

## ANEXO Nº 2. APLICACIÓN DE LA METODOLOGÍA EXPUESTA, SOBRE UNA “MUESTRA PILOTO” (EXTRACTO)

El objeto del presente trabajo es el estudio de determinados aspectos relacionados con el dialecto murciano: [...] en segundo, una somera aproximación al conocimiento del léxico asociado a dicha variedad [...].

Para ello, se ha contado con los resultados de una encuesta con diversas cuestiones que hacen referencia a dichos aspectos y que, con respuestas ordenadas y calificadas escalarmen- te, ha sido respondida por 39 individuos pertenecientes a un mismo ámbito laboral (firma profesional dedicada a la auditoría, consultoría empresarial y el asesoramiento fiscal y laboral) y, a pesar del sesgo que puede conllevar el hecho de la limitación del ámbito escogido (posibilidad de filtración a causa de constituir una *comunidad de habla* reducida durante buena parte del día laboral), la muestra resulta suficientemente variada en cuanto a los parámetros edad, sexo/género y entorno social habitual, en este último caso se ha distinguido entre capital (Murcia ciudad) y no capital, ya que no resultaba procedente la distinción urbano/rural dado que ninguno de los encuestados se desenvolvía en este ámbito; sin embargo, no resultaba variada en lo referente al nivel de estudios (prácticamente la totalidad de los encuestados tienen estudios universitarios), por lo que este último aspecto no ha sido considerado en los análisis de resultados.

Dicho análisis ha sido llevado a cabo sin la menor manipulación de los datos obtenidos, por lo que cabe considerar la posibilidad de que exista –y, de hecho, así sucede– alguna respuesta con falta de coherencia, bien por error, bien por mala interpretación de la pregunta; en todo caso, la incidencia de dichos “errores”, que siempre han de ser considerados en encuestas “no guiadas”, no resulta lo suficientemente significativa como para desvirtuar los resultados globales.

En cuanto a la metodología, para cada uno de los tres aspectos considerados, se ha efectuado una tabulación global de resultados y se ha desglosado, posteriormente, de acuerdo con los parámetros ya mencionados de edad, género y entorno, en los dos primeros aspectos (es decir, se le ha dado un enfoque ligeramente sociolingüístico). Asimismo, se han confeccionado algunos gráficos (en los casos en que se ha considerado que eran convenientes) y se han practicado algunos contrastes estadísticos (paramétricos y no paramétricos), si resultaban oportunos. Con todo ello se procede, finalmente, al análisis de resultados y las correspondientes conclusiones, en cada uno de los tres apartados de que consta el trabajo.

Se preguntó a los encuestados acerca de su grado de conocimiento de diez palabras propias (aunque alguna de ellas no sea totalmente exclusiva) del dialecto murciano. Este conocimiento fue puntuado de 1 a 3 en cada una de las respuestas, lo que da un recorrido de puntuación de 10 a 30; un total de veintidós valores posibles, que han sido agrupados en tres intervalos de siete valores cada uno y señalan, como es lógico, un grado de conocimiento léxico bajo, medio y alto, respectivamente.

Esta división de la puntuación ha sido aplicada a cada una de las tres categorías en las que se dividió la muestra: edad, género y entorno social (predominante), y se ha traspasado la información a tres tablas 2x3 de contingencia, donde la cifra que aparece en cada una de las casillas de la parte central de la tabla indica el número de individuos (frecuencia conjunta absoluta) de la muestra que cumple lo que denota la fila y columna que le corresponde; así, por ejemplo, el número de individuos que tienen entre 25 y 55 años y obtuvieron entre 10 y 16 puntos es de 10.

Inmediatamente debajo de cada uno de los valores de frecuencia conjunta real aparece, entre paréntesis, un número que indica la frecuencia teórica que se debería haber dado en el supuesto de que los pares de variables que han sido consideradas en cada caso (*edad/puntuación*, *género/puntuación* y *entorno/puntuación*) no tuviesen relación entre sí. Por supuesto, la coincidencia entre el valor real y el teórico solo podría deberse al azar, pero podemos averiguar si las diferencias producidas son lo suficientemente pequeñas como para atribuir las al proceso de muestreo y, consiguientemente, aceptar la

independencia (incorrelación) entre las variables o, por el contrario, son grandes y nos llevan a concluir que existe relación entre las variables consideradas.

