



Grau de Lingüística

Treball de Fi de Grau

Curs 2017-2018

“CAERSE DE UN PINO” O “CAERSE DE UN GUINDO”?

Errors de la parla en expressions multiparaula

NOM DE L'ESTUDIANT: Ana Belén Laguna

NOM DEL TUTOR: Irene Castellón

Barcelona, 20 de juny de 2018

RESUM

En la parla espontània sempre hi ha presents algun tipus de refrany o de frase feta, ja que formen part de la tradició lingüística per expressar una realitat complexa. Aquestes unitats no són composicionals, ja que el significat no depèn de les unitats lèxiques, sinó que és un significat conjunt. Llavors, perquè hi ha errors de la parla que també es produeixen en aquestes unitats, les quals són estructures fixades a la llengua? En aquest treball s'analitzaran els errors de la parla en aquestes expressions multiparaula i s'estudiarà si la freqüència més alta de les paraules error influeix, ja que potser la causa dels errors sigui la baixa freqüència de les paraules originals.

Paraules clau: errors de la parla, corpus, expressions multiparaula, composicionalitat

ABSTRACT

In spontaneous speech, there is always the presence of proverbs and idioms, since they are a part of the linguistic tradition to express a complex reality. These units are not compositional, because their meaning doesn't depend on the meaning of every lexical unit in the construction, but instead the construction has meaning as a whole. If so, why are there mistakes in these units, if they have a fixed linguistic structure? In this paper, the speech mistakes in these multiword expressions will be analysed, as well as the influence of the higher frequency of the words (and expressions) that contain this kind of errors, since the cause for these errors might lie in the low frequency of the original words.

Key words: speech mistakes, corpus, multiword expressions, compositionality

ÍNDEX

1. INTRODUCCIÓ	4
2. MARC TEÒRIC	6
2.1 Idiomatics i col·locacions.....	6
2.2 Les expressions multiparaula	7
2.3 Els errors de la parla	7
3. METODOLOGIA	8
3.1 Anotació del corpus	8
3.2 Criteris per l'anotació dels errors de la parla	10
3.3 Criteris per l'anotació de les expressions multiparaula	13
4. ANÀLISI DE LES DADES I RESULTATS	14
4.1 Anàlisi i resultats dels errors de la parla	15
4.2 Anàlisi i resultat de les freqüències de les paraules afectades i les expressions error ...	17
4.2.1 Anàlisi de les freqüències de les expressions error.....	17
4.2.2 Anàlisi de la freqüència de les paraules error	19
5. CONCLUSIONS.....	21
6. ANNEX.....	24
7. BIBLIOGRAFIA	24

1. INTRODUCCIÓ

Un dels termes més emprats en la Semàntica i en la Lingüística en general és el concepte d'idiomatisme. El fet que els idiomatismes no compleixin el principi de composicionalitat fa que l'estudi d'aquestes unitats sigui interessant per a la recerca semàntica. Aquest terme engloba una sèrie de construccions lingüístiques fixades a la llengua que tradicionalment han rebut el nom de frases fetes, refranys, unitats fraseològiques... La motivació d'aquest treball va venir precisament de l'observació que si aquestes unitats estan fixades a la llengua, el seu ús és freqüent en la llengua de parla espontània i, normalment, el parlant les utilitza regularment per designar conceptes complexos, com és que hi ha parlants que poden cometre errors dins d'aquestes unitats i, tot i així, designar el mateix significat?

Des de l'estudi diacrònic de la llengua, s'han analitzat aquestes construccions de manera descriptiva. Aquesta descripció inclou les expressions originals, la seva etimologia i, com a molt, s'han classificat segons la tipologia de les paraules que componen el seu significat (García-Borrón, 2013) Però, com ja s'ha dit abans, els parlants produeixen errors en la producció oral d'aquestes construccions i un dels motius pot ser que aquests no coneixen l'etimologia de totes les frases fetes i refranys, sinó que coneixen el significat de les construccions i, per això, tot i que es produeixin errors en la producció, els parlants som capaços d'associar el mateix significat a l'expressió.

Per tot això, es va decidir fer un estudi des de la sincronia, ja que es pressuposa que els parlants no tenen aquest coneixement lingüístic de les expressions multiparaula que produeixen i de les seves etimologies i que, simplement, es diuen perquè formen part de la tradició lingüística d'una llengua i de la seva comunitat, per expressar conceptes complexos. De fet, les expressions com les frases fetes, els refranys, etc. formen part de l'adquisició més tardana del llenguatge, pel procés de metaforització que impliquen, ja que aquestes unitats acostumen a ser metàfores d'una realitat complexa i es necessita un coneixement lingüístic ampli per tal d'adquirir-les.

És per aquests motius que l'objectiu principal inicial d'aquest treball és saber perquè es produeixen aquests errors en la parla freqüent: si depenen del parlant i del seu coneixement lingüístic i del món per tal d'adaptar paraules desconegudes a paraules que sí coneix i, en segon terme, si aquesta substitució està influenciada per l'alta freqüència de les paraules i, per tant, si les paraules originals són de més baixa freqüència que les paraules utilitzades en les expressions error.

Per tant, les hipòtesis d'aquest treball són:

- 1) El parlant adapta paraules desconegudes d'expressions multiparaula a paraules conegudes. Aquest fet implica que el parlant utilitzarà vocabulari conegut i no paraules noves
- 2) En aquest tipus d'error, la freqüència serà un paràmetre important. Això implica que la o les paraules de l'error seran de més alta freqüència que les paraules originals

Els errors de la parla és un tema prou estudiat des de la disciplina de la psicolingüística, ja que són errors que produeix el parlant sense adonar-se que realment són errors i, per tant, no n'és conscient. Tampoc són errors que produeixin canvis regulars en la llengua; simplement, depenen del parlant. Aquesta disciplina i, precisament, en aquest tema, afirma que: els errors que es produeixen en la parla estan adaptats al coneixement lingüístic del parlant, perquè no es tracta de pseudoparaules, sinó de paraules pròpies de la llengua, i aquests errors també són sistemàtics, ja que no presenten formes aleatòries (Belinchón et al. 1992). El problema és que no hi ha gaire bibliografia d'aquests errors que afectin a construccions multiparaula, sinó que la bibliografia existent es basa en les produccions de paraules dins dels límits d'una frase que no segueixen cap estructura en concret. Per aquest motiu, en aquest treball s'haurà d'adaptar les propostes de la bibliografia existent al nostre tema d'estudi.

Però, realment, què és una frase feta, un refrany o una construcció multiparaula? Des de la semàntica sincrònica, que és la perspectiva des d'on s'ha enfocat tot el treball, no són gaire comuns els termes frases fetes o refranys, ja que on són els límits entre uns i altres? És per això que és més comú parlar d'expressions multiparaula (estudiades des de la Lingüística Computacional també) o d'idiomatismes i col·locacions, tot i que aquests últims termes també tenen controvèrsia pel mateix motiu: on és la frontera entre un idiomatisme i una col·locació? Val a dir que els termes emprats en l'estudi diacrònic de la llengua, com són refranys o frases fetes, són els termes més tradicionals i els més coneguts entre tots els parlants i, per això, en l'inici d'aquest treball també es van haver de definir alguns límits per tal de poder tenir clar l'objecte d'estudi. En conseqüència, primer es va observar la idiomacitat de les construccions o, el que és el mateix, que el significat no fos composicional de les paraules que les componen, sinó que les expressions tinguessin un significat propi. Per altra banda, ens vam acostar a les unitats multiparaula, que des de la lingüística computacional són formalment més detectables. Finalment, es van adaptar una sèrie de criteris sobre el nostre corpus i això ha permès delimitar les expressions a tractar.

