

LAS PALABRAS FAMILIARES

Este artículo describe una de las primeras herramientas de cálculo ideadas para matizar las frecuencias obtenidas en experiencias de vocabulario y que son un antecedente claro de los intentos para estimar los coeficientes de *disponibilidad léxica* que caracterizan a una comunidad de habla.

La construcción de un modelo matemático con esta intención debe reproducir de la manera más exacta posible, tanto el hecho lingüístico en sí mismo, como el contexto en que el acto de habla se produce. Ello implica que cualquier elemento con incidencia en el fenómeno debe tener una potencia numérica en la ecuación proporcional a su significación real establecida con criterios estrictos de objetividad y rigor. Tampoco puede olvidarse las servidumbres derivadas del método experimental utilizado para recoger los datos a partir de los cuales se opera.

Es ineludible, por ello, realizar un examen pausado de los factores que constituyen la esencia de la *disponibilidad* y establecer las condiciones matemáticas que los integrará en el procedimiento adoptado para medirla. Para comprender las causas que determinan la tendencia de algunas palabras a ser actualizadas aquí y ahora con preferencias a otras más usuales y frecuentes, es necesario empezar por descubrir las circunstancias, facultades, sucesos o leyes que actúan sobre los individuos cuando deciden establecer una conversación acerca de un tema familiar y cercano. Interesa especialmente ponderar el componente método. No se olvide que la investigación no se erige a partir de la observación directa del acto de la comunicación, sino que articula una situación artificial de “laboratorio” con la intención específica de intuir, a partir de ella, lo que puede llegar a ocurrir en un entorno espontáneo de cotidianidad. Procede, después, delimitar los ámbitos y mecanismos de todos y cada uno de los componentes descubiertos, así como evaluar la intensidad con que operan, se interfieren o favorecen.

Permítase empezar aclarando que sólo en este sentido se analizan con mirada crítica algunas de las formulaciones manejadas en importantes investigaciones pioneras

sin que, en ningún caso, se pretenda devaluar sus conclusiones, ni cuestionar las construcciones metodológicas imaginadas por quienes las realizaron. Muy al contrario, con este ejercicio se aspira a interpretar mejor los mecanismos que desarrollaron, las razones con que justificaron sus axiomas y postulados y los excesos, si es que los hubo, en que pudieron incurrir. Se trata de indagar, mediante un ensayo reflexivo, en su epistemología para descubrir aquello que cada una aporta a la investigación lingüística y que no se puede perder en modo alguno, junto a sus direcciones más cuestionables por poco productivas.

La primera tentativa moderna cercana a la línea de investigación de la *disponibilidad* pertenece más al ámbito de la psicología que al de la lingüística, aunque relaciona estrechamente ambas disciplinas. Se centra en la búsqueda, ya mencionada, de las *palabras familiares*.

Los métodos utilizados para establecer los índices de *familiaridad léxica* están claramente vinculados a los planteamientos sobre los que se construye este concepto. Recogen, en principio, dos aspectos de un mismo fenómeno: la conciencia que posee el propio sujeto acerca de los términos que le son más cercanos y la rapidez con que el individuo reconoce determinadas palabras y responde al estímulo que representan.

Basados en la primera premisa se encuentran bastantes trabajos. Se comentan a continuación los realizados por P. Fraisse en París y por G. Noizet y C. Flament en Aix y Marsella que proclamaron de forma clara e inequívoca en sus conclusiones la alta correlación existente entre la *frecuencia de uso* y el tema del discurso.

Las experiencias de París, Aix y Marsella [FRAISSE, NOIZET y FLAMENT. 1977a:180-181] consistieron en proponer a un grupo de estudiantes una lista de 25 sustantivos (pertenecientes a las bandas logarítmicas con *frecuencia* de 1, 4, 16, 64 y 256, sobre $3/2$ de término medio, según las listas de Gougenheim de 1956), mezclados al azar y presentados en una sola columna, para solicitar de los encuestados su ordenación de acuerdo a los siguientes criterios: se debía asignar un 1 a las cinco palabras consideradas como “más familiares” y un 5 a igual cantidad de términos reconocidos como “menos familiares”; un 2 y un 4, respectivamente, a otros cinco vocablos calificados como “bastante” o “poco familiares”; se adjudicaría, finalmente, el valor 3 a las voces no puntuadas con anterioridad.

