diccionaris)⁷.

D'altra banda, el treball amb llengües no romàniques planteja un altre tipus de problemes. Per exemple, quant al rus, a banda de l'alfabet, els especialistes en aquesta llengua han observat que la mancança de diccionaris monolingües i fraseològics actuals és una dificultat per al seu treball (des dels anys 80 cap ací). A més, la diferència de conceptualització o de veure i classificar els esdeveniments i els sentiments implica una adaptació a les llengües romàniques que no sempre suposa una traducció exacta del sentit del fraseologisme (ex. esp. *coger al toro por los cuernos* - rus *vzjat' byka za roga*: mentre que en espanyol presenta connotacions culturals, en rus es basa en el fet que un bou és un animal perillós i agafar-lo per les banyes pot suposar un acte de valor). Finalment, hi ha la problemàtica derivada de la diversa filiació lingüística, que origina problemes, per exemple, en el fet que la traducció al rus no es correspon amb una estructura fixa (ex. en rus s'omet el verb *ser* en present, però existeix en imperfet). Aquests problemes són semblants al polonès.

Respecte al xinès, es repeteixen les qüestions observades pel que fa al rus i en part al polonès: el sistema gràfic i les diferències de concepció de la realitat, que s'accentua en no tractar-se d'una llengua indoeuropea.

3. PROJECTE DE FUTUR

En aquests fulls hem presentat un projecte que acabem tot just de començar i que es preveu que s'allargarà durant molts anys. És per això que les conclusions són més un desideràtum i una visió del projecte de futur que no un punt i final de l'estat de la qüestió.

Així doncs, la faena pròxima prevista consisteix a:

- a) perfeccionar la base de dades d'elaboració i de la interfície de consulta;
- b) completar les bases de dades de cada idioma a partir dels materials descrits i d'altres nous que apareguen al mercat;
- c) incorporar d'altres idiomes i, consegüentment, d'altres investigadors, cosa que farà augmentar el nombre dels problemes examinats i d'altres nous, en funció de la filiació lingüística de cada idioma; i, finalment,
- d) mirar de resoldre la problemàtica que afecta cada llengua.

Si aconseguim superar els obstacles descrits, la intenció final és aconseguir una eina que aprofite per a cobrir, fonamentalment, les necessitats dels traductors.

⁶ Ara bé, aquesta qüestió no acaba de ser ben bé un problema, ja que el que ens interessa realment és oferir un llistat al més exhaustiva possible en cada llengua. I és, doncs, en aquest cas que pot ser possible que no hi haja una unitat fraseològica que corresponga a una altra d'un altre idioma.

⁷ En qualsevol cas, no és objecte del nostre projecte entrar a valorar aquestes qüestions ni criticar els nostres diccionaris de referència, sinó intentar oferir una aplicació que resulte útil al traductor i a l'usuari en general. En aquest sentit, serà el propi traductor qui haurà de decidir a l'hora de triar una traducció o una altra en funció dels criteris lingüístics (nivell d'ús, freqüència, etc.) i extralingüístics (destinatari, particularitats de l'encàrrec de traducció, etc.).

REFERÈNCIES BIBLIOGRÀFIQUES

- Alcaraz Varó, E. i M. A. Martínez Linares (1997): *Diccionario de lingüística moderna*. Barcelona, Ariel.
- Corpas Pastor, G. (1997): *Manual de fraseología española*. Madrid, Gredos.
- DCVB = Alcover, A. M. i F. de Borja Moll (1928-1962): *Diccionari Català-Valencià-Balear*. 10 vols. Palma de Mallorca, Moll.
- DBIL = *Diccionari castellà-català*. Barcelona, Enciclopèdia Catalana, 1985.
- Dubois, J. (1979): *Diccionario de lingüística*. Madrid, Alianza [Traducció d'Inés Ortega i Antonio Domínguez].
- Espinal, M. T. (2004): *Diccionari de sinònims de frases fetes*. Barcelona, Universitat Autònoma de Barcelona-Servei de Publicacions de la Universitat de València-Publicacions de l'Abadia de Montserrat.
- Firth, F. R. (1957): "Modes of Meaning". En *Papers in Linguistics 1934-1951*. Londres, Oxford University Press, pàgs. 190-215.
- Mogorrón Huerta, P. (2002): *La expresividad en las locuciones verbales en francés y en español*. Alacant, Publicacions de la Universitat d'Alacant.
- Mogorrón Huerta, P. (2004): "Los diccionarios electrónicos fraseológicos (perspectivas para la lengua y la traducción)", *Estudios de Lingüística de la Universidad de Alicante*, 12.
- Mogorrón Huerta, P. (2005): "As locucions verbais nos dicionarios", *Cadernos de fraseología galega*, 7, pàgs. 149-176.
- Pérez Saldanya, M. (1998): *Diccionari de Lingüística*. Oliva, Colomar Editors.
- Pusch, C. D. i W. Raible (eds.) (2002): *Romanistische Korpuslinguistik – Korpora un gesporchene Sprache Romance Corpus Linguistics – Corpora and Spoken Language*. Tübingen, Gunter Narr.
- Ruiz Gurillo, L. (1997): *Aspectos de fraseología teórica española*, Anejo XXIV de Cuadernos de Filología. València, Universitat de València.
- Santamaría Pérez, M. I. (2003): *La fraseología española en el diccionario bilingüe español-catalán: aplicaciones y contrastes*. Alacant, Publicacions de la Universitat d'Alacant.
- Solà, J. (1994): "La lexicografía general i la normativa", *Caplletra*, 17, pàgs. 273-289.