

FRECUENCIA E INSTRUMENTALIDAD

Ya hemos sugerido anteriormente¹ que se defina la auxiliaridad en términos de frecuencia relativa. Si se considera que un verbo auxiliar es una forma instrumental, tendrá en común con las demás formas instrumentales cierto número de características y entre ellas una frecuencia relativamente elevada. Hemos admitido que las formas instrumentales de una lengua son frecuentes y que las formas más frecuentes son instrumentales. En las páginas que siguen esta aserción se somete a un primer examen.

El examen es sólo posible en un corpus cerrado; en nuestro caso se trata del mismo que ha servido en los precedentes análisis sobre la auxiliaridad: un corpus de 19 autores españoles y latinoamericanos modernos compuesto de 101.613 ocurrencias.² Aquí se examinan los 204³ lemas⁴ más frecuentes. Son por orden de frecuencia decreciente:

el	11.158	= artículo, al, del, + que (456), + cual (68), + de (171)
de(l)	7.681	
y (e)	3.259	
a(l)	2.628	
en	2.645	
que	1.848	= pronombre

¹ J. De Kock, *Pour une nouvelle définition de la notion d'auxiliarité*, en *La Linguistique*, 11 (1975), 2, 81-92 (también en J. De Kock, *Linguistique automatique et langues romanes*, París, 1977, 163-166), *La notion d'auxiliaire contrôlée (de la fréquence relative à la différence de fréquence relative)* en *Linguistica Antverpiensia*, 13 (1979), 103-108, y *La noción de auxiliaridad. Las construcciones intransitivas con adjetivo verbal*, en *Philologica Hispaniensia in honorem Manuel Alvar*, II (1985), 283-303.

² Se trata del corpus base de J. De Kock, *Gramática española: enseñanza e investigación*, Lovaina, 3a ed. parcial y experimental, 1986, 5 tomos. El corpus de 19 textos constituye el t. 3.

³ Hasta agotarse la frecuencia 45.

⁴ La lematización ha sido efectuada según las normas siguientes. Las formas declinadas y conjugadas han sido reunidas bajo la forma masculina singular y el infinitivo respectivamente. También es así para las formas regente y régimen de los pronombres personales (**yo, me, mi, nosotros y nos**, por ejemplo, figuran bajo **yo**), las formas apocopadas (**cualquier, gran**, etc.) y las formas distintivas para cosa y persona (**algo, alguien**). Se reúnen las formas adjetivas y pronominales (**este y éste**, por ejemplo) y se juntan las formas interrogativas a las no interrogativas correspondientes (**cómo, como**, etc.). No se mencionan separadamente las preposiciones según que precedan a un sustantivo, un infinitivo o la conjunción **que**; tampoco se restan los empleos adverbiales de los adjetivos correspondientes si tienen la misma forma (**solo, tanto, mejor**, etc.). Aún ha sido contado con **aun**. **Al** y **del** han sido contados dos veces, una vez como artículos y otra como preposiciones. Se distingue entre formas verbales y nominales en el caso del adjetivo verbal, y entre formas adjetivas y sustantivas en todos los casos. Las formas que puedan ser tanto instrumentales como léxicas, están reunidas bajo un mismo lema. Siempre se apunta a frecuencia de las formas sumadas. La lematización de los lexemas ha sido efectuada por N. Delbecq con vistas a un estudio de vocabulario, de próxima aparición.

ser	1.763	
un	1.751	= artículo
se	1.602	
que	1.455	= conjunción
no	1.381	
haber	1.277	
él	1.215	= pronombre: él ... (351), lo ... (580), le ... (284)
por	1.096	
su(s)	1.029	
con	845	
o(u)	804	
como	715	+ cómo (40)
más	657	
este	608	+ éste ... (227)
para	584	
yo	574	+ me (246), mi (28), nosotros (35), nos (160)
todo	516	adj. + pron. (126)
poder	445	solo verbo
otro	353	adj. + pron. (121)
pero	325	
ese	321	+ ése ... (76)
tener	314	+ que + inf. (45), + A.V. (17), + A.C. (2) = (64)
hacer	306	+ inf. (45), + A.C. (39), tercio-personal (25) = (109)
decir	290	
si	270	
sin	254	
ni	241	
mismo	240	
ya	236	
sino	234	
alguno	222	algo (55), alguien (15), alguno ...
estar	221	+ A.V. y A.C. (103), + ger. (45) = (148)
mi	203	
español	202	adj.
sólo	201	adj. (31), adv. (170)
vez	196	
tan	188	
sobre	178	
hombre	175	
España	173	
porque	172	+ porqué (1)
entre	170	