Para ello utilizaremos un contraste (test) no paramétrico de independencia basado en la llamada  $\chi^2$  de Pearson que, para  $(2-1) \times (3-1) = 2$  grados de libertad y un nivel de confianza del 95%, toma el valor 5,99.

En el primer caso, emparejando las variables *edad* y *puntuación*:

		PUNTUACIÓN			
		10-16	17-23	24-30	
EDAD	25-55	10 (8,72)	19 (17,44)	5 (7,85)	34
	56 o más	0 (1,28)	1 (2,56)	4 (1,15)	5
		10	20	9	39

$$\Sigma \Sigma d_{ij}^2 = 24,38 > 5,99 = \chi^2_{0,05;2} \Rightarrow \text{se rechaza independencia.}$$

En el segundo, para las variables *género* y *puntuación*:

		PUNTUACIÓN			
		10-16	17-23	24-30	
GÉNERO	M	4 (4,87)	9 (9,74)	6 (4,38)	19
	F	6 (5,13)	11 (10,26)	3 (4,62)	20
		10	20	9	39

$$\Sigma \Sigma d_{ij}^2 = 3,32 < 5,99 = \chi^2_{0,05;2} \Rightarrow \text{se acepta independencia.}$$

Por último, para las variables *entorno* y *puntuación*:

		PUNTUACIÓN			
		10-16	17-23	24-30	
ENTORNO	C	8 (6,92)	13 (13,85)	6 (6,23)	27
	$\bar{C}$	2 (3,08)	7 (6,15)	3 (2,77)	12
		10	20	9	39

$$\Sigma \Sigma d_{ij}^2 = 3,88 < 5,99 = \chi^2_{0,05;2} \Rightarrow \text{se acepta independencia.}$$

Estos resultados nos permiten concluir que sí existe relación de dependencia (correlación) entre las variables *edad* y *puntuación* y que, por su propia naturaleza, esta es de tipo unidireccional; más

concretamente, que la edad influye en la puntuación obtenida. Por el mismo razonamiento concluimos que el género y el entorno no tienen influencia en el conocimiento del léxico murciano.

Otro objetivo de este estudio podría ser la comprobación estadística de que la media de la puntuación obtenida por los hablantes de mayor edad es superior a la obtenida por los de menor edad, a un nivel suficientemente significativo. En definitiva, que los mayores tienen un conocimiento del léxico murciano superior al de los jóvenes. Por los datos muestrales obtenidos ya se puede suponer que las variables *género* y *entorno* no presentarán diferencias dignas de estudio (las medias muestrales fueron de 20,95 y 19,75, para masculino y femenino, respectivamente, y de 19,74 y 21,67, para capital y no capital, es decir, pequeñas diferencias a favor del género masculino y del entorno no capital).

Para profundizar más en el estudio del apartado que nos parece más interesante (grado de conocimiento léxico en función de la edad) utilizaremos los contrastes paramétricos para medias y diferencia de medias basados en la *t* de Student, con un nivel de confianza del 95%.

Llamando  $m_1$  a la puntuación media (grado de conocimiento léxico) de los hablantes con edades comprendidas entre los 25 y los 55 años, y  $m_2$  a la de los hablantes de 56 años o más, contrastaremos unilateralmente la hipótesis de que  $m_1 \geq m_2$  (hipótesis nula) con la alternativa de que  $m_1 < m_2$ . Por otro lado, si es aceptada la hipótesis alternativa, construiremos un intervalo de confianza para la diferencia  $m_1 - m_2$ .

La región de aceptación resultante para el contraste planteado es **A**:  $(-1,69; +\infty)$  y el valor del estadígrafo de contraste es de -3.63, que queda fuera de dicha región, por lo que, al nivel de confianza contemplado, hay que rechazar la hipótesis nula y aceptar la alternativa, es decir, se puede concluir que  $m_1 < m_2$  y podemos proceder con el siguiente paso previsto.

De acuerdo con los cálculos efectuados, el intervalo de confianza, al 95%, para la diferencia  $m_1 - m_2$ , es  $I_c = (-8,65 ; -4,99)$ , lo que significa que, al nivel que se ha elaborado dicho intervalo, la puntuación esperable en los hablantes de más de 56 años es superior a la de los de edades comprendidas entre 25 y 55 años entre casi 5 y casi 9 puntos (en el entorno de los 6,8 puntos), lo que indica una diferencia de alrededor de un 35%.