Los investigadores, después, van a otorgar a estos rangos un peso decreciente que oscila entre 4, para la elección más *familiar*, y 0. Los *índices de familiaridad* de cada palabra se obtienen de promediar la suma de las puntuaciones alcanzadas. En

definitiva, los cálculos realizados para cada palabra responden a la siguiente ecuación:

$$\text{Familiaridad palabra} = \frac{4 f_{p1} + 3 f_{p2} + 2 f_{p3} + 1 f_{p4}}{N}$$

donde f_{p1}	es frecuencia de la palabra en la posición 1
f_{p2}	es frecuencia de la palabra en la posición 2
f_{p3}	es frecuencia de la palabra en la posición 3
f_{p4}	es frecuencia de la palabra en la posición 4
N	es número de entrevistados

(Nótese la desaparición del sumando correspondiente a la 5ª posición al atribuirle un peso de 0).

La comparación de los resultados obtenidos, *escala de familiaridad*, con la escala de *frecuencia de uso* de Gougenheim mediante el *coeficiente de correlación de Pearson* alcanzó un valor de 0.90 que demuestra la alta dependencia existente entre ambas medidas. No obstante, la detección de términos discrepantes con la norma general lleva a los investigadores a manifestar que:

“El distingo entre la primera palabra y la segunda es mucho más importante en familiaridad que en logaritmo de *frecuencia* de uso. Lo cual entraña una estrechura hacia lo alto de la mayor parte de la escala”. [FRAISSE, NOIZET y FLAMENT. 1977a:185].

El segundo procedimiento para precisar la *familiaridad* analiza la fase de recepción del acto comunicativo oral o escrito. Para ello, con ayuda del *taquitoscopio*, se controla lo que el individuo tarda en percibir el significado de las palabras midiendo el tiempo transcurrido entre su presentación y la reacción que provoca.

Ya se ha comentado en el capítulo anterior la importancia que, en Psicología, han tenido los experimentos basados en los *tiempos de reacción*. Los estudios taquitoscópicos aplicados al vocabulario comprueban que los términos que acreditan alta *frecuencia* producen latencias de respuesta más cortas y menor número de errores que los términos de baja *frecuencia* [HOWES y SOLOMON. 1951:401-410; SOLOMON, R. L. y HOWES, D. H. 1951:256-270. Desde ese momento los psicolingüistas aceptan que el *umbral de reconocimiento* de la *frecuencia de uso* depende de la longitud de la palabra, de la actitud del sujeto y de sus intereses o necesidades en el momento de la prueba. Entienden también que el proceso perceptivo ha de provocar necesariamente una acción discriminatoria o identificativa expresa [FRAISSE. 1977b:192]. Sólo entonces se considera

finalizada la percepción. Lo confirman midiendo el tiempo que cada sujeto necesita para reconocer y reaccionar ante palabras de *frecuencia* y *familiaridad* conocida presentadas como estímulo.

“No éramos los primeros en encarar este problema. En 1947, Bruner y Postman, y en 1954 Delucia y Stagner, lo habían estudiado. En ambos casos el material verbal tenía connotaciones afectivas, y en ambas investigaciones los autores habían encontrado una relación curvilínea entre umbral de reconocimiento y tiempo de reacción verbal”. [FRAISSE. 1977b:192]

Las investigaciones de Fraisse [FRAISSE. 1977b] avalan los resultados de Howes y Solomon y demuestran, sin dejar lugar a la duda, que el *umbral de reconocimiento* desciende en la medida que aumenta la *familiaridad* de la expresión. Cuantifican la correlación así:

$$\rho = + 0.83$$

No se alcanzan, sin embargo, índices tan elevados cuando se correlaciona *umbral de reconocimiento* y *frecuencia*. En este caso el coeficiente desciende hasta la estimación

Los valores anteriores corroboran que el sujeto aprecia como más suyas aquellas

$$\rho = + 0.62$$

palabras que designan objetos próximos o habituales, frente a otras, quizás más usuales en el habla cotidiana, pero de significación menos familiar.

Al margen de las conclusiones, puede resultar de interés exponer algunas reflexiones acerca de los métodos utilizados. Se trabaja siempre sobre palabras seleccionadas por cumplir los requisitos apriorísticos de la investigación. Los vocablos no fluyen espontáneamente ante los sujetos entrevistados. Es posible que la percepción de proximidad hubiese variado de integrarse dentro de un contexto comunicativo determinado.