2. MARC TEÒRIC

Un cop vista una petita introducció, es presentarà en aquesta secció una sèrie de conceptes que són importants per emmarcar el treball. Per tal de poder situar al lector, es definiran diferents termes que són clau per la comprensió de l'anàlisi de les dades. Es tracta de dos conceptes relacionats amb la semàntica no compositiva: els idiomes i les col·locacions (2.1), les expressions multiparaula (2.2) i, en darrer cas, amb els errors del llenguatge (3.3).

2.1 Idiomes i col·locacions

Segons Carreter, 1953, un idiomisme és un tret lingüístic (lèxic, morfològic o sintàctic) que és propi de l'idioma en concret; l'idiomisme sintàctic també rep el nom de gir idiomàtic i, si és una frase completa, és una frase idiomàtica. Com es pot observar, és una definició prou àmplia i tradicional que tampoc serveix per establir uns límits concrets, perquè què és un tret lingüístic propi de l'idioma en concret? A més, des de la Semàntica, els idiomes s'han abordat des de moltes perspectives i, fins i tot, dins d'elles no hi ha unanimitat per establir els idiomes. Tot i això, els criteris més comuns són: el significat total d'una expressió idiomàtica és diferent dels significats dels seus constituents lèxics i, per tant, constitueixen una unitat semàntica no composicional; no accepten la substitució d'alguns dels seus elements per un sinònim o un quasisinònim; alguns aspectes gramaticals, com el temps verbal, poden ser lliures o no en un idiomisme –alguns autors com Katz (1964) consideren que si una construcció permet fer-se en passiva sense perdre el sentit figurat, no és un idiomisme- i, sobretot, aquestes construccions es fixen a la llengua i no permeten canvis.

Per contra, les *col·locacions* també es poden interpretar com un conjunt de paraules on el seu sentit depèn de la seva aparició o, en un sentit més estadístic, com una combinació de paraules amb una alta freqüència d'aparició juntes, les quals també poden patir un procés de fossilització. Això equival al procés de metaforització, pel qual una construcció concreta o una paraula pateix una irradiació metafòrica (una ampliació del significat) mentre el parlant segueix percebent el significat inicial i, per tant, es produeix una analogia. Es diu que una metàfora s'ha fossilitzat quan s'ha normalitzat en la comunitat lingüística aquest nou significat i els parlants ja no tenen present l'origen de l'expressió o la base de l'analogia i, llavors, la col·locació seria un idiomisme.

2.2 Les expressions multiparaula

Recentment, i sobretot en la lingüística computacional i cognitiva, s'utilitza el terme d'*expressions multiparaula* per poder establir unes fronteres més clares i no fer distinció entre col·locació i idiomatisme (tot i que tampoc hi ha unanimitat en la definició dels criteris). Aquestes unitats no segueixen un patró estructural lingüístic i tampoc segueixen el principi de composicionalitat dels seus termes, sinó que són unitats complexes on el significat és un tot de les paraules que la formen, les quals tenen també una alta freqüència d'aparició, i és un significat únic d'entre tots els termes, i és per això que tampoc poden ser predictibles a partir de les paraules que la formen. Segons Hawwari et al. (2014) una expressió multiparaula no creua mai els límits de la frase i sempre té un comportament propi per la seva interpretació no composicional. D'altres autors com Masini (2005) o Tsvetkov & Wintner (2018) les defineixen més des de la psicolingüística, ja que consideren que els parlants nadius tenen aquestes expressions guardades en el seu lèxic com un únic concepte lingüístic, ja que tot i que siguin dues paraules o més, només tindran un referent. Bannard (2006) i Belinchón (1999) també les classifiquen segons el comportament d'aquestes expressions i la conseqüent determinació del seu significat, ja que Bannard considera que aquestes construccions sí tenen certs graus de composicionalitat gradualment: des de completament composicionals (que seria el que s'ha definit abans com a col·locació en el sentit més estadístic) fins a les expressions idiomàtiques. En conclusió, aquest concepte permet un sentit més ampli i una definició que pugui abastar les diferents construccions que són característiques i pròpies de cada llengua i aquest terme, amb uns criteris establerts prèviament, han sigut els utilitzats en aquest treball.

2.3 Els errors de la parla

Els errors de la parla o, tradicionalment anomenats *lapsus linguae*, són errors que es produeixen en la parla espontània sense que el parlant en sigui conscient i es produeix una desviació entre la producció que el parlant pretenia dir i la que realment diu. Tots ells estan relacionats amb components de la gramàtica i es produeixen per manca d'algun element, addició d'un element innecessari o alteració en l'ordre de la producció. Segons el tipus d'unitat afectada, hi ha una classificació d'errors de la parla que és la següent -la classificació d'aquesta descripció i de l'anàlisi de les dades es basa en el treball d'Alves Mogedas (2014)-:

- Errors per substitució: es produeixen quan es substitueix una paraula (o una part d'aquesta) per una altra paraula que no apareix al discurs.

- Errors per intercanvi: es produeixen quan s'intercanvien l'ordre de dues paraules dins d'una mateixa oració.
- Errors per addició: es produeixen quan s'afegeix a l'oració una paraula o un sintagma innecessari.
- Errors per elisió: es produeixen quan s'elideix una paraula, un morfema o un grup de fonemes respecte de l'element que es volia produir originàriament. Com a resultat, s'obté una oració sense sentit o una paraula que no tingui cap significat en la llengua.

3. METODOLOGIA

Com s'ha introduït abans, l'objectiu d'aquest treball es comprovar si els errors de la parla que es produeixen en expressions idiomàtiques s'adapten al coneixement lingüístic dels parlants i si aquests errors també estan influenciats per l'alta freqüència de les paraules error vers la baixa freqüència de les paraules originals de les expressions. Per tal d'arribar a assolir aquests objectius, s'ha hagut de fer una petita recopilació de dades en un corpus on s'ha anotat les característiques més rellevants per l'extracció de les dades i la seva posterior anàlisi; després, s'han seleccionat i adaptat els criteris dels errors del llenguatge i, finalment, s'han seleccionat i adaptat els criteris per a l'anotació de les expressions multiparaula.

Per tant, d'una banda es descriurà l'anotació realitzada i la recollida de dades (3.1); d'altra banda, s'explicaran els criteris utilitzats per l'anotació dels *errors de la parla* (3.2) i, finalment, s'explicaran els criteris utilitzats per descartar les construccions que no eren *multiparaula* (3.3).

3.1 Anotació del corpus

El corpus d'aquest treball consisteix en una base de dades que recull tot el treball de camp; és a dir, recull totes les expressions multiparaula del castellà que s'han anat recollint, sobretot, de tres fonts: del llibre *No seas pájaro de paraguero* (Abadía et al., 2018) de dues pàgines web sobre frases cèlebres de personatges famosos¹ i de produccions reals de gent que s'han anat escoltant durant l'elaboració del treball. Precisament, aquesta espontaneïtat i modalitat oral dels parlants en produir aquestes construccions ha estat la que ha dificultat la recollida de dades, atès que només hi ha reculls d'aquestes construccions en la seva forma original sense cap error. En un primer moment, quan es va plantejar aquest treball, només haurien de ser frases fetes i

¹ Les adreces de les quals són <https://vandal.elespanol.com/foro/mensaje/605502/frases-celebres/> i <http://blogs.gamefilia.com/indhira-gh11/31-03-2010/31915/frases-estupidas-de-famosos-vol-4>

refranys però, a causa de l'escassa bibliografia trobada, es van incloure més construccions multiparaula ja que era la que podia abastir més unitats lingüístiques recollides al treball

El criteri principal per anotar les construccions ha estat que provinguessin de la llengua oral i de la tradició, com poden ser refranys o frases fetes. Per emmagatzemar aquestes dades, s'ha utilitzat el programa Excel i s'ha utilitzat també com a eina d'anotació, on també s'han anotat les expressions originals, s'han classificat les dades segons els criteris dels *lapsus linguae* i s'han aplicat els filtres de les expressions multiparaula (s'han marcat en vermell aquelles que violen aquests criteris) (figura 1).

