

Corpus, frecuencia y gramaticalidad

Treinta construcciones contrincantes en tres corpus

Josse DE KOCK

En *Gramática española. Enseñanza e investigación* se recurre de manera sistemática y exhaustiva a 19, 20 o, ambos juntos, 39 textos¹. Todos los textos proceden del mismo registro de lengua: prosa informativa –escrita sin intenciones literarias o preocupaciones estéticas y por lo tanto carente de efectos especiales– de 39 autores consagrados: ensayistas, novelistas, poetas, eruditos del siglo XX –de Ortega y Gasset (1922) a García Márquez (1988)– de origen español (25 autores) o americano (14).

El análisis del corpus proporciona colecciones de ejemplos de construcciones morfosintácticas interpretadas, acompañadas de índices de frecuencia comparables entre sí. El análisis de 19 textos muestra, por ejemplo, 1623 verbos conjugados pronominalmente en las tres personas. Se construyen de maneras diversas e incluso divergentes y con frecuencias muy dispares tanto por lo que respecta a la extensión del corpus como entre ellos mismos. En el corpus la frecuencia de aparición de las construcciones examinadas oscila entre 1348 y dos por cien mil ocurrencias. El 89 por ciento están construidos con un grupo nominal único en función de regente y el 3 por ciento en función de régimen directo marcado; un siete por ciento no tienen ni regente ni régimen y en un 0,43 por ciento figura un régimen preposicional introducido con *por*, complemento agente. Las construcciones con un grupo nominal único en plural y la forma verbal en singular constituyen el 0,12 por ciento del total.

Incluso si salta a la vista que las construcciones sin regente y régimen marcado, con régimen preposicional, complemento agente, y sin concordancia constituyen una clara minoría y que difieren en frecuencia entre ellas, no deja de ser difícil evaluar el alcance de índices de frecuencia aislados. El común origen de todos los inventarios hace que se puedan clasificar todas las construcciones en una misma escala de frecuencia. La comparación permite evaluar la frecuencia de cada una de ellas, lo que es particularmente interesante cuando se trata de construcciones marginales y polémicas. A continuación figuran quince construcciones en minoría que alternan con una construcción

Frecuencia absoluta (F) y relativa (f) de las construcciones pronominales del verbo, con relación al número de ocurrencias ($f_{/0000}$) y al número total de verbos conjugados pronominalmente (%) por orden de frecuencia decreciente²

19 textos (99.058 occ.)			
	F	$f_{/0000}$	f(%)
+ regente	1446	1460	89,09
- regimen	1335	1348	88,35
v tr.	1233	1245	75,97
	1157	1168	76,57
v. intr.	212	214	13,06
	177	179	11,71
- regente	119	120	7,33
- regimen	119	120	7,88
v. tr.	47	47	2,90
	47	47	3,11
v. intr.	72	73	4,44
	72	73	4,77

contrincante en 39 textos del mismo registro de lengua, sin que se produzca una diferencia de significado.	- regente	50	50	3,08
	+ regimen	50	50	3,31
	+ regente	7	7	0,43
	+ por	7	7	0,46
	- regente	2	2	0,12
	sin concord.	2	2	0,13
Total	1623		100,00	
	1511		100,00	

Le(s) = lo(s), persona masculina³

A Cratilo *le* preocupó el problema de la lengua, (...) CEL 10
 Además es probable que me faltara a mí la habilidad literaria indispensable para referirme a España y a los españoles sacándoles del contexto de (...) AYA 252
 Es la imagen (...) que inspira ese «dilatarse mi corazón» que del salmista toma Teresa de Cepeda, cuando empieza la transformación que *le* llevará a la santidad. ZAM 228
 (...) las preciosas evitaban decir 'ridicule' porque *les* asustaba el sufijo; CEL 343

frente a

Al escritor, al artista, hay que dejarlo en paz. SAL 316
 Y hay que agregar que a los mismos individuos de la ciudad de México *los* solicitan universales de fuera de la ciudad (...) GAO 485

Le(s) = lo(s), la(s), persona femenina y cosa⁴

Pues aunque la vinculación de todo texto con el corpus literario de su tiempo es siempre más intensa que la que *le* une al contexto, (...) GOY 313
 No creo que el hecho de no haber retocado los textos que reimprimo (...) *les* perjudique demasiado, (...) GOM 70
 No significa esto que la misión de la Universidad concluya en formar profesores socialmente útiles: si se *le* quiere reducir a eso, (...) LAZ 233
 (...) al condenar a silencio a las opiniones razonables (...) se *les* equipara (...) a las opiniones irrazonables; PER 145

frente a

(...) los caballeros de aquel país no pueden coger dulcemente del brazo a una dama para ayudarle a cruzar la calle, sino que se ven obligados a agarrarla (asirla fuertemente) CEL 349
 ¿Y qué, si, con achaque de evitar desórdenes, se *los* provoca (...)? UNA 38
 (...) por eso a la cauística los Jesuitas *la* llaman praxis, (...) PER 325

Lo(s) = le(s)

yo a estos chicos *los* cortaría el pelo (...) Ma⁵

frente a

(...) a este señor no *le* caben [estas cosas] en la cabeza. Ma 17.136

La(s) = le(s)

(...) y entonces *la* digo: "Bah" ... mujer, no, no te enfades, (...) Ma 19.555

frente a

No, yo es es que tengo una niña (...) y *le* doy de comer a las cuatro (...) Ma 23.42

Se, régimen indirecto plural, *lo(a)s = se lo(a)*⁶

Bueno, entonces *se los* comentamos a los viejos pausadamente ¿no? BA 22.1

frente a

(...) vende a los turistas. Ella *se lo* sabe explicar en inglés y en francés (...) BA 29.74

*Le = les*⁷:

Dar *le* voz y nombre a quienes no los tienen. FUE 90

frente a

De este modo *les* quitan a los cubanos de la calle una ocasión de expresarse ante el mundo. GAR 186

Pronombre relativo, régimen directo, reduplicado⁸

Este último es un término taurino, que, como acabo de aprenderlo (...), quiero lucir *lo*. UNA 148