bien	169	
quien	165	+ quién (16)
obra	164	
mucho	163	adj. (101), adv. (49), pron. (13)
vida	163	
pueblo	158	
menos	157	
literatura	154	
aquel	149	+ aquél ... (46)
dar	145	
desde	141	
uno	141	+ uno(a)s: pronombre
deber	139	+ (de) + inf. (120)
cuando	138	+ cuándo (4)
tanto	136	adj. (86) + adv. (51)
ir	133	+ A.V. y A.C. (10) + ger. (50) + inf. (43) = (103)
siglo	133	
así	132	
aun	132	+ aún (63)
también	132	
tal	127	
escribir	123	
historia	121	
propio	121	
poesía	120	
querer	120	+ inf. (85)
saber	120	+ inf. (19)
escritor	119	
poco	119	adj. (57), pron. (13), adv. (49)
muy	118	
político	116	adjetivo
espíritu	116	
caso	114	
social	114	
tiempo	113	
mundo	112	
qué	112	
hasta	111	
parecer	110	+ A.V. y A.C. (41), + inf. (39) = (80)
aunque	108	
cierto	108	
ver	108	verse + A.V. y A.C. (12); ver + A.V. (9) = (21)
grande	105	

problema	105	
parte	104	
nuevo	103	
llamar	101	
palabra	101	
realidad	101	
venir	101	+ A.V. y A.C. (3), + ger. (23), + inf. (17) = (43)
donde	100	+ dónde (11)
dos	100	
cada	98	
siempre	97	
sí	97	
poner	96	
cual	95	+ cuál (11), el cual (68)
humano	94	
bueno	93	buen (24), bueno (69)
cuanto	93	+ cuánto (5), pron. rel. (21)
llegar	92	+ a + inf. (33)
cultura	91	
estado	91	sust.
pues	89	
entender	83	
sentido	83	
creer	83	
hablar	81	
mejor	81	adj (51) + adv. (29)
poeta	81	
cosa	80	
modo	80	
idea	79	
libro	77	
año	76	
dejar	76	+ A.V. (9), + A.C. (1) = (10)
país	76	
forma	75	
histórico	75	
voz	75	
ahora	73	
autor	73	
gobierno	73	
hoy	73	
nada	73	
crítica	72	sust.

fuerza	72	
aquí	71	
mayor	71	adj. + adv.
pasar	70	
cualquier	69	adj. + pron.
hecho	69	sust.; de hecho (9), el hecho (de) (que) (28)
literario	68	
conciencia	67	
razón	67	
arte	66	
crítico	65	sust.
pensar	65	
vivir	65	
fin	64	
posible	63	
ejemplo	61	
lector	61	
nombre	61	
llevar	60	
público	60	
seguir	60	
casi	59	
económico	59	
encontrar	59	
leer	59	
nacional	59	
nación	59	
lenguaje	58	
orden	57	
popular	57	
quedar	57	+ A.V. y A.C. (31)
general (jeneral)	56	adj.
último	56	
guerra	55	
claro	54	
lengua	54	
antes	53	
concepto	53	
publicar	52	
verdad	52	
novela	51	
punto	51	
considerar	50	

luego	50		
tanto	50		
español	49	sust.	
libertad	49		
poder	49	sust.	
sociedad	49		
tierra	49		
trabajo	49	sust.	
contrario	48	adj.	
dentro	48		
dictadura	48		
tratar	48		
fe	47		
momento	47		
universal	47		
volver	47		
demás	46		
día	46		
etcétera	46		
nunca	46		
personal	46	adj.	
tomar	46		
valor	46		
verdadero	46		
conocer	45		
opinión	45		
tampoco	45		

Salvo 22, todas las formas instrumentales del corpus figuran entre los 204 lemas más frecuentes. Son los artículos, los pronombres personales y relativos, los pronombres y los adjetivos demostrativos, posesivos e indefinidos, las preposiciones, las conjunciones, los adverbios cortos y los auxiliares.⁵

Algunas de las formas que faltan tienen una frecuencia vecina a la del último lema registrado (45), otras vienen muy a la zaga. Estas formas y su frecuencia son:

⁵ No hay que olvidar que este análisis se basa en la lengua escrita, la cual refleja sólo parcialmente la realidad lingüística. El número de morfemas es mucho más elevado que el que la lengua escrita distingue al transcribirlos de manera aislada. No se han anotado las desinencias y terminaciones, los prefijos y sufijos que se ortografían aglutinados al lexema. Si se tuvieran en cuenta, la cabeza de la lista de frecuencia se incrementaría con numerosas formas instrumentales, mientras que el número y la frecuencia de los lexemas no sufriría modificaciones.

ante	42
soler	41
todavía	41
quizás	40
después	38
hacia	37
nadie	35
siquiera	35
acaso	33
entonces	33
tras	30
pronto	28
fuera	26
jamás	23
apenas	22
(sin) embargo	22
suyo	22
además	18
respecto (a)	16
mas	13
pese (a)	5
usted	3

Son formas que pertenecen a series cerradas—no podrían forjarse otras—en las que las formas más frecuentes redundan a favor de las escasas formas ocasionalmente menos frecuentes—en otro corpus éstas podrían ser más frecuentes que aquéllas. Tal solidaridad no puede existir entre las formas léxicas al contarse éstas en número infinito. Llama también la atención que las formas con un índice de frecuencia menor son sobre todo adverbiales, probablemente la clase menos gramaticalizada de las formas instrumentales.⁶

Las formas instrumentales constituyen la cabeza de la lista y van escaseándose a medida que la frecuencia decrece. Los lexemas sólo ocurren poco a poco y acaban siendo mayoría. Pero algunas formas pueden ser tanto instrumentales como léxicas y antes de fijarse en su rango conviene examinar las construcciones en las que ocurren. Esto se da sobre todo con los verbos.

Ya sabemos que **ser** (1763), **haber** (1277) y **poder** (445) sólo son auxiliares,—no lo son tan pocas veces que se puede omitir el número de ejemplos en los que esto ocurre. No es así para **tener** (314), **hacer** (306) y **decir** (290), que tampoco son sólo lexemas: **tener** se construye 45 veces con **que** + infinitivo, y otras 17 y 2 veces con

⁶ A pesar de tener forma lingüística los números, tanto cardinales como ordinales, no pertenecen en realidad a la lengua, sino a otro sistema de comunicación en el que todos los signos son equiprobables.

adjetivo verbal y calificativo respectivamente. **Hacer** se conjuga con un infinitivo en 45 ejemplos, con un adjetivo calificativo en 39, y en forma terciopersonal en 25. Si separamos estos ejemplos, 64 para **tener** y 109 para **hacer**, de los demás, su frecuencia de empleo como lexema es sólo de 250 y 197, de los demás, mientras que la que alcanza en empleos eventualmente instrumentales es análoga a la de otras formas instrumentales, y para algunos suficientemente alta para ser considerados como auxiliares. Para los verbos más frecuentes las cifras son como sigue:

	total	empleos eventualmente instrumentales	empleos léxicos
ser	1763		
haber	1277		
poder	445		
tener	314	64	250
hacer	306	109	197
decir	290		
estar	221	148	73
dar	145		
deber	139	120	19
ir	133	103	20
escribir	123		
querer	120	85	35
saber	120	9	101
parecer	110	80	30
ver	108	21	87
llamar	101		
venir	101	43	58
poner	96		
llegar	92	33	59
entender	83	1 (entenderse)	
creer	81	1 (creerse)	
hablar	81		
dejar	76	10	66
pasar	70	1	
pensar	65		
vivir	65	6	59
llevar	60		
encontrar	59		
leer	59		
seguir	59	32	27
quedar	57	31	26
existir	53		

Decir forma a menudo parte de expresiones fijas, pero estas son más difíciles de contar por la multiplicidad de sus variantes: **es decir** (26 veces), **dicho sea, mejor dicho, dicho de otra manera** (12), **por (no) decir(lo) (sí)** (8), son las que más se destacan. De todos modos, en el corpus **decir** pertenece más bien al metalenguaje: sirve casi exclusivamente para referirse al propio texto u otros textos citados. De los 39 ejemplos con **dice**, por ejemplo, 22 acompañan una cita. Etc. En una situación análoga se encuentran **escribir** (123), **llamar** (101) y **entender** (83).