A part d'aquest anàlisi, es pretén veure la freqüència de les expressions tant error com les originals per veure si hi ha alguna expressió error més freqüent que l'expressió original i també, a totes aquelles on la unitat afectada de l'expressió és a nivell de paraula, ja que es pretén veure si la freqüència de la paraula error és més alta que la de la paraula original o meta², per tal d'afirmar la segona hipòtesi d'aquest treball. Per això, s'ha utilitzat el cercador Google com a corpus, per mirar les ocurrencies de les expressions, i el Corpus del Español³ per tal d'extreure la freqüència de les paraules que les componen. El motiu pel qual Google s'ha utilitzat com a corpus és perquè, ja que és molt difícil trobar aquestes expressions en llengua escrita i, per tant, no hi ha cap eina que ho pugui fer, el cercador inclou un volum important de pàgines web on es poden trobar tant llengua escrita formal com llengua escrita informal més propera a l'oralitat (xats, fòrums, etc.).

1	IDIOMATISMOS CON ERROR	IDIOMATISMOS ORIGINALES	UNIDAD AFECTADA	NATURALEZA DEL ERROR
2	Yo no quiero irme con los perros de Úbeda	No quiero irme por los cerros de Úbeda	Fonema	Substitución 1
3	No seas pájaro de paraguero	No seas pájaro de mal agüero	Sintagma	Elisión
4	Es más astuta que las gallinas	Es más puta que las gallinas	Palabra	Substitución 1
5	Y todo quedará en agua de borrascas	Todo quedará en agua de borrajas	Palabra	Substitución 1
6	Hacia tanto calor que no corría ni una hoja de aire	Hacia tanto calor que no corría nada de aire	Palabra	Substitución 4
7	No se pueden poner vallas al cielo	No se pueden poner vallas al campo	Palabra	Substitución 3
8	Caían chuzos y punto	Caían chuzos de punta	Sintagma	Substitución 1
9	Esto es como enhebrar una aguja en un pajar	Esto es como encontrar una aguja en un pajar	Palabra	Substitución 1
10	No siempre se puede estar en lo más alto de la cresta de la ola	Estar en lo más alto/en la cresta de la ola	Sintagma	Substitución 4
11	No quiero ser correctamente político	No quiero ser políticamente correcto	Sintagma	Intercambio 2
12	Esto es de cajón de pato	Esto es de cajón/esto es cojón de pato	Sintagma	Substitución 4
13	Esto, por debajo de mi cadáver	Por encima de mi cadáver	Palabra	Substitución 6
14	Me parece súper interesante lo que me cuentas así que, por favor, manténme al	Por favor, manténme al tanto	Palabra	Substitución 2
15	Y dicen que les engañamos. ¡Nada más cierto de la realidad!	Nada más alejado de la realidad	Palabra	Substitución 2
16	Yo, ni entro ni corto	Ni pincho ni corto / ni entro ni salgo	Palabra	Substitución 4
17	Este es un tema vudú	Ser un tema tabú	Fonema	Substitución 1
18	Claro, ya sé que lo dije para adornarme la píldora, pero...	Dorarme la píldora	Palabra	Substitución 1
19	Callando, callando y con el mazo dando	A Dios rogando y con el mazo dando	Sintagma	Substitución 2
20	Si os parece interesante nos sentamos encima de una mesa y nos ponemos de ac	Nos sentamos frente a una mesa y nos ponemos de acuerdo	Palabra	Substitución 6
21	Esto se está sacando fuera de tiesto	Sacar las cosas de quicio	Sintagma	Substitución 4
22	Y lo digo sin actitud hacia nadie	Decirlo sin actitud hacia nadie	Fonema	Substitución 1
23	Tienes más cuento que Picio	Más feo que Picio / Más cuento que Calleja	Sintagma	Substitución 4
24	Eso sí que fue una puñalada baja	Ser una puñalada traperera / Ser un golpe bajo	Palabra	Substitución 4
25	Té creó a pies puntillas	Crear a pies puntillas	Fonema	Substitución 1
26	Estás rizando el bucle totalmente	Estás rizando el rizo totalmente	Palabra	Substitución 1
27	Tiró de la manta con todo el equipo	Tirar de la manta / Caer con todo el equipo	Sintagma	Substitución 4
28	El alquillino no me paga	El inquilino no me paga	Morfema	Substitución 4

Figura 1. Imatge del corpus

² S'entén com *element error* aquell que s'ha produït en el *lapsus linguae* i que, com indica, és l'error de l'expressió; per contra, l'*element meta* és el que hauria de ser en la producció, l'element de la expressió original.

³ Accessible a: <https://corpus.byu.edu/iweb/>

Finalment, el corpus consta de 115 expressions escollides, de les quals se n'han descartat 22 per no complir amb els criteris de multiparaula i l'estructura per anotar ha estat:

- En un full, s'ha anotat per columnes les expressions error, les expressions originals, la unitat afectada dins de la construcció i la classificació segons quin fos l'error de la parla (figura 1).
- En un altre full, s'ha tornat a anotar per columnes les expressions error, les expressions originals, la freqüència de les paraules error, la freqüència de les paraules originals (això només de les expressions on la unitat afectada fos la paraula), la freqüència de la construcció error i la freqüència de la construcció original. (figura 2).