(...) como alguien *a quien le* obligasen a mirar por el ojo de una cerradura. FER 144

frente a

Este término, *que* usan todavía en sus ingenuos escritos nuestros casticistas, no significa (...) ORT 297

(...) no deseaba muy instruidos *a quienes* sólo quería esclavos. REY 95

Verbo conjugado pronominalmente sin concordancia⁹

El juicio reprobatorio es unánime en los escritores de Hispanoamérica que tienen conciencia de que *ha de echarse raíces* en la propia tierra y no *zarcillos* en el aire. MAR 116

frente a

(...) en este congreso *habrán de abordarse cuestiones análogas*. CAR 54

Verbo conjugado pronominalmente con régimen preposicional con *por*, complemento agente¹⁰:

(...) no sería seguro que *se leyese* todo lo que se publica *por todos aquellos* para quienes se publica (...) GAO 291

frente a

(...) que alguien (...) llegue a decir “digno y altivo” referido a una persona es algo tan inefable que bien merecería *ser recogido* por todos los humoristas españoles, (...) FER 33

Verbo conjugado pronominalmente con régimen, pero sin regente¹¹

(...) se les toman sus héroes de guillotina y horca, *se los aliña* y *se los transforma* en máscaras heroicas. MAR 444, 445

Yo no entiendo como *se puede llamar* reconquista *a una cosa* que duró ocho siglos.

ORT 374

frente a

se corteja a la mujer, se sonríe a la flor, *se trucida el enemigo* y se goza del bosque. ORT 439,440

Estos modelos abstractos se llaman ‘fonemas’; GIL 43

-ra = -ría, cuatro verbos¹²

Y a falta de una poda sabia y consciente, *podiera* ser bueno el huracán. MAC 204 (“como *debieran* estar todos los cristianos”, piensa el lector que pensaría Fray Luis) CAS 563

Y puesto que estamos aquí entre profesores (...) *quisiera* preguntarles tan solo, (...) GOY 68

debe entenderse que este *apartado hubiera podido* llamarse (...) CEL 682

frente a

Pero no *podría* llamarse ‘lost generation’ a la de estos poetas (...) GUI 268
(...) pero tampoco *debería* extrañarnos demasiado semejante apreciación (...) FER 20
no *querría* ofender a nadie (...) LAZ 161
(...) la incógnita de si la renuncia no *habría matado* la esencial posibilidad de su ser. ALA 83

-ra = indicativo pluscuamperfecto o pretérito perfecto simple¹³

había señalado con dedo acusador a “los delicados que eran hombres y no querían hacer el trabajo de hombres”, y, profético como lo *fuera* muchas veces, vislumbró la abyecta traza de “los desertores (...)”. CAR 167

frente a

Yo conocía a un ilustre hispanista, cuya situación *había sido* más o menos la aquí descrita, (...) RAM 47

*Este, ese, aquel + que*¹⁴

No hay otra tiranía sino *aquella que* se afana y ufana en sofocar la libertad de espíritu. PER 129

(...) no vacila Castro en atribuir la formación *de eso que* hoy entendemos por España a la convivencia de cristianos, moros y judíos (...) AYA 304
sociedades y comunidades (...) dentro de las cuales, *esto que* llamamos “pensamiento” y “comportamiento” tradicional es ante todo: BAR 153

frente a

Creo que ahora se entenderá mejor *lo que* antes he dicho: ORT 536

*Quien(es) + antecedente, regente*¹⁵

(...) es el mismo [consenso] de Azorín y de Maeztu, *quienes* no encontraban ninguna relación entre nuestra literatura y nuestra realidad social. MAR 119

frente a

(...) algunos intelectuales (...) del aparentemente inmaculado equipo Kennedy, *que* mintieron a sabiendas en importantes ocasiones (...) BEN 115

Nombre propio geográfico, régimen directo, sin *a*¹⁶

[los jesuitas que] abarcaron [.] Michoacán y Guadalajara, Puebla y Veracruz, (...) REY 131

frente a

Pero desde comienzos del siglo XIX se definen (...) zonas rivales: la germánica (...); la inglesa, que abarca *a* Inglaterra (...) y *a* los Estados Unidos; HEN 141

Prosa informativa. Frecuencia absoluta (F) y relativa (f) en relación con el número de ocurrencias (⁰/0000), por orden decreciente según la frecuencia relativa

	corpus	F	f (⁰ /0000)
verbo pron. + régimen, - regente	39 textos, 197.886 oc.	87	44
<i>le(s) = lo(s)</i> , persona masculina	39 textos, 197.886 oc.	86	43
<i>-ra = -ría</i> , 4 verbos	39 textos, 197.886 oc.	78	39
<i>aque, l ese, este + que</i>	39 textos, 197.886 oc.	48	24
<i>le(s) = lo(s)</i> , <i>la(s)</i> , persona femenina y cosa	39 textos, 197.886 oc.	29	15
<i>-ra = ind. pluscuamperf., pret. perf. simple</i>	39 textos, 197.886 oc.	15	8
verbo pron. + <i>por</i> , compl. agente	39 textos, 197.886 oc.	11	6
pron. rel., régimen directo, reduplicado	39 textos, 197.886 oc.	10	5
<i>quien + antecedente</i> , regente	39 textos, 197.886 oc.	6	3
nombre propio geográfico, rég. directo, - <i>a</i>	39 textos, 197.886 oc.	6	3
<i>le = les</i>	39 textos, 197.886 oc.	3	2
verbo pron. sin concordancia	39 textos, 197.886 oc.	2	1
<i>lo(s) = le(s)</i>	39 textos, 197.886 oc.	0	0
<i>la(s) = le(s)</i>	39 textos, 197.886 oc.	0	0
<i>se</i> , régimen indirecto plural, <i>lo(a)s = se lo(a)</i>	39 textos, 197.886 oc.	0	0

Puede ocurrir, sin embargo, que una construcción poco frecuente en sí, alcance, en cambio, una proporción relativamente elevada del conjunto de las construcciones que compiten con ella¹⁷. Por ejemplo, los nombres propios de lugar, régimen directo, sin *a* constituyen, pese a todo, el 16 por ciento de los nombres en tal función. *Le(s)* en función de régimen directo para una persona femenina o una cosa, y *-ra* con valor de indicativo pasado, al contrario, solo se producen, respectivamente, en tres y uno por ciento de los casos, lo que les hace retroceder considerablemente en la escala de frecuencia relativa. La mayoría de las construcciones de frecuencia baja en relación con la extensión del corpus son, sin embargo, igualmente poco frecuentes en relación con la construcción rival¹⁸.