Hay que advertir también que los datos manejados en los experimentos del primer tipo se encuadran en la categoría de *ordinales*. Esta clase de medición coloca cada observación en una posición relativa con respecto a todas las demás, pero no especifica la distancia entre dos mediciones adyacentes, es decir, a partir de la ordenación se conoce que “a” es más, o menos, que “b”, pero no existe legitimidad

para cuantificar la amplitud real de su diferencia. La tabla de Siegel, *Tabla 1.2.1 - 2*, ya comentada, reduce a *mediana, percentil y coeficientes de correlación de Spearman y Kendall*, los *estadísticos* que le son aplicables.

Para incrementar la información proporcionada por esta clase de dato se suele recurrir a otorgar *pesos* a las posiciones. Eso significa ponderar la influencia relativa que cada lugar ostenta con relación a la variable tomada en su conjunto. Se concede un valor numérico a cada rango de modo que, en su variación, se concrete la medida estimada del entramado jerárquico en que se inserta. Se adjudican puntajes mayores a las primeras posiciones. El procedimiento es habitual en la *toma de decisiones* y tiene como base el sentido común.

Sin embargo, esta aproximación intuitiva está condicionada, ya que podría sesgar el análisis [GASS, S. 1991]. Transforma la escala ordinal en numérica, lo que incrementa la potencia y fiabilidad de los *estadísticos*, pero supone una ingerencia en las respuestas. Siempre que los expertos [FOX, D. J. 1981:725] asignan un *peso*, aún cuando la decisión haya sido fruto de profunda y bienintencionada reflexión, se puede demostrar que habría conclusiones diferentes a partir de parámetros distintos.

Las *Tablas 2.1 - 1* y *2.1 - 2* resumen las variaciones que experimentarían los índices de familiaridad en un hipotético experimento realizado sobre 15 informantes. Con mover sólo una unidad el rango de puntuación, de 0 - 4 a 1 - 5, se produce un cambio en los *coeficientes de familiaridad* que obliga a la reordenación de palabras. Hay un intercambio de los lugares 1ª y 2ª entre las palabras “c” y “d”. También los términos “a” y “e” se alternan en 4ª y 5ª posición.

Es fácil concluir, a la vista de lo expuesto, lo aventurado de extraer afirmaciones categóricas a partir de escalas ordinales convertidas en numéricas por este camino.

Tabla 2.1. - 1

Cálculos efectuados a partir de los *pesos* aplicados en la experiencia parisina

	<i>Peso</i>	4	3	2	1	0	<i>Coeficiente</i>	
	<i>Posición</i>	1	2	3	4	5	<i>familiaridad</i>	<i>POSICIÓN</i>
<i>Palabras</i>	<i>a</i>	1	1	3	1	9	0,933333	5
	<i>b</i>	4	2	2	4	0	2	3
	<i>c</i>	4	4	4	6	4	2,8	2
	<i>d</i>	3	7	5	1	2	2,933333	1
	<i>e</i>	3	1	1	3	0	1,333333	4
<i>Total en la posición</i>		15	15	15	15	15		

Tabla 2. 1 -2

Cálculos efectuados cambiando los pesos originales para dar el valor 1 a la 5ª posición

	Peso	5	4	3	2	1	Coficiente	
	Posición	1	2	3	4	5	familiaridad	POSICIÓN
Palabras	a	1	1	3	1	9	1,933333	4
	b	4	2	2	4	0	2,8	3
	c	4	4	4	6	4	4,266667	1
	d	3	7	5	1	2	4,133333	2
	e	3	1	1	3	0	1,866667	5
Total en la posición		15	15	15	15	15		

REFERENCIAS

FRAISSE, Paul (1977b). “La percepción de las palabras”. En *Introducción a la psicolingüística*, 191-200. PIAGET, J. (ed). Traducción de *Problèmes de psycho-linguistique*, Presses Universitaires de France, 1967. París. Ediciones Nueva Visión SAIC, Buenos Aires.

----; **NOIZET, Georges y FLAMENT, Claude.** (1977a). “Frecuencia y familiaridad del vocabulario”. En *Introducción a la psicolingüística*, 179-190. PIAGET, J. (Ed.). Traducción de *Problèmes de psycho-linguistique*, Presses Universitaires de France, 1967. París. Ediciones Nueva Visión SAIC, Buenos Aires.

HOWES. D. (1954). “On the interpretation of word frequency as a variable affecting speed of recognition”. *Journal of Experimental Psychology*, 48:106-112.

---- y **SOLOMON, R. L.** (1951). “Visual duration threshold as a function of word probability”. *Journal of Experimental Psychology*, 41:401-410.