	A	B	C	D	E	F	G
	IDIOMATISMOS CON ERROR	IDIOMATISMOS ORIGINALES	FRECUENCIA PALABRA	FRECUENCIA PALABRA	OCURRENCIAS COLOCACI	OCURRENCIAS COLOCACION META	
1	Yo no quiero irme con los perros de Úbeda	No quiero irme por los cerros de Úbeda			4	11.200	
3	No seas pájaro de pagadero	No seas pájaro de mal agüero			45.600	45.600	
4	Es más astuta que las gallinas	Es más pura que las gallinas	315	776	983	58.000	
5	Y todo quedará en agua de borrajas	Todo quedará en agua de borrajas	425	13	15.000	137.000	
6	Hacia tanto calor que no corría ni una hoja de aire	Hacia tanto calor que no corría nada de aire			0	1.770	
7	No se pueden poner vallas al cielo	No se pueden poner vallas al campo	33568	20216	6	519	
8	Cifan chuzco y punto	Cifan chuzco de punta			3	416.000	
9	Esto es como enhebrar una aguja en un pajar	Esto es como encontrar una aguja en un pajar	8	4183	1.160	41.800	
10	No siempre se puede estar en lo más alto de la cresta de la ola	Estar en lo más alto en la cresta de la ola			4	700.000 / 74.100	
11	No quiero ser correctamente político	No quiero ser políticamente correcto			259	143.000	
12	Esto es de cañón de pato	Esto es de cañonero es coñón de pato			1	678.000 / 2.330	
13	Esto, por debajo de mi cadáver	Por encima de mi cadáver	7271	8022	29.300	36.200	
14	Me parece súper interesante lo que me cuentas así que por favor mantenme al rrr	Por favor, mantenme al tanto	1576	89791	1.420	6.780	
15	Y dicen que les engañamos. Nada más ciento de la realidad!	Nada más alijado de la realidad			94.500	163.000	
16	Yo, ni entro ni cocto	Ni picho ni cocto	15083	211.19385	7	7.030 / 12.500	
17	Este es un tema vuóó	Ser un tema tabú			600	218.000	
18	Claro, ya sé que lo dijo para adornarme la píldora, pero...	Dorarme la píldora	880	132	7.510	36.100	
19	Callando, callando y con el mazo dando	A Dios rogando y con el mazo dando			1	85.100	
20	Sí os parece interesante nos sentamos encima de una mesa y nos ponemos de	Y los sentamos frente a una mesa y nos ponemos de acuerdo			58.700	42.800	
21	Esto se está sacando fuera de tiesto	Sacar las cosas de quicio			82	17.400	
22	Y lo digo sin actitud hacia nadie	Decirlo sin actitud hacia nadie			1.050	2.760	
23	Tienes más cuento que Picio	Más feo que Picio / Más cuento que Calleja			3	18.200 / 15.900	
24	Esto sí que fue una puñalada baja	Clavar una puñalada por la espalda / Ser una puñalada traspera / Ser un golpe bajo	19139	28	1.340	105.000 / 426.000	
25	Te creo a pies puntillas	Cree a pies puntillas			475	21.600	
26	Estás ritando el bueie totalmente	Estás ritando el ribo totalmente	20	90	4.710	205.000	
27	Tiré de la manta con todo el equipo	Tirar de la manta / Caer con todo el equipo			0	86.900 / 35.500	
28	El alquiler no me paga	El inquilino no me paga			44	44.600	
29	Seguimos tu carrera meteorífica	Seguimos tu carrera meteorífica					
30	Es un estafador. Con todas sus palabras	Es un estafador. Con todas sus letras	20229	6588	681.000	216.000	
31	Todo cambió con el algoritmo de Google	Todo cambió con el algoritmo de Google	11	9	3.010	221.000	
32	A mí me pagaron una parte en dinero y otra en especias	A mí me pagaron una parte en dinero y otra en especie	110	9677	3.740	125.000	
33	Hay que mejorar los costes sin decremento de la calidad	Ir en decremento / Evitar el decremento			1.090	1.370.000 / 203.000	
34	La formación en valores se da en casa y, o, barra, en la escuela	Se da en casa y/o en la escuela.					
35	Tu padre como siempre al pie del camión	Estar al pie del cañón	375	775	5	58.600	
36	Decidid rescindir de sus servicios	Decidid prescindir de tus servicios o rescindir el contrato			21.200	603.000 / 632.000	
37	Está de agua, pan y moja	Está de pan y moja			2	16.900	
38	En todas partes cuecen almas	En todas partes cuecen habas	37655	59	52	203.000	

Figura 2. Imatge del corpus

3.2 Criteris per l'anotació dels errors de la parla

Els criteris que s'han seguit són els del treball d'Alves Mogades (2014), que es basa en l'agrupació que fa l'autora del Viso (2002) i la seva pròpia. S'ha utilitzat la seva agrupació perquè el seu objecte d'estudi es centra només en les unitats lèxiques (les més properes a l'objecte d'estudi d'aquest treball) i perquè era la classificació més clara i directa trobada fins aleshores per analitzar-ho.

Com ja s'ha dit anteriorment (secció 1.3), són quatre els tipus d'errors de la parla: substitució, intercanvi, addició i elisió. Com s'ha de suposar, els criteris han estat adaptats a l'objectiu d'aquest treball i al context, ja que en l'original els *lapses* s'apliquen només a paraules i en aquest treball s'ha hagut d'aplicar i adaptar també a unitats multiparaula. Per tant, els criteris que s'han seguit són els presentats en la figura 3, que seguidament expliquem i els exemples esmentats són extrets del corpus d'aquest treball, tret que no hi hagi cap exemple, llavors s'han utilitzat exemples de la bibliografia.

SUBSTITUCIÓ	(1) A causa de raons prosòdiques, fonètiques sintàctiques i semàntiques
	(2) No hi ha semblança fonètica entre els elements
	(3) No hi ha compatibilitat semàntica entre l'element nucli i l'element error que ha de fer d'argument
	(4) L'element error és la combinació de dos elements meta
	(5) Substitució d'un sufix per un altre
	(6) Els elements meta i error pertanyen al mateix camp semàntic
INTERCANVI	(1) Paraules no adjacents
	(2) Paraules adjacents
ADDICIÓ	
ELISIÓ	

Figura 3. Classificació dels errors de la parla

-Classificació per substitució. Hi ha sis subtipus de substitucions:

-Substitució 1 prosòdia, sintaxi i semàntica: poden ser per raons prosòdiques, fonètiques i gramaticals, on l'element error i l'element meta⁴ comparteixen el patró accentual, el número de síl·labes, la categoria gramatical, el gènere i el nombre, entre d'altres o per raons semàntiques i sintàctiques, on els dos elements tenen una relació semàntica més estreta perquè els dos formen part del mateix camp semàntic i també podrien ser seleccionats pel mateix nucli verbal (a).

(a) Es más *astuta* (*puta*) que las gallinas

-Substitució 2 d'absència de semblança fonètica: en aquest grup l'element error i l'element meta no comparteixen semblança fonètica però sí prosòdica, gramatical i sintàctica (b).

(b) Mantenme al *margin* (*tanto*)

-Substitució 3 d'incompatibilitat semàntica: els elements d'aquest grup comparteixen semblances fonètiques, prosòdiques i gramaticals però l'element error no pot ser un argument del verb perquè no comparteix raons semàntiques per estar regit per aquest nucli amb l'element meta (c).

(c) En todas partes cuecen *almas (habas)*

-Substitució 4 de combinació de dos termes: en aquest grup l'expressió error és el resultat d'haver barrejat dos elements meta. El resultat d'aquesta combinació és una expressió no existent en la llengua, tot i estar formada per elements que sí són reals i aquest tipus d'error és el més comú en expressions multiparaula (d).

(d) Le han pillado *con las manos en el carrito del helado (con las manos en la masa/con el carrito del helado)*

-Substitució 5 de sufixos: l'element afectat és a nivell morfològic dins de la frontera de la paraula i, més concretament, afecta un sufix. En el nostre treball només hi ha un cas (e) i s'ha adaptat la substitució, ja que s'ha senyalat només pel fet que l'element afectat és un sufix gramatical de flexió.

(e) Ir de *incògnita (incògnito)*

-Substitució 6 de coincidència de camp semàntic: els dos elements no comparteixen similitud fonètica però sí formen part del mateix camp semàntic (i, per tant, poden ser arguments del mateix verb) i comparteixen la mateixa categoria gramatical (f).

(f) A este tío no te llega *a la punta del tacón (a la suela del zapato)*

-Classificació per intercanvi. Dins d'aquesta categoria es troben dos subtipus:

-Intercanvi 1 de paraules no adjacents: es produeix entre elements que comparteixen la mateixa categoria gramatical, l'estructura accentual i el número de síl·labes o, simplement, que tenen la mateixa categoria gramatical però no tenen en comú cap similitud fonètica. Els elements de l'intercanvi 1 no han d'anar seguits dins de l'oració. En algunes ocasions, s'ha caracteritzat com intercanvi 1 expressions que intercanvien l'ordre d'un element dins de la oració, tot i que el segon element no estigui present en l'element meta; és a dir, en l'exemple (g), *trinchera* s'ha intercanviat amb l'element meta *agujero*, tot i que després aquest no surti en l'expressió error.

(g) *En tiempos de guerra, toda trinchera es buena (Todo agujero es trinchera)*

-Intercanvi 2 de paraules o parts de paraules adjacents: aquest error es produeix entre paraules que van seguides dins de l'oració i tenen una similitud fonològica. Tot i això, han de tenir diferents categories gramaticals (h).