Prosa informativa. Frecuencia absoluta (F) y relativa (f) en relación con el número de oportunidades (%), por orden decreciente según la frecuencia relativa

	corpus	F total	F	f (%)
-ra = -ría / -ría, 4 verbos	39 textos	229	77	33,62
le(s) = lo(s) / lo(s), persona masc.	39 textos	263	86	32,70
nombre propio geográfico, rég. directo, - a /+ a	39 textos	37	6	16,22
verbo pron. + régimen, - regente / + regente, de pasiva	19 textos	899	50	5,56
verbo pron. + por / adjetivo verbal + por, complemento agente	39 textos	201	11	5,47
aquel, ese, este + que / el que	19 textos	477	19	4,21
le(s) = lo(s), la(s) / lo(s), la(s), persona femenina y cosa*	39 textos	982	29	2,95
le = les / les	39 textos	147	3	2,04
quien + antecedente, regente / que (280), el cual (1) + antecedente humano	19 textos	281	5	1,78
verbo pron. sin concordancia / con concordancia en plural	39 textos	129	2	1,55
pron. rel., régimen directo, reduplicado / pron. rel. régimen directo	39 textos	843	10	1,19
-ra = ind. pluscuamperf., pret. perf. simple / ind. pluscuamperf. (145), pret. perf. s.(1502)	39 textos	1662	15	0,90
lo(s) = le(s) / le(s)	19 textos	181	0	0
la(s) = le(s) / le(s)	19 textos	39	0	0
se, régimen indirecto plural, lo(a)s = se lo(a) / se lo(a)	39 textos	10	0	0

* Al contrario de lo que se ha hecho en *GE.El*, t. II.6, *Norma, libertad y probabilidad: El leísmo. Una confusión opcional*, el cálculo de *lo* para una cosa incluye aquí los pronombres neutros ('generalizadores') en función de régimen directo.

Nos consta que en el mismo corpus otras construcciones optativas, plenamente admitidas y que nunca se ponen en tela de juicio, alcanzan una frecuencia similar a aquella con complemento agente y cuyo verbo no tiene regente pero sí régimen y que otras, que no se recomiendan o juzgadas poco recomendables, se colocan a distancia análoga a la que, con respecto a la más frecuente, se encuentra la fórmula pronominal sin concordancia.

Se ve que la construcción sin concordancia, con un 1,55 por ciento de las oportunidades, se encuentra entre las menos frecuentes, las más discutidas y las menos aceptadas las que más controversias suscitan: *se lo(a)s* para *se lo(a)*, *la(s)* para *le(s)*, *lo(s)* para *le(s)*, ausentes, *-ra* con valor de indicativo pasado (0,90), la reduplicación de un pronombre relativo, régimen directo, en el núcleo verbal (1,19), *quien(es)* con antecedente en función de regente de la oración relativa (1,78), *le* para *les* (2,04), *le(s)* para *lo(s)*, *la(s)*, cosa o persona femenina (2,95) La construcción que lleva un régimen pero sin regente

(5,29 por cien de las oportunidades) y la construcción pronominal con complemento agente (5,47), al contrario, forman parte de las más frecuentes, las que se admiten sin discusión: *-ra* para *-ría* con *haber* más adjetivo verbal, *deber*, *poder*, *querer* (33,62), *le(s)* para *lo(s)*, persona masculina (32,70), la ausencia de *a* delante de un nombre propio geográfico en función de régimen directo (16,22), pronombres relativos compuestos con una forma demostrativa (5,29).

Salta a la vista que las construcciones menos frecuentes son las únicas que se ponen en tela de juicio en la práctica, aunque sea desordenadamente. En efecto, hay coincidencia entre frecuencia baja y duda en cuanto a la gramaticalidad de cada una de ellas. El criterio de frecuencia relativa permite solucionar con uniformidad las estimaciones de gramaticalidad, a menudo variables según la construcción y más de una vez contradictorias dentro de la misma, al estar fundadas en argumentos heterogéneos. Es una pauta de (a)gramaticalidad en el registro de lengua representado.

* * *

Es un hecho, sin embargo, que la frecuencia de aparición de muchas construcciones varía de un registro de lengua a otro. La ausencia de concordancia entre un verbo conjugado pronominalmente y el grupo nominal, muy escaso en la lengua escrita (el dos por ciento de las oportunidades) es siempre más frecuente en la lengua hablada: oscila, en el habla culta de 12 capitales, entre un 7 y un 35 por ciento, o sea un promedio del 13 por ciento¹⁹

Verbos conjugados pronominalmente en singular con grupo nominal en plural

	total F	con concordancia F	f(%)	sin concordancia F	f(%)
Prosa informativa					
39 Textos	129	127	98	2	2
Habla culta					
Bogotá	166	150	90	16	10
Buenos Aires	106	96	91	10	9
Caracas	76	66	87	10	13
La Habana	81	67	83	14	17
La Paz	112	73	65	39	35
Lima	38	29	76	9	24
Madrid	41	35	85	6	15
México	80	73	91	7	9
San Juan	141	133	94	8	6
Santiago	209	184	90	21	10
Sevilla	46	43	93	3	7
total	1092	949	87	143	13

Mientras que en la lengua escrita un nombre propio geográfico en función de régimen directo está introducido por *a* en un 84 por ciento de casos, en el habla culta hay un promedio de 72 por ciento sin preposición²⁰.