El primer sustantivo que aparece es **vez** (196), y a continuación aparecen **hombre** (175), **España** (173), **obra** (164), **vida** (163), **literatura** (154), **pueblo** (158), **siglo** (133), etc. Son tan frecuentes como las formas instrumentales que los rodean y más frecuentes que muchas otras que las siguen. Pero a la diferencia de estas cuyas frecuencias son análogas en cualquier corpus, los sustantivos alcanzan posiciones muy distintas según los corpus. No son de uso general, sino particular y sólo figuran entre los lemas más frecuentes en repertorios de la misma temática. Esta inestabilidad permite considerarlos como apariciones casuales y que no ponen en tela de juicio la relación entre el carácter instrumental de una forma y su alta frecuencia.

Los 19 textos son demasiado cortos para poder analizarse al respecto de manera significativa; la mayoría además pertenecen a un mismo género y tratan de temas análogos. Baste señalar que en este corpus **hombre** ni siquiera ocurre en 3 textos, **España** está ausente en 6, **obra** en 1, **vida** en 2, **literatura** en 6, así como **pueblo** y **poesía** en 10, **escritor** en 5, etc.⁷

Si se confronta el corpus de los 19 autores en su conjunto con cada uno de los cinco géneros analizados por A. Juilland,⁸—teatro, novela, ensayo, textos de periódico y textos técnicos,—todos de 100.000 ocurrencias, y si se calcula para cada uno de los sustantivos más frecuentes ($f > 100$) la dispersión (D) en los seis corpus según la fórmula propuesta por el mismo Juilland,⁹ se observan los valores siguientes:

⁷ Véase *Índices, Lista de frecuencia*, t. 4, pág. 269.

⁸ A. Juilland y E. Chang-Rodríguez, *Frequency Dictionary of Spanish Words*, La Haya, 1965.

⁹
$$D = 1 - \frac{s}{x \sqrt{n}}$$

en la que s es la desviación tipo, o sea

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

en la que x es la frecuencia relativa (aquí también absoluta ya que todos los corpus tienen la misma frecuencia) de una forma en un corpus, \bar{x} el promedio de las frecuencias de la misma forma en todos los corpus, o sea

$$\bar{x} = \frac{\sum x_i}{n}$$

y n el número de corpus, aquí 6.

Recordamos que el resultado oscila entre 0 y 1 y que 1 significa que la distribución es homogénea, y 0 lo contrario.

	19A	Teatro	Novela	Ensayo	Textos Periód.	Textos Técnicos	D	
vez	196	131	184	160	107	134	0,908	
hombre	175	314	168	264	94	101	0,807	
España	173	eliminado en Juilland						
obra	104	24	28	134	88	133	0,748	
vida	163	166	164	252	104	62	0,826	
literatura	154	0	3	55	9	11	0,367	
pueblo	158	33	68	127	110	47	0,779	
siglo	133	13	20	101	56	103	0,719	
historia	121	16	21	124	38	94	0,705	
poesía	120	3	0	28	10	2	0,299	
escritor	119	2	4	45	14	24	0,480	
espíritu	116	17	30	85	57	30	0,720	
caso	114	64	35	45	48	96	0,809	
tiempo	113	89	112	135	74	94	0,914	
mundo	112	88	86	138	65	66	0,876	
problema	105	3	5	56	45	64	0,660	
parte	104	39	39	111	92	192	0,760	
palabra	101	55	49	75	59	50	0,874	
realidad	101	8	28	62	59	36	0,730	

Se puede apreciar la magnitud de los valores calculados para estos sustantivos si se le compara con la de los valores que se calculan para las formas más frecuentes y sin duda ninguna instrumentales:

	19A	Teatro	Novela	Ensayo	Textos Periód.	Textos Técnicos	D
el	11.158	6.543	11.289	12.223	12.977	11.732	0,915
de(l)	7.681	3.631	6.581	7.811	8.623	8.498	0,893
y (e)	3.259	2.291	3.219	3.104	3.050	3.282	0,950
a(l)	2.628	2.975	2.732	2.435	2.971	2.607	0,968
en	2.645	1.580	2.647	2.781	2.778	3.136	0,917
un	1.751	1.686	2.358	1.809	1.506	1.382	0,921
que ^a	1.848	2.559	1.223	1.529	2.654	2.202	0,870
ser	1.763	2.892	1.634	2.066	1.578	1.737	0,896
se	1.602	1.288	1.720	1.556	1.673	1.644	0,960
que ^b	1.455	2.409	1.382	1.660	624	1.443	0,844

^a Pronombre.

^b Conjunción.

Si se comparan los valores de dispersión en las dos muestras mediante el test de Tukey-Kramer,¹⁰ se verifica que, a un nivel de confianza superior a 0,95, hay una diferencia significativa entre ambas, y que los sustantivos se diferencian pues de las formas instrumentales en cuanto a su dispersión.