(h) *Discuto disfrutando (Disfruto discutiendo)*.

-Classificació per addició.

En principi, aquest lapsus seria la repetició de dues paraules redundants, com *bajar hacia abajo*. En el propòsit d'aquest treball, s'ha adaptat aquest lapsus a aquelles expressions que han quedat alterades per l'addició d'un element que té una relació semàntica amb l'element meta però que no forma part de cap expressió idiomàtica relacionada (i).

(i) *Igual suena la flauta de Hamelin (igual suena la flauta)*

-Classificació per elisió.

L'únic cas anotat al corpus és (j) i s'ha classificat d'aquesta manera perquè s'ha elidit una part de la paraula i, com a resultat, s'ha produït una paraula que sí existeix en la llengua però no forma part de cap expressió multiparaula similar; per tant, l'element resultant no forma part de la llengua.

(j) *No seas pájaro de paragüero (de mal agüero)*

Un cop revisada la classificació dels errors de la parla es presentarà a continuació els criteris i l'adaptació que s'ha realitzat per a l'anotació de les expressions multiparaula.

3.3 Criteris per l'anotació de les expressions multiparaula

Per poder fer una selecció objectiva de totes les expressions que s'havien recollit per formar part del corpus, s'han utilitzat els criteris de Castellón et al. (2014) amb l'objectiu d'establir les restriccions que ha de complir un expressió multiparaula. Les autores divideixen els criteris en tres parts, aquells que afecten al significat, aquells que afecten a l'estructura sintàctica i aquells que afecten a les paraules:

-Semàntics: el tret principal és la composicionalitat. Els termes d'una unitat lèxica multiparaula no s'han de poder substituir per la seva definició lexicogràfica, sinó que el conjunt dels termes són els que formen el significat total de l'expressió. Conseqüentment, si els termes no es poden substituir per tal d'obtenir un significat concret, això vol dir que aquesta unitat també tindrà una lectura figurada i, per tant, els significats es componen d'una fase prèvia de metaforització de l'expressió multiparaula.

-Sintàctics: en general, *les unitats lèxiques multiparaula difícilment accepten transformacions sintàctiques i, per tant, tendeixen a presentar-se com a estructures fixes* (Castellón et al. 2014). Tot i això, hi ha un seguit de trets que caracteritzen aquestes expressions:

-Aquelles formades per un substantiu i un adjectiu no admeten modificacions lèxiques, com afegir un adverbi modificador.

-Si una expressió està formada per un subjecte i el verb o un verb i el seu argument intern i un d'ells té una lectura figurada, serà una expressió lèxica multiparaula.

-No s'accepta la relativització d'un dels termes d'aquesta unitat multiparaula ni la passivització.

-Moltes d'aquestes construccions es formen a partir d'un verb lleuger i d'un substantiu i, per això, s'ha establert que els verbs *dar*, *tomar*, *hacer*, *poner* i *tener* són els verbs de suport per al castellà. El verb *hacer* més un infinitiu no es considerarà unitat lèxica multiparaula, ja que aquesta formació aporta sentit de causalitat.

-Respecte a les unitats formades per un sintagma preposicional, no es consideren unitats lèxiques aquelles que designen una pertinença de la entitat a un grup o unitat.

-Lèxics: els criteris lèxics que s'han establert són més concrets que els sintàctics, ja que parla de grups específics que sí són expressions multiparaula com poden ser dates o termes científics expressats per més d'una unitat lèxica. Tot i això, aquestes restriccions no s'han aplicat en aquest treball perquè no ha estat necessari.

-Per defecte: al final, la caracterització d'una unitat multiparaula, tot i que hagi de passar per tots els filtres previs, serà caracteritzada d'aquesta forma si l'annotador així ho considera, atesa la dificultat de la caracterització d'aquestes unitats lèxiques.

4. ANÀLISI DE LES DADES I RESULTATS

En aquest apartat es descriurà com s'han analitzat les dades i quins són els resultats de l'anàlisi dels errors de la parla (4.1) i quins són els resultats de la freqüència de les paraules afectades i les expressions error (4.2).

En primer lloc, es van classificar totes les expressions abans d'aplicar els criteris de multiparaula segons les unitats afectades i els tipus d'errors. Les dades són que 9 expressions estan afectades a nivell fonema, 3 a nivell morfema, 59 a nivell de paraula i 44 a nivell de sintagma. D'aquestes i segons la classificació dels errors de la parla, només una expressió conté una elisió, tres contenen una addició, 6 contenen errors d'intercanvi i 105 són causades per errors de substitucions.

En segon lloc, es van aplicar els criteris d'expressions multiparaula i moltes d'aquestes construccions es van rebutjar pel criteri de composicionalitat, ja que el seu significat es podia deduir de les paraules que formaven la construcció, com en la construcció error *No quiero ser correctamente político (No quiero ser políticamente correcto)*.

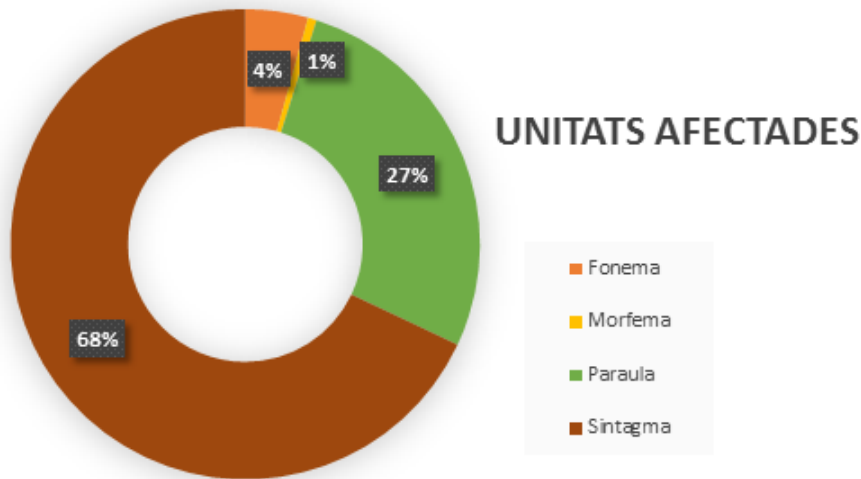
Finalment, en un altre full de l'Excel es van anotar la freqüència de totes les expressions (error i meta) i de totes les paraules que formaven part d'expressions error, on la unitat afectada era la paraula. En aquesta segona part, només es van anotar les expressions que no havien estat rebutjades segons els criteris d'expressions multiparaula.

Totes les expressions multiparaula que formen el corpus tenen una estructura oracional amb el verb principal i els seus arguments i la majoria són oracions simples, ja que no tenen cap oració subordinada dins de l'estructura.