Nombres propios geográficos en función de régimen directo con y sin preposición

	+ a		- a	
	F	f	F	f
Prosa informativa				
39 textos	31	83,78	6	16,22
Habla culta				
Bogotá	10	59	7	41
Buenos Aires	3	12	23	88
Caracas	5	31	11	69
La Habana	9	50	9	50
La Paz	10	63	6	37
Lima	3	7	38	93
Madrid	5	38	8	62
México	6	29	15	71
San Juan de Puerto Rico	3	43	4	57
Santiago de Chile	3	14	25	86
Sevilla	3	25	9	75
Total	60	28	155	72

A las diferencias entre registros de lengua se vienen a añadir las que median entre áreas geográficas. El *leísmo* es, como se sabe, propio de la península Ibérica: alcanza un 68 por ciento de las oportunidades referido a un ser animado, masculino singular, en la prosa informativa, mientras que es inhabitual en América: el 15 por ciento en iguales condiciones; un 12 por ciento frente a un cuatro por ciento en general²¹.

Al contrario, *le* por *les*, que no aparece en los textos de autores de origen español, se produce hasta en un siete por ciento de las oportunidades en los escritores americanos²².

Formas etimológicas y no etimológicas en función de régimen directo

	prosa informativa peninsular (25 t., 122.266 oc.)				prosa informativa americana (14 t., 75.620 oc.)			
	f. etimológica		forma no etim.		f. etimológica		forma no etim.	
	F	f	F	f	F	f	F	f
persona								
masc. sing.	27	31,76	58	68,24	83	84,69	15	15,31
plural	38	76,00	12	24,00	27	96,43	1	3,57
fem. sing.	22	91,67	2	8,33	2	100,00	0	0,00
plural	4	57,14	3	42,86	4	100,00	0	0,00
no persona								
masc. sing.	283	95,61	13	4,39	2549	98,28	3	1,72
plural	50	98,04	1	1,96	211	100,00	0	0,00
fem. sing.	189	97,42	5	2,58	773	99,05	1	0,96
plural	67	98,53	1	1,47	245	100,00	0	0,00
total	680	87,74	95	12,26	448	95,73	20	4,27

Le = les

prosa informativa peninsular (25 t., 122.266 oc.)				prosa informativa americana (14 t., 75.620 oc.)			
f. etimológica		forma no etim.		f. etimológica		forma no etim.	
F	f	F	f	F	f	F	f
106	100,00	0	0,00	38	92,68	3	7,32

Si se compara la frecuencia relativa de las quince construcciones estudiadas en la prosa informativa con la del habla culta de Madrid²³ y de Buenos Aires²⁴, por ejemplo, se observan numerosas y grandes diferencias. Además de las divergencias ya mencionadas salta a la vista la frecuencia elevada de *se* (régimen indirecto plural) *lo(a)s* por *se lo(a)*: un 68 por ciento en Buenos Aires, ausente de la península Ibérica tanto en el registro hablado como en el escrito, la frecuencia muy baja (en Buenos Aires) y la ausencia (en Madrid) de la construcción pronominal sin régimen pero con regente, que alcanza el 5 por ciento en la prosa informativa, al contrario de lo que ocurre con la construcción sin concordancia, esporádica en los textos (un dos por ciento) y corriente en el habla (un 9 por ciento en Buenos Aires y un 15 en Madrid).

Habla culta de Madrid. Frecuencia absoluta (F) y relativa (f) en relación con el número de oportunidades (%), por orden decreciente según la frecuencia relativa

	corpus	F total	F	f (%)
nombre propio geográfico, rég. directo, - a / + a	24 enc.	13	8	61,54
<i>le(s) = lo(s) / lo(s), persona</i>	110 enc.	778	459	59,00
<i>-ra = -ría / -ría, 4 verbos</i>	24 enc.	49	24	48,98
<i>la(s) = le(s) / le(s)</i>	110 enc.	470	77	16,38
verbo pron. sin concordancia / con concordancia en plural	24 enc.	41	6	14,63
pron. rel., régimen directo, reduplicado / pron. rel. régimen directo	8 enc.	74	10	13,51
<i>aquel ese, este + que / el que</i>	8 enc.	232	14	6,03
<i>le(s) = lo(s), la(s) / lo(s), la(s), persona femenina y cosa</i>	110 enc.	4213	182	4,32
verbo pron. + régimen, - regente / + regente, de pasiva	8 enc.	104	1	0,96
<i>lo(s) = le(s) / le(s)</i>	110 enc.	1596	11	0,69
<i>-ra = ind. pluscuamperf., pret. perf. simple / ind. pluscuamperf. (42), pret. perf. s.(648)</i>	24 enc.	679	0	0,00
<i>quien + antecedente, regente / que (74), el cual (0) + antecedente humano</i>	8 enc.	74	0	0,00
<i>le = les / les</i>	8 enc.	30	0	0,00
<i>se, régimen indirecto plural, lo(a)s = se lo(a) / se lo(a)</i>	24 enc.	6	0	0,00
verbo pron. + <i>por</i> / adjetivo verbal + <i>por</i> , complemento de agente	8 enc.	2	0	0,00

Habla culta de Buenos Aires. Frecuencia absoluta (F) y relativa (f) en relación con el número de oportunidades (%), por orden decreciente según la frecuencia relativa

	corpus	F total	F	f (%)
nombre propio geográfico, rég. directo, - a / + a	33 enc.	26	23	88,46
se, régimen indirecto plural, lo(a)s = se lo(a) / se lo(a)	33 enc.	15	10	66,67
-ra = -ría / -ría, 4 verbos	33 enc.	171	60	35,09
le = les / les	7 enc.	105	16	15,24
verbo pron. sin concordancia / con concordancia en plural	33 enc.	106	10	9,43
le(s) = lo(s) / lo(s), persona masculina	7 enc.	153	10	6,54
aquel, ese, este + que / el que	7 enc.	359	22	6,13
pron. rel., régimen directo, reduplicado / pron. rel. régimen directo	7 enc.	336	16	4,76
le(s) = lo(s), la(s) / lo(s), la(s), persona femenina y cosa	7 enc.	721	24	3,33
verbo pron. + régimen, - regente / + regente, de pasiva	7 enc.	349	1	0,86
lo(s) = le(s) / le(s)	7 enc.	374	1	0,28
verbo pron. + por / adjetivo verbal + por, complemento de agente	7 enc.	7	0	0,00
la(s) = le(s) / le(s)	7 enc.	239	0	0,00
quien + antecedente, regente / que (74), el cual (0) + antecedente humano	7 enc.	225	0	0,00
-ra = ind. pluscuamperf., pret. perf. simple / ind. pluscuamperf., pret. perf. s.	7 enc.	1246	0	0,00

* * *

Censurar o recomendar, tolerar o preferir una u otra construcción se apoya en motivos muy diversos, no siempre coherentes, a menudo basados en prejuicios, análisis divergentes, razonamientos ajenos al funcionamiento de la lengua, rechazos instintivos, una documentación mal calibrada. No son en absoluto ilegítimos cuando se trata de la práctica individual, pero la cuestión de su "legitimidad" se plantea cuando el objetivo es estudiar la lengua.