Entre los sustantivos se observan algunos cuyo valor de dispersión es más elevado que los que los sustantivos suelen tener y se acerca a los de las formas instrumentales: **tiempo**: 0,914, **vez**: 0,908, **mundo**: 0,876, son los más altos. **Vez** aparece a menudo en expresiones fijas que debilitan su naturaleza léxica y que, como locuciones, se comportan hasta cierto punto como instrumentos: **a veces** (25 ejemplos), **a la (su) (mi) vez** (19), **en vez de (que)** (5), **de vez en cuando** (1). Pero no es así con **tiempo** y **mundo**, o por lo menos no en la misma medida.

Las mismas verificaciones pueden hacerse para los diez adjetivos calificativos más frecuentes ($f > 62$) que aparecen entre las formas más frecuentes del corpus. La mayoría tiene valores de dispersión relativamente bajos:

	19A	Teatro	Novela	Ensayo	Textos Periód.	Textos Técnicos	D
español	202	21	31	138	129	78	0,716
político	116	4	5	75	48	17	0,586
social	114	5	10	53	39	17	0,581
gran(de)	105	81	161	165	200	188	0,872
nuevo	103	36	69	104	96	75	0,867
buen(o)	93	231	78	78	29	56	0,694
humano	94	8	18	89	18	36	0,646
histórico	75	0	5	44	16	25	0,584
literario	68	5	4	71	19	8	0,556
posible	63	35	18	37	24	30	0,815

A la diferencia de otras formas instrumentales los instrumentos adjetivos **todo**, **otro**, **mismo**, **solo**, **mucho**, **tanto**, **tal**, **propio**, **poco** y **cierto** no se distinguen de los adjetivos calificativos, lexemas, por su forma. Sólo pueden ser disociados por su frecuencia relativamente alta a la par que la relativa regularidad de ésta en corpus diferentes.

$$^{10} \mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm t \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

en la que \bar{x} representa el promedio de los valores de dispersión de una serie de vocablos, s la desviación tipo, es decir

$$\sqrt{\frac{\sum (x_i - \bar{x})^2}{N - 1}} \quad (\text{véase supra, pág. 486 nota 3}), n \text{ el número de observaciones o vocablos de la serie, y } t \text{ se}$$

deriva de la **tabla t** según n y el nivel de confianza escogido.

	19A	Teatro	Novela	Ensayo	Textos Periód.	Textos Técnicos	D
todo	516	577	519	553	523	464	0,970
otro	353	256	302	346	241	429	0,911
mismo	240	133	155	209	164	260	0,892
solo	201	133	120	153	128	162	0,919
mucho	163	215	186	144	118	178	0,915
tanto	136	103	78	84	87	86	0,909
tal	127	27	60	94	71	87	0,822
propio	121	18	39	77	40	53	0,743
poco	119	108	183	118	84	101	0,910
cierto	108	44	44	76	56	76	0,851

También llama la atención que la mayoría de estos adjetivos se colocan siempre, preferentemente o a menudo delante del sustantivo al cual se refieren, como lo hacen las formas instrumentales. Es curioso observar que los tres adjetivos calificativos que tienen los valores de dispersión más altos (**gran(de)**, **nuevo** y **posible**) son los más propensos a la anteposición entre los adjetivos examinados.¹¹

La comparación de las muestras en cuanto a los valores de dispersión calculados, revela a su vez una diferencia significativa entre las dos muestras de adjetivos, como entre la muestra de adjetivos calificativos y la de las formas instrumentales más frecuentes. Al contrario no hay diferencia significativa entre esta última muestra y la de los adjetivos instrumentales como tampoco la hay entre la muestra de adjetivos calificativos y la de los sustantivos. Las muestras de los adjetivos instrumentales y de los sustantivos, sin embargo, se revelan significativamente distintas.¹²

El análisis aproximativo e incompleto y por lo tanto provisional de la lista de los lemas más frecuentes del corpus de los 19 autores confirma que no es falso que las formas más frecuentes sean en su mayoría instrumentales y que la mayoría de las formas instrumentales se encuentren entre las formas más frecuentes, y que, por consiguiente, se puede acudir a este rasgo para definir la auxiliaridad, función instrumental.

