

OPINIONES DE ALUMNOS DE CENTROS ESPECIALES SOBRE LA ESCUELA: UNA TIPOLOGÍA BASADA EN ANÁLISIS DE TEXTOS POR ORDENADOR

por

*Norberto Corral Blanco
Samuel Fernández Fernández
Raquel-Amaya Martínez González
José Luis San Fabián Maroto*

ORDENADORES E INVESTIGACIÓN CUALITATIVA

Una de las características que definen el ámbito de la investigación en ciencias sociales durante los últimos años es sin duda el creciente interés despertado por el empleo de metodologías cualitativas. Rebasada ya la fase de confrontación entre diseños cuantitativos y cualitativos y una vez que éstos han adquirido un cierto reconocimiento y aceptación en la comunidad científica, se abre una nueva y decisiva etapa dirigida a dar respuesta a las expectativas creadas.

A diferencia de lo que ocurre en la metodología cuantitativa, la descripción de las técnicas de análisis de datos no ha ocupado un lugar central en los métodos cualitativos. Con frecuencia, la «calidad» de los datos cualitativos se ha visto rebajada ante la inexistencia