No cabe duda de que para una mente amante del orden el recurso exclusivo a *lo(s)*, *la(s)* en función de régimen directo y a *le(s)* para el indirecto constituye la mejor solución. Dado el elevado número de pronombres relativos puede parecer irrazonable añadir otros más, sobre todo si, a primera vista, son préstamos de otras lenguas. Si se reflexiona únicamente en términos semánticos, introducir un nombre propio geográfico en función de régimen directo mediante una preposición puede ser una incongruencia. Cuando se considera que la redundancia en la comunicación, lejos de ser natural, es un mal, la reduplicación de un pronombre relativo en función de régimen tiene que resultar excesiva.

De análisis gramaticales divergentes surgen evaluaciones de gramaticalidad dispares. Quien considere «activa» la construcción pronominal sin concordancia entre el verbo y el único grupo nominal en plural y otorgue a *se* la función de regente, no puede por menos de encontrar normal la ausencia de concordancia y absurda la aparición de un complemento agente; lo contrario ocurrirá para el que, atribuyendo a *se* un estatuto de morfema auxiliar de conjugación, vea en la construcción un carácter esencialmente pasivo. También existe una construcción cuyo único grupo nominal está introducido por la preposición *a*, marca inequívoca de un régimen directo, o constituido por un pronombre personal régimen. Los partidarios de la construcción sin concordancia hallan aquí la confirmación de la interpretación «activa», pero anteponer una preposición ante un nombre de cosa también puede, a su vez, dar lugar a dudas; para sus adversarios la inserción de la preposición invalida la falta de concordancia, incluso si la presencia de un régimen directo pone en entredicho la interpretación “pasiva”.

El desdén hacia la lengua hablada o una actitud conservadora o inmovilista pueden incitar a rechazar todo lo que sea característico de esta: las construcciones pronominales sin concordancia, el loísmo y el láismo, *se lo(a)s* por *se lo(a)*, los pronombres relativos régimen reduplicados, los nombres de lugares sin preposición en función de régimen directo. A la inversa, estar en contra de la autoridad, el gusto de la novedad, la demagogia, las ganas de llevar la contraria, pueden arrastrar a promocionar lo oral enfrentándose con la norma académica.

El que sea de origen peninsular puede tener tendencia a pasar por alto las construcciones propias del continente americano: *le* por *les*, *se lo(a)s* por *se lo(a)*, mientras que el de origen americano quizás no admita los leísmos, en general o los que no se atienen a las recomendaciones de la Academia, y que son característicos de la Península ibérica.

En el extranjero se tiene tendencia a comentar prolijamente algunas de las construcciones propias del español, lo que les otorga una especie de estatuto especial que, a su vez, les confiere implícitamente una gramaticalidad acrecentada o superior a la que se les concede por lo general en el país de origen. Es lo que ocurre con las construcciones pronominales de verbo sin concordancia, que gozan de afición particular, el empleo automático y sin tasa de *quien* para personas, con o sin antecedente, en función de regente o de régimen, olvidando los otros pronombres, y el recurso a *le(s)* en lugar de *lo(s)* que pasa de ser optativo a preferible cuando no se considera obligatorio.

El conocimiento y el respeto del pasado pueden conducir a pensar que las soluciones antiguas son más gramaticales que otras más recientes, sin que por ello quede muy claro cuál es el pasado ideal. Se puede leer a veces que lo mejor es utilizar las formas etimológicas a la hora de elegir un pronombre personal régimen. ¿Se puede esgrimir el mismo argumento para justificar el empleo de las formas en *-ra* en lugar de un pluscuamperfecto de indicativo? Y si se hace intervenir el criterio etimológico ¿qué actitud hay que adoptar en el caso de las formas en *-ra* con valor de condicional? Y si, en el caso contrario, se hace abstracción de la etimología y solo se llega hasta el siglo XVII, habría que generalizar el leísmo y desterrar las formas en *-ra* equivalentes a un indicativo.

Los argumentos citados precedentemente no resultan convincentes. No es evidente que la gramática constituya siempre un todo plenamente lógico o razonable, ni que sus utilizadores recurran a ella con pleno conocimiento de causa, de manera sistemática y sin caprichos ni desviaciones, consciente o inconscientemente. Para evaluar la gramaticalidad no basta un análisis basado en consideraciones teóricas o competencias individuales, vengan de quien vengan, sea gramático o utilizador. Habría que saber si hay que

admitir que lo que fue el uso pero ya no lo es, salvo ocasional o parcialmente, sigue siendo gramatical, o que lo que representa el uso en el registro de lengua más prestigiado sea lo gramatical para todos los registros, inclusive si en estos no es habitual, de ninguna manera o en menor grado. ¿Puede pretenderse que lo que es acostumbrado en España sea siempre gramatical allende los mares o al revés?

Pero lo que causa más desazón es que los argumentos esgrimidos son heterogéneos, se invocan desordenadamente, con frecuencia son contradictorios e inaplicables al conjunto de las construcciones examinadas; para cada construcción rige un criterio diferente de (a)gramaticalidad. Ante todo es necesario tener una medida única.

Si se admite que el uso prevalece sobre prescripciones inspiradas de consideraciones ajenas a la lengua, que el uso actual predomina sobre usos anteriores, que tal uso puede no ser el mismo en todos los registros de lengua y todos los lugares, que la frecuencia relativa de una construcción representa un aspecto del uso y que existen grados de (a)gramaticalidad, puede pensarse que la frecuencia relativa de una construcción morfosintáctica constituye un índice de (a)normalidad, (in)aceptabilidad, (i)legitimidad o (a)gramaticalidad.

La frecuencia puede tasarse con precisión a favor de la observación sistemática y exhaustiva de corpus cerrados. Las frecuencias anotadas dentro de un mismo corpus pueden ser ordenadas según una escala, lo que las hace comparables. De ahí que la gramaticalidad de una construcción se calibre según su posición en la escala y la diferencia de gramaticalidad se incremente si la distancia en la escala aumenta. De este modo se llega a una evaluación objetiva, relativa y graduada de la gramaticalidad.

La clasificación de las quince construcciones por orden de frecuencia relativa en cada corpus por separado muestra que dicha escala es muy diferente de un corpus a otro, tanto entre el corpus de lengua escrita y los dos corpus de lengua hablada, como entre los de lengua hablada entre sí. Más de una construcción que se sitúa entre las menos frecuentes de un corpus, o ni siquiera aparece, y por consiguiente no puede considerarse gramatical, se revela frecuente y gramatical en otro.

Toda evaluación de (a)gramaticalidad debería tener en cuenta esta variabilidad, a menos que se tenga por definitivamente agramatical cualquier construcción insuficientemente frecuente en una u otra escala de frecuencia, o gramaticales todas aquellas cuya frecuencia es suficiente en una escala cualquiera, o dar la preferencia a un tipo de lengua determinado y juzgar que, en general, es agramatical toda construcción demasiado poco frecuente en esta escala en particular.

También se podría estimar que, por lo que respecta a la gramaticalidad, una construcción que es frecuente solo en un registro o un área geográfica no puede ser tenida en cuenta, incluso si alcanza una frecuencia elevada, o que la gramaticalidad de una construcción aumenta cuando se reproduce con una frecuencia similar en corpus de origen diverso. No obstante, de las quince construcciones que se han examinado solo el empleo de las formas en *-ra* con valor de condicional es relativamente frecuente en los tres corpus explorados. El *leísmo* referido a una persona de género masculino aparece en los tres, pero con índices muy divergentes. La falta de concordancia con un verbo conjugado pronominalmente es propia del registro hablado y el empleo de la forma pronominal con régimen y sin regente lo es del escrito. *Le* por *les* así como *se lo(a)s* por *se lo(a)*, características del continente americano, no aparecen en la península Ibérica; el *laísmo* se limita al habla culta de Madrid.

En ningún caso es lícito recurrir a escalas diferentes para aceptar o rechazar una construcción y no otra, afín de satisfacer convicciones o conveniencias personales, teóricas o políticas.

Intentar determinar en la escala de frecuencia un umbral que delimite las construcciones gramaticales y agramaticales es una idea tentadora. Sin embargo, no será posible establecerlo hasta que no se haya descrito la totalidad de las construcciones que constituyen el corpus. Cabe pensar que tal umbral –si es que existe– es móvil: probablemente será más elevado en el registro escrito –más exigente– que en el hablado –más laxista– y diferente en regiones con gran tradición o de desarrollo reciente, por ejemplo, ya que depende tanto del que interpreta la escala de frecuencia como de la organización económica del código lingüístico. Si se desea que sea un instrumento gramatical productivo, una construcción debe alcanzar cierta frecuencia de empleo para que el usuario la pueda adoptar es necesario que haya sido documentada reiteradamente. La convergencia de estas dos exigencias es lo que dicta el límite a partir del cual una construcción funciona económicamente o se gramaticaliza.

Se puede suponer que en cada escala de frecuencia, a altura variable, existe una zona de incertidumbre, a su vez fluctuante, más o menos extensa e indefinida según las colectividades, instituciones, clases sociales, los individuos, su pertenencia social, educación, etc., de continuo sometida a perturbaciones desordenadas y contradictorias por el azar, las decisiones humanas y las gramáticas vecinas o dominantes. La influencia recíproca de los registros, de las regiones, de los géneros²⁵, etc. es probablemente el factor perturbador más importante y quizás el más solapado. La mezcla voluntaria o inconsciente, sutil en unos casos y burda en otros, de rasgos morfosintácticos de origen distinto es permanente. Los escritores la cultivan –es uno de los objetos de la estilística²⁶–, mientras que la gramática normativa intenta excluirla o canalizarla imponiéndole límites.

De todos modos, la gramática normativa es, en fin de cuentas, la única a la que concierne la cuestión de saber si una construcción es gramatical o no. La lingüística que se limite a tener en cuenta únicamente las construcciones verificadas en la producción espontánea de otros –consciente o inconsciente y que no se haya provocado intencionalmente con objetivos de estudio gramatical– no tiene por qué plantearse el dilema. Le basta con tomar en consideración la frecuencia relativa de cada construcción en condiciones determinadas.

Es cierto que la gramática normativa tiene la obligación de formular reglas que se ajusten al máximo a la realidad, pero sin olvidar el aspecto práctico: deben ser fáciles de comprender, asimilar y aplicar por la gran mayoría de los utilizadores, tanto extranjeros como autóctonos. Con este fin y en grado diverso, de acuerdo con el destinatario, el tipo de lengua y el nivel de conocimiento deseado, tiene que escoger, reducir y simplificar. No le queda más remedio que descartar las alternativas en minoría o proponer como modelo el uso de una comunidad restringida.

Cada cual, y no solo el gramático, con más o menos razones y éxito, podría elegir una escala de frecuencia relativa y asignarse un umbral de gramaticalidad que le permitiese comportarse –admitir o rechazar tal o cual construcción– de manera uniforme, sistemática y objetiva. No quiere decir esto que cualquiera deba aprovechar esta libertad, pero en todo intento normativo siempre se debería estar convencido de la necesidad como también de la relatividad de un punto de referencia de este tipo.

Precisemos que sugerir que una gramática determinada, la de la prosa informativa, por ejemplo, o cualquier otra gramática, constituya el modelo, no va en menoscabo de las demás. Recuérdese asimismo que la gramática es el resultado de convenciones colectivas que cada cual puede rechazar en todo momento y con cualquier motivo, de manera ocasional o sistemática, sean cuales sean los riesgos de ser llamado al orden.