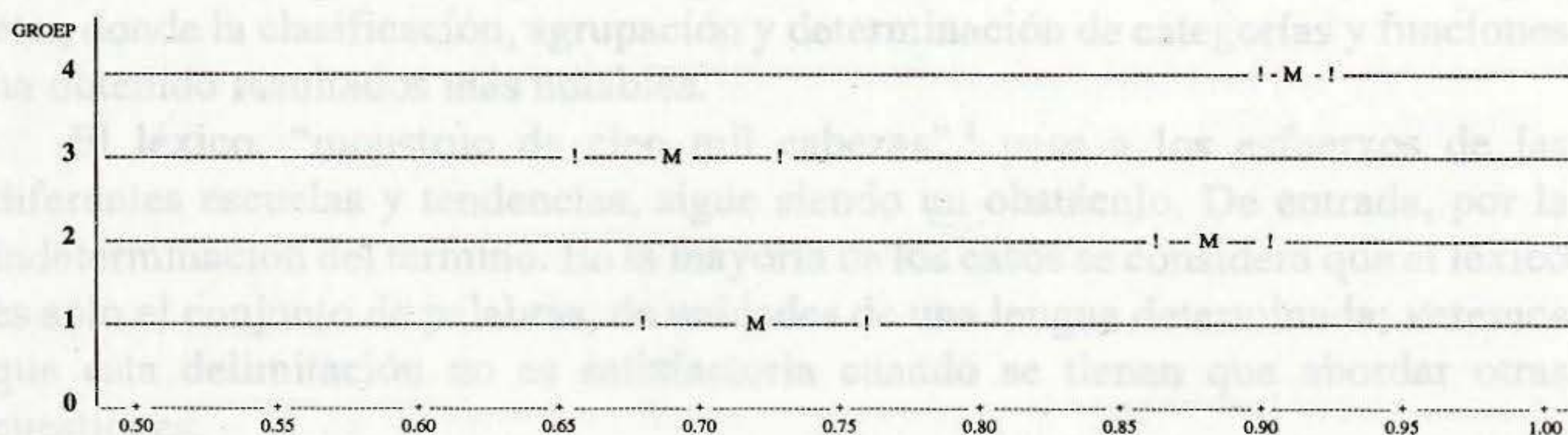
Es cierto que no puede trazarse un límite de frecuencia preciso entre ambas clases de formas. Pero también es verdad que dentro del conjunto del vocabulario, las formas instrumentales, acumuladas al principio de la lista de frecuencia decreciente, se agotan de manera relativamente rápida, mientras que los lexemas sólo aparecen progresivamente, y que sólo en apariencia esta distribución de

¹¹ El 92% de las ocurrencias de **gran(de)**, asociado a un sustantivo es antepuesto, el 76% de **nuevo** y el 48% de **posible**.

¹² Se observa que a un nivel de confianza de 0,99% la diferencia de dispersión entre estas dos últimas muestras es la única que cesa de ser significativa.

frecuencia se ve perturbada por formas ambivalentes y lexemas localmente frecuentes.

Nota: La diferencia o similitud de los promedios de valores de dispersión para las cuatro muestras examinadas pueden representarse gráficamente de la manera siguiente:



MEAN

M representa el promedio de dispersión ($= \bar{x}$) en cada muestra; los dos puntos de exclamación delimitan la zona dentro de la cual el promedio puede variar para ser significativo. **4** representa las formas instrumentales más frecuentes (**el, de(l), a(l)**, etc.), **3** los sustantivos (**vez, hombre, obra**, etc.), **2** los adjetivos indefinidos (**todo, otro, mismo**, etc.) y **1** los adjetivos calificativos (**español, político, social**, etc.).

Jesse de Kock

Es sabido que no se usan bien los diccionarios, que se consideran elemento auxiliar en la enseñanza, que no se obtiene de ellos todo el rendimiento que sería de desear. La información de cualquier artículo de un diccionario es muy abundante. En la planta de estas obras hay un objetivo fundamental: el significado de la unidad; pero también puede aparecer la etimología, categoría gramatical, nivel social o geográfico de uso, sinónimos, antónimos, etc.

Me voy a detener en un aspecto poco tratado en el estudio del léxico: la fraseología² y su tratamiento en los diccionarios. Sobre esta cuestión los

² G. Solvèr, "Dinámica estructural y variaciones del vocabulario", *Semántica y morfología del español*, Madrid, 1954, p. 67.

³ En el artículo bibliográfico sobre el tema, en el libro o "Fraseología y enseñanza" en prensa en el *Boletín de*