4.1 Anàlisi i resultats dels errors de la parla

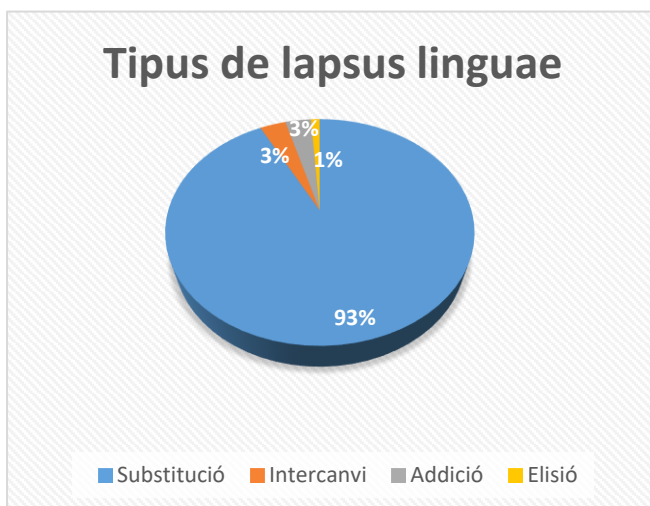
Del total de 115 expressions recollides inicialment, se'n van descartar 22 per no passar els criteris per ser expressions multiparaula. Com ja s'ha dit anteriorment, si es considera que els errors de les expressions multiparaula són errors de la parla, tal i com es defineix en la disciplina de la psicolingüística, això ja confirma la primera hipòtesi del treball, ja que si hi ha errors en les expressions multiparaula (que en principi tenen estructures fixes i una alta freqüència d'aparició, a diferència dels errors de la parla comuns) és perquè el parlant, sense ser-ne conscient, adapta una paraula desconeguda d'aquesta estructura fixada per una altra paraula que forma part del seu coneixement lingüístic. Per tot això, s'ha fet una classificació dels errors de la parla més comuns en aquestes expressions, per tal de veure també com s'adaptaven en aquests contextos, on no depèn només d'una paraula, sinó de dues o més.

D'una banda es van analitzar les unitats afectades i el resultat va ser que les principals unitats eren les paraules i els sintagmes (*gràfic 1*).

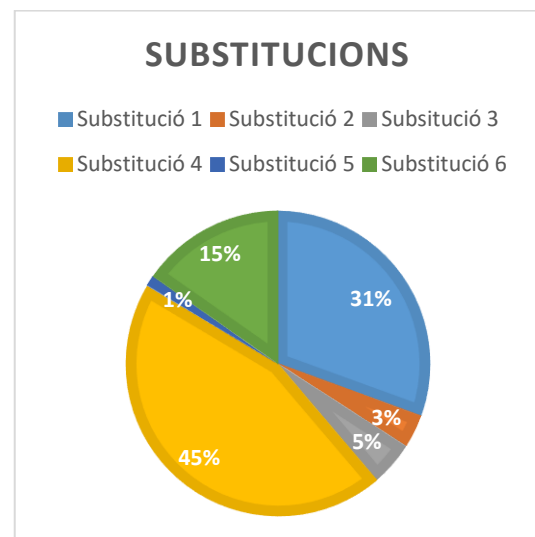


1. Gràfic de les unitats lingüístiques afectades per els lapsus linguae (elaboració pròpia)

D'altra banda, la classificació segons els *lapsus linguae* ens indica que el 92,39% d'aquests són substitucions, el 3% intercanvis, un altre 3% addició i només un 1% elisió (gràfic 2). Finalment, es va comprovar que el 44,71% d'aquestes substitucions són del tipus 4 (citad dins de l'apartat 2.2 *Criteris per l'anotació dels errors de la parla*); és a dir, són errors que provenen de la fusió de dos elements meta i que, per tant, són les més comunes en aquest tipus d'errors (gràfic 3). Aquest tipus d'error ha estat com el "calaix de sastre" per l'anotació dels errors de la parla d'aquestes construccions, ja que ha estat el més freqüent i perquè quasi bé tots els errors que es produeixen en les expressions multiparaula són resultat de la fusió de dues altres expressions multiparaula.



2. Gràfic dels tipus de *lapsus linguae* analitzats al corpus (elaboració pròpia)



3. Gràfic que expressa el percentatge de lapsus de substitucions al corpus (elaboració pròpia)

Conseqüentment, aquestes dades avalen la hipòtesi en certa forma: sí certifiquen que les paraules error provenen del coneixement lingüístic del parlant i que, per tant, aquest les coneix (ja que sinó no es podria parlar d'errors de la parla des d'una perspectiva psicolingüística) però no tots els errors provenen de la desconexença d'alguna unitat de les expressions originals de la mostra, sinó que també es poden produir errors per la fusió de dues expressions (com *Tienes más cuento que Picio*) i que, de fet, aquests són els més comuns en aquest tipus d'expressions. A continuació, es presentaran en els punts següents l'anàlisi i els resultats de les freqüències de les expressions error i l'anàlisi de les freqüències de les paraules afectades.

4.2 Anàlisi i resultat de les freqüències de les paraules afectades i les expressions error

L'altre aspecte que es pretenia observar amb l'elaboració del corpus, com s'ha mencionat anteriorment, és la freqüència de les paraules afectades i de les expressions error. Recordem que aquesta anàlisi permetria analitzar la freqüència de les expressions error, per tal de veure si hi havia algunes més freqüents amb error que la seva forma original, la qual cosa també permetria confirmar la segona hipòtesi: si la paraula meta té una freqüència més baixa que la paraula error, aquesta diferència de freqüència podria ser a la causa del lapsus. Això pot tenir a veure amb el fet que el parlant desconex la paraula de l'expressió meta.

4.2.1 Anàlisi de les freqüències de les expressions error

En aquest apartat es presenta l'anàlisi de les ocurrències que tenen aquestes expressions a Google, tant les que contenen l'error com les originals. Després de buscar-les al cercador, es pot afirmar que només 12 expressions són les que superen les 10.000 ocurrències:

- *Y todo quedará en agua de borrascas (agua de borrajas)*
- *Esto, por debajo de mi cadáver (por encima de mi cadáver)*
- *Es un estafador. Con todas sus letras (palabras)*
- *Este equipo se hunde, ha tocado techo (tocar fondo)*
- *Me he leído el texto así, en horizontal (en diagonal)*
- *No sé cómo todavía esto no se sabe y no ha saltado la luz (salir a la luz)*
- *Muy bien, ya hemos roto el fuego (romper el hielo o abrir el fuego)*
- *¡Vamos, ale ale que es gerundio! (¡Arreando que es gerundio!)*
- *Hoy hay garrafas de viento (ráfagas de viento)*
- *Eso quedó en el barril de los recuerdos (el baúl de los recuerdos)*
- *Me gustan los toreros que están en el candelabro (en el candelero)*
- *Estos de OT interesan mucho a la prensa amarilla (la prensa rosa)*

Alguna d'elles té més ocurrencies al buscador perquè l'ha pronunciat algú famós i es repeteix la narració d'aquest fet, d'altres són construccions que sí existeixen en la llengua però no són vàlides en aquest context semàntic on es produeixen... Tanmateix, aquest fet demostra que: les expressions error són menys freqüents a la llengua que les expressions originals, ja que quasi el 44% de les construccions error del corpus tenen menys de deu ocurrencies a la xarxa i que els errors de la parla també depenen del parlant: cadascú canvia les paraules de les expressions segons les paraules més freqüents de la seva comunitat lingüística i, per tant, el coneixement lingüístic és variable depenent de la cognició lingüística del parlant. Tot i això, s'han trobat cinc expressions error que superen el nombre d'ocurrencies respecte de les expressions originals i aquests són:

- a. *No seas pájaro de paragüero (de mal agüero)*
- b. *Es un estafador. Con todas sus letras (palabras)*
- c. *Me he leído el texto así, en horizontal (en diagonal)*
- d. *Vamos, ale ale que es gerundio! (Arreando que es gerundio)*
- e. *He sido cocinera antes que fraila (Ser monaguillo antes que fraile)*

La primera expressió seria una excepció, ja que també és el títol del llibre que recull expressions multiparaula que s'ha fet servir en aquest treball i, potser, ha estat un factor que ha augmentat les seves aparicions a Google. Tot i això, és prou rellevant l'alta freqüència de la resta: la b i la c són expressions on la paraula modificada és per aproximació semàntica, la d és per desconexió del significat original i del de la paraula *gerundio* (de fet, demostra la fixació d'aquestes expressions multiparaula en la llengua, perquè, tot i que la paraula de l'inici de l'expressió error no sigui un gerundi, és segueix dient la terminació de la frase i se n'obté el mateix significat) i la e és la fusió de dos refranys (*Ser monaguillo antes que Fraile* i *Quien fue cocinero antes que fraile, lo que pasa en la cocina bien sabe*). Sobre aquesta última, sí està ben estesa aquesta construcció i, potser, ja és una expressió multiparaula diferent i normalitzada.

I, per finalitzar, hi ha una expressió que té un número d'ocurrencies molt similar: "Después de comer siempre me entra la *morriña* vs. la *modorra*" (1.020 i 1.940 respectivament). Val a dir que l'ús de l'expressió error és bastant freqüent aquí, a Catalunya; el que no es pot afirmar és si aquesta expressió error és menys comuna depenent del context territorial, ja que les paraules són d'origen gallec.

4.2.2 Anàlisi de la freqüència de les paraules error

Per tal de comprovar la segona hipòtesi, es va comprovar al Corpus del Español la freqüència de les paraules error, però només les construccions que havien quedat afectades a nivell de paraula, no a nivell fonètic ni de sintagma (en total, 44 expressions). Els resultats han estat que 18 de les paraules error (40,91% dels casos) tenen una freqüència més alta que la paraula original, però aquestes dades no són suficients per tal de confirmar la hipòtesi presentada en general, sinó que només és aplicable en casos concrets (*carse de un pino vs. de un guindo, adornarme la píldora vs. dorarme la píldora...*), ja que no arriba ni al 50% dels casos. A continuació, al següent quadre, es presentaran les dades.

Paraules error	Paraules meta	Freqüència paraula error	Freqüència paraula meta
Astuta	Puta	315	776
Borrascas	Borrajás	425	13
Cielo	Campo	33566	20216
Enhebrar	Encontrar	8	4183
Debajo	Encima	7271	8022
Margen	Tanto	1576	89791
Entro	Pincho	15083	19365
Adornarme	Dorarme	860	132
Golpe	Puñalada	19139	28
Bucle	Rizo	20	90
Palabras	Letras	20229	6588
Logaritmo	Algoritmo	11	9
Especias	Especie	110	9677
Camión	Cañón	375	775
Almas	Habas	37655	59
Rascarse	Rasgar	54	93
Arriesgar	Rasgar	234	93
Calzoncillo	Calzón	21*	34*
Parece	Padece	9776	2463
Techo	Fondo	1769	9421
Colchones	Algodones	305	1727
Sobre (prep.)	Entre (prep.)	131474	120061

Magia Borràgia	Magia Potàgia/Borràs	*	*
Palco	Listón	333	144
Zambomba	Mambo	14	19
Echar	Crecer	6322	1370
Suelo	Aire	14533	14572
Culo	Lengua	210	16517
Morriña	Modorra	3	151
Ciego	Tuerto	483	456
Majara	Marajá	1	0
Ginecológico	Genealógico	7	67
Barril	Baúl	224	327
Hormiga	Oveja	344	748
Candelabro	Candelero	66	248
Hormona	Horma	178	65
Cocinera	Monaguillo	362	57
Amarilla	Rosa	1486	4455
Pesadilla	Pescadilla	489	11
Loro	Oro	373	26145
Cerrado	Abierto	2069	3420
Cagado	Mojado/meado	18	473 / 11
Pino	Guindo	833	31
Bombilla	Bombero	80	49

Figura 4. Taula de les freqüències extretes del Corpus del Español

Dos casos especials han estat *Magia Potàgia/Magia Borràs* que no s'ha pogut trobar al Corpus del Español i el cas de *hablar a calzón/calzoncillo quitado*. Aquest corpus disposa d'informació des del 1800 fins a l'actualitat i, per això, *calzón* és més freqüent. Tanmateix, actualment la més freqüent és *calzoncillo* i, en aquest cas, si es possible validar la hipòtesi presentada.

Un altre factor important és la correcció o la incorrecció de paraules 'vulgars' per d'altres que no ho són. En concret, es parla de *puta* i *cagado/meado*. En el primer cas, tot i que la paraula més freqüent sigui aquesta, es creu que el parlant fa una ultracorrecció, i potser es deu per la connotació negativa de la paraula. En el segon cas, tot i que la paraula *mojado* sigui la original i la que més freqüència té, el parlant la canvia per dues paraules d'un registre més baix, potser

a causa de la informalitat d'aquestes frases fetes. Però això, de moment, només són suposicions i s'hauria d'analitzar en un estudi futur.

5. CONCLUSIONS

En aquest treball s'han hagut de realitzar diferents tasques per tal d'arribar als objectius inicials. Primer de tot, s'ha presentat un estudi que inclou la recopilació d'un corpus d'expressions multiparaula que contenen errors i s'han revisat els criteris per a la detecció d'aquestes expressions multiparaula i per poder realitzar la classificació dels errors. Seguidament, s'han classificat els errors d'aquestes expressions segons els criteris dels errors de la parla que es donen des de la psicolingüística i, finalment, s'ha presentat un petit estudi freqüencial de les paraules afectades i les expressions error. Amb tot això, es pretén demostrar o rebutjar les hipòtesis inicials:

- 1) El parlant adapta paraules desconegudes d'expressions multiparaula a paraules conegudes. Aquest fet implica que el parlant utilitzarà vocabulari conegut i no paraules noves
- 2) En aquest tipus d'error, la freqüència serà un paràmetre important. Això implica que la o les paraules de l'error seran de més alta freqüència que les paraules originals

Les conclusions finals respecte d'aquests dos objectius marcats a l'inici d'aquest treball són que:

-Els parlants sí adapten els errors produïts en les expressions multiparaula a paraules del seu domini lingüístic perquè, de fet, tots els errors de la parla produeixen paraules existents en la llengua, no pas pseudoparaules. És per això que durant aquest treball s'han classificat els errors de les expressions multiparaula en els errors de la parla o, tradicionalment dit, *lapsus linguae* seguint la classificació que normalment és la més freqüent entre els manuals de psicolingüística. Els resultats obtinguts han estat que la unitat afectada més freqüent és a nivell de sintagma i que el tipus més freqüent d'error de la parla en aquestes construccions és la substitució 4, que és la fusió de dos elements meta i provoca com a resultat un element error que no existeix; en aquest cas, seria la fusió de dues expressions multiparaula que provocaria una expressió que no existeix en la llengua, tot i que els seus elements sí.

-La freqüència de les paraules error o meta de les expressions multiparaula no sembla rellevant per produir el canvi o, si més no, no ho ha estat en la petita mostra que forma el corpus del treball. Els errors els produeixen els parlants depenent del seu coneixement d'aquestes

expressions multiparaula i del seu coneixement lingüístic, però les raons també poden ser per semblança fonètica dels termes error i meta. Tampoc hi ha gaire expressions multiparaula que s'hagin estès amb aquests errors pel mateix motiu: cada parlant té unes característiques lingüístiques diferents i els errors que produiran seran diferents. Les més esteses són degudes a que les ha pronunciat un personatge famós o que en una certa comunitat o territori lingüístic s'han estès (com el cas de la *morriña* i la *modorra*).

Aquestes són les conclusions d'aquest treball, però potser amb un corpus molt més extens els resultats canviarien. El tema dels errors en les expressions multiparaula fixades a la llengua és un tema poc explorat i la bibliografia és escassa però, igual que en aquest treball s'ha aplicat una anàlisi des de diferents disciplines, com la Semàntica, la Psicolingüística i la Lingüística de Corpus, encara es podrien afegir més per tal de fer-lo més enriquidor. Des de la Adquisició del llenguatge, s'ha estudiat l'adquisició de frases fetes i refranys, ja que s'adquireixen per el seu significat i no per les construccions estrictament, però no els errors que podem cometre els parlants. Una de les principals causes d'aquest fet, des de l'Adquisició, pot ser que aquestes construccions s'aprenen oralment i no es presenten sovint de manera escrita i, per tant, els parlants poden tenir problemes de segmentació d'aquestes expressions multiparaula que provoquin els errors de la parla. Una altra disciplina des d'on es podrien estudiar aquestes construccions és la Sociolingüística, ja que es podria observar si alguns errors són més freqüents en uns territoris que en d'altres a causa del bilingüisme, per exemple (en el corpus del treball hi ha un exemple que és *Sacar las cosas fuera de tiesto* per *Sacar las cosas de quicio*, on l'element error pot venir provocat per la fusió de l'expressió catalana *Pixar fora de test*, tot i que no es pot afirmar perquè el llibre d'on s'ha extret aquesta expressió no diu de quin territori s'ha extret), si en el cas de la *morriña* o *modorra* aquest error es dona més freqüentment en territoris més allunyats de Galícia (d'on és original la paraula *morriña*) o si algunes expressions són més pròpies d'una comunitat o territori lingüístic que d'unes altres. I aquestes dues disciplines en són dos exemples de les formes d'enriquir aquest treball, com també podria ser estudiar les ultracorreccions que els parlants produeixen en aquest tipus de construccions per tal d'evitar paraules malsonants (com *astuta* per *puta*) o estudiar aquestes construccions multiparaula en Lingüística Computacional per tal d'extreure'n la manera de processar-les automàticament pel seu significat, tot i que tinguin qualsevol error en la seva estructura.

En conclusió, opino que des de la perspectiva diacrònica de l'estudi de la llengua és un tema que ja està prou explotat i descrit però des de la perspectiva sincrònica encara queda molt per fer i per explorar. Les expressions multiparaula descrites en aquest treball són la llengua més viva dels parlants, la del carrer, la llengua que utilitzem tots els parlants en el context del dia a dia i, si realment com a lingüistes volem estudiar la llengua real, aquest tema en seria un molt bon exemple per fer-ne una aproximació.

6. ANNEX

Com s'ha explicat durant el treball, s'ha hagut de fer un corpus en el programa Excel per tal de recollir les dades i analitzar-les. Per realitzar consulta, l'enllaç al corpus és:

<https://www.dropbox.com/s/e0cv61ijj2iokis/corpus.xlsx?dl=0>

7. BIBLIOGRAFIA

- ABADÍA, G., ABADÍA, J., ABADÍA, J., ABADÍA, R., & ABADÍA, A. (2018). *No seas pájaro de paraguero y otras blabladurías*. Espasa.
- ALMOHIZEA, M. I. (2016). The Placement of Idioms in Traditional and Non-Traditional Approaches. *International Journal of Language and Linguistics*, 3(5). Recuperat de http://ijllnet.com/journals/Vol_3_No_5_November_2016/5.pdf
- ALVES MOGEDAS, A. B. (2014). Los Lapsus linguae desde una perspectiva gramatical. Recuperat de <https://dugi-doc.udg.edu/handle/10256/10036>
- BANNARD, C. J. (2006). *Acquiring phrasal lexicons from corpora* (Doctoral dissertation, University of Edinburgh).
- BELINCHÓN CARMONA, M. (1999). *Lenguaje no literal y aspectos pragmáticos de la comprensión. Psicolingüística del español*. Recuperat de <http://sire.ub.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsdnp&AN=edsdnp.997025ART&lang=es&site=eds-live>
- BELINCHÓN CARMONA, M., RIVIÈRE, A., & IGOA GONZÁLEZ, J. M. (1992). *Psicología del lenguaje : investigación y teoría*. [Madrid] : Trotta. Recuperat de http://cercabib.ub.edu/iii/encore/record/C__Rb1150159__Spsicolog%EDa del lenguaje, investigaci%F3n y teor%EDa__Orighresult__U__X7?lang=cat
- CARRETER, F. L. (1953). *CARRETER - Diccionario de términos filológicos*. Madrid: Editorial Gredos. Recuperat de <http://www.aragob.es/pre/cido/lazaroca.htm>
- CASTELLÓN, I., LLOBERES, M., & MARIMON, M. (2014). Spanish WordNet 3.0: les expressions multiparaula. En M. A. MARTÍ ANTONIN & M. TAULÉ DELOR (Ed.), *Homenatge a Sebastià Serrano* (p. 101-112). Barcelona: Publicacions i Edicions de la Universitat de Barcelona,.
- COLÓN DOMÈNECH, G., & GIMENO BETÍ, L. (2006). *Els noms i els conceptes : noves tendències en l'estudi del lèxic*. Universitat Jaume I.

- FROMKIN, V. A. (s.d.). Los aspectos gramaticales de los errores del habla. En L. Eguren (Ed.), *Panorama de la lingüística moderna de la Universidad de Cambridge* (p. 149-173). Madrid: Visor.
- GARCÍA-BORRÓN, J.-P. (2013). *Semántica de la palabra*. Barcelona : Publicacions i Edicions de la Universitat de Barcelona,. Recuperat de http://cercabib.ub.edu/iii/encore/record/C__Rb2086084__Sgarc%EDA-borr%F3n__Orightresult__U__X4?lang=cat
- GONZÁLEZ PÉREZ, R. (2008). Reanálisis semántico, procesos metafóricos y polisemia. *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)*, 239-245. Recuperat de <http://www.unav.es/linguis/simposiosel/actas/>
- HAWWARI, A., ATTIA, M., & DIAB, M. (2014). A Framework for the Classification and Annotation of Multiword Expressions in Dialectal Arabic, 48-56. Recuperat de http://www.attiaspace.com/Publications/2014_MWE_W14-36.pdf
- KATZ, J. J. (1964). *An Integrated theory of linguistic descriptions / Jerrold J. Katz, Paul M. Postal.* : MIT Press,. Recuperat de http://cercabib.ub.edu/iii/encore/record/C__Rb1183638__Skatzpostal__Orightresult__U__X2?lang=cat
- RABINOVICH, E., TSVETKOV, Y., & WINTNER, S. (2018). Native Language Cognate Effects on Second Language Lexical Choice. Recuperat de <http://sire.ub.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsarx&AN=edsarx.1805.09590&lang=es&site=eds-live>
- VISO PABÓN, S. del. (2002). Los lapsus linguae como fuente de datos en el estudio de la producción del lenguaje: un corpus de errores en castellano. Recuperat de <http://sire.ub.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsrac&AN=edsrac.61621&lang=es&site=eds-live>



Declaració d'autoria

Amb aquest escrit declaro que sóc l'autor/autora original d'aquest treball i que no he emprat per a la seva elaboració cap altra font, incloses fonts d'Internet i altres mitjans electrònics, a part de les indicades. En el treball he assenyalat com a tals totes les citacions, literals o de contingut, que procedeixen d'altres obres. Tinc coneixement que d'altra manera, i segons el que s'indica a l'article 18, del capítol 5 de les Normes reguladores de l'avaluació i de la qualificació dels aprenentatges de la UB, l'avaluació comporta la qualificació de "Suspens".

Barcelona, a _20 de juny de 2018

Signatura:

Membre de:

LE
RU

Reconeixement internacional de l'excel·lència



B:KC

Barcelona
Knowledge
Campus



Health Universitat
de Barcelona
Campus