Ni que decir tiene que las construcciones examinadas son solo el embrión de una lista ideal más extensa en la que se recogerían todas aquellas construcciones que se ponen en duda, o podrían ponerse en duda, o que se admiten comúnmente. Pero es posible, desde ya, confrontar con la lista existente cualquier construcción cuya gramaticalidad se pone en tela de juicio, aunque solo sea para sondear la eficacia de dicha lista y, de paso, enriquecerla, puesto que la (a)gramaticalidad no se define según criterios ajenos al código, heterogéneos y subjetivos –teóricos, lógicos, sociales, políticos, didácticos, prácticos, personales, etc.–, sino que emana de un criterio único, formalmente identificable, que se puede definir con precisión y se desprende de la observación de la lengua en marcha, y por lo tanto aplicable a todos sus productos.

Apéndice

Gramática española: Enseñanza e investigación, publicada por Ediciones Universidad de Salamanca, Salamanca, desde 1990, se compone de cinco partes: I, *Apuntes metodológicos*, II, *Gramática*, III, *Textos*, IV, *Índices*, V, *Concordancias*. Hasta la fecha han aparecido:

- I.1. J. DE KOCK, *Apuntes metodológicos*, 1, 1990, 156 págs.
- I.2. J. DE KOCK, *Enseñanza*, 1994, 223 págs.
- I.3. J. DE KOCK, C. GÓMEZ MOLINA, N. DELBECQUE, F. MORENO FERNÁNDEZ, J. DE BAERE, *De la relatividad en lingüística*, 1995, 263 págs.
- I.4. J. DE KOCK, N. DELBECQUE, L. MELIS, J.L. RODRÍGUEZ GÓMEZ y E. ROEGUEST, *Lingüística contrastiva*, 1996, 211 págs.
- I.4. J. DE KOCK, N. DELBECQUE, L. MELIS, J.L. RODRÍGUEZ GÓMEZ y E. ROEGUEST, *Lingüística contrastiva*, 1996, 211 págs.
- I.5. J. DE KOCK y G. DEMELLO, *Lengua escrita y habla culta en América y España*
- I.6. R. CARAVEDO, *Lingüística del corpus. Cuestiones teórico-metodológicas aplicadas al español*, 1999, 292 págs.
- I.7. J. DE KOCK (Ed.), *Lingüística con corpus. Catorce aplicaciones sobre el español*, 2001, 404 págs.
- II.1. J. DE KOCK, C. GÓMEZ MOLINA y N. DELBECQUE, *Gramática didáctica*, 1990, 188 págs.
- II.2. J. DE KOCK y C. GÓMEZ MOLINA, *Las formas pronominales del verbo y la pasiva*, 1990, 140 págs.
- II.3. J. DE KOCK, *Del pretérito perfecto compuesto o de la importancia del contexto y de la cuantificación. La noción de auxiliaridad*, 1990, 168 págs.
- II.4. N. DELBECQUE, *El orden de los sintagmas: la posición del regente*, 1991, 319 págs.
- II.5. J. DE KOCK, C. GÓMEZ MOLINA y R. VERDONK, *Los pronombres demostrativos y relativos*, 1992, 184 págs.
- II.6. J. DE KOCK y C. GÓMEZ MOLINA, *Norma, libertad y probabilidad. Ocho soluciones del español*

- II.7. G. DE STERCK, *Registros y áreas geográficas en lingüística. Valores y usos de las formas verbales en -ra, -se, -ría y -re*, 2000, 382 págs.
- III.1. J. DE KOCK, R. VERDONK y C. GÓMEZ MOLINA, *19 textos*, 1ª reimpresión, 1996, 280 págs.; edición electrónica, 1996, 1 disquete.
- III.2. J. DE KOCK, C. GÓMEZ MOLINA, P. GARCÍA MOUTON y N. DELBECQUE, *20 textos*, 1992, 276 págs.
- IV.1. J. DE KOCK, *Índice alfabético, alfabético inverso y de frecuencia de 19 textos*, 1991, 499 págs.
- IV.2. J. DE KOCK, *Índice alfabético, alfabético inverso y de frecuencia de 20 textos*, 1992, 526 págs.
- V.1. J. DE KOCK, *Concordancia alfabética de 19 textos*, 1990, 1964 págs.
- V.2. J. DE KOCK, *Concordancia alfabética de 20 textos*, 1990, 1927 págs.

De próxima aparición:

- I.8. J. DE KOCK y C. GÓMEZ MOLINA, *Lingüística aplicada. La lengua: meta, materia y referencia en investigación, enseñanza y estilística*, 548 págs. + ilustración en Internet, en prensa

¹ *Gramática española. Enseñanza e investigación* (de aquí en adelante *GE.EI*), Salamanca, Ediciones Universidad Salamanca, 1990 ss. t. III.1, *19 textos*, 1.ª reimpresión + disquete, 1996, y t. III.2, *20 textos*, 1992.

² Puesto que las variaciones de construcción gramatical de las formas pronominales del verbo no son en absoluto las mismas en la tercera persona y, por otro lado en la primera y la segunda, y que la frecuencia relativa de las tres personas puede variar notablemente de un registro a otro, se dan, para facilitar la comparación con otros corpus, las cifras para las tres personas y las que se refieren a la tercera persona (en negritas y bastardillas) por separado.

³ Para más detalles véase *GE.EI*, t. II.6, *Norma, libertad y probabilidad: El leísmo. Una confusión opcional*.

⁴ Véase *idem*.

⁵ Citado por A. Quilis, M. Cantarero, M. J. Albalá y R. Guerra, *Los pronombres le, la, lo y sus plurales en la lengua española hablada en Madrid*, Madrid, 1985, pág. 58.

⁶ Véase *GE.EI*, t. I.5, *Lengua escrita y habla culta en América y España: Se lo(a)s por se lo(a)*.

⁷ Véase *idem*: *Le por les*.

⁸ Véanse *idem*, *La reduplicación del pronombre relativo, régimen directo*, y t. II.6, *Norma, libertad y probabilidad: La reduplicación instrumental del régimen directo e indirecto*.

⁹ Véanse *GE.EI*, t. I.8, *Lingüística aplicada, La lengua, materia de enseñanza: Las construcciones pronominales del verbo y la pasiva*, *En 39 textos: Tres construcciones marginales*, t. I.5, *Lengua escrita y habla culta en América y España: Construcciones pronominales del verbo sin concordancia*, t. I.3, *De la relatividad en lingüística: La pasiva en 19 textos, artículos de periódico de Miguel de Unamuno y lengua hablada*.

¹⁰ Véanse *GE.EI*, t. I.8, *Lingüística aplicada, La lengua, materia de enseñanza: Las construcciones pronominales del verbo y la pasiva*, *En 39 textos: Tres construcciones marginales*, t. I.5, *Lengua escrita y habla culta en América y España: Verbo pronominal con por + agente*.

¹¹ Véase *GE.EI*, t. I.8, *Lingüística aplicada, La lengua, materia de enseñanza: Las construcciones pronominales del verbo y la pasiva*, *En 39 textos: Tres construcciones marginales*.

¹² Véanse *GE.EI*, t. II.7, *Registros y áreas geográficas en lingüística*, § II.2, *Las formas en -ra, -se y -ría en contextos de subjuntivo*, § II.3, *en contextos de indicativo con valor de condicional*, § II.4, t. I.5, *Lengua escrita y habla culta en América y España: Formas verbales en -ra/-se con valor de condicional*, t. I.2, *Enseñanza: Índices y concordancias en la clase de gramática*, § 2, *La competencia entre el subjuntivo en -ra y el condicional*.

¹³ Véase *GE.EI*, t. II.7, *Registros y áreas geográficas en lingüística*, § 4, *Las formas en -ra y -se en contextos de indicativo (con valor de pasado)*.

¹⁴ Véanse, para *que*, *quien*, *el cual*, *cuyo* y *cuanto*, *GE.EI*, t. II.5, *Los pronombres demostrativos y relativos: Los pronombres relativos...*; para *este*, *ese* y *aquel*, *idem* y *Los pronombres demostrativos...*, § 2.2.2.3, t. II.6, *Norma, libertad y probabilidad. Ocho soluciones del español: Este, ese y aquel, pronombres. De los esquemas teóricos y de la realidad*, así como t. I.3, *De la relatividad en lingüística*, § II, *La relatividad gramatical en registros y áreas geográficas diferentes*, § 3, *Los pronombres relativos en 19 textos, lengua hablada y prensa peninsular*.

¹⁵ Véanse *GE.EI*, t. I.8, *Lingüística aplicada, La lengua, objeto de investigación: Variaciones colectivas e individuales, espontáneas e intencionales: quien(es)*, y *Divergencias y convergencias en dos registros de lengua diferentes*, § 2.

¹⁶ Véase *GE.EI*, t. I.5, *Lengua escrita y habla culta en América y España: El régimen directo preposicional: los nombres propios geográficos*.

¹⁷ La comparación no está exenta de escollos, porque no siempre es evidente cuál es la construcción competidora, ni si las condiciones de empleo son exactamente las mismas. Incluso si no se pueden excluir todas las dudas según el contexto y la situación, no cae de su peso que a construcciones diferentes correspondan automáticamente significados, intenciones o efectos de sentido distintos, aún de manera inconsciente. Las interpretaciones posteriores y la justificación de elecciones ajenas no constituyen pruebas suficientes.

¹⁸ Conviene tener presente que el número de ocurrencias examinado en el mismo corpus puede variar de una construcción a otra, y según se trate de la frecuencia relativa en relación con el número de ocurrencias o con el número de oportunidades para la misma construcción. Respecto a la prosa informativa puede tratarse de 39 o de 19 textos, respectivamente 197.886 o 99.058 ocurrencias.

¹⁹ Véase *GE.EI*, t. I.5, *Lengua escrita y habla culta en América y España: Construcciones pronominales del verbo sin concordancia*.

²⁰ Véase *idem*, *El régimen directo preposicional: los nombres propios geográficos*.

²¹ Véase *GE.EI*, t. I.8, *Lingüística aplicada, La lengua, objeto de la investigación: Divergencias y convergencias en dos corpus de registros de lengua diferentes*, § 3.

²² Véase *GE.EI*, t. I.5, *Lengua escrita y habla culta en América y España: Le por les*. Recordemos, sin embargo, que hay menos diferencias entre continentes en el registro escrito de las que existen entre registros, incluso dentro de un mismo continente. En la prosa informativa solo son importante en cuanto al leísmo y el empleo de *le* por *les*. La distancia entre el empleo de formas en *-ra* con valor de condicional (un 39 por ciento de las posibilidades en España frente a un 32 por ciento en los textos americanos) o con valor de pretérito de indicativo (un 0,72 por ciento en España, un 1,03 por ciento en los autores americanos) es demasiado pequeña e insegura para que se pueda subrayar.

²³ Según M. Esgueva y M. Cantarero (Eds.), *El habla de la ciudad de Madrid*, Madrid, 1981. Téngase en cuenta que el número de ocurrencias explorado puede variar de una construcción a otra. Puede tratarse de todas las encuestas grabadas (110: 540.998 ocurrencias), de las encuestas publicadas (24: 141.750 ocurrencias) o de las encuestas en forma de diálogo (8: 44.339 ocurrencias).

²⁴ Según A.M. Barrenechea (Ed.), *El habla culta de la ciudad de Buenos Aires*, Buenos Aires, 1987. Téngase en cuenta que el número de ocurrencias explorado puede variar de una construcción a otra. Puede tratarse de todas las encuestas publicadas (245.463 ocurrencias) o de siete (21 a 24, 27 a 29: 75.256 ocurrencias).

²⁵ En este caso no se han tenido en cuenta.

²⁶ Véanse *GE.EI*, t. I.8, *Lingüística aplicada, La lengua, punto de referencia en estilística: Lengua escrita con gramática hablada: el caso Unamuno*, y *Variaciones colectivas e individuales, espontáneas e intencionales: quien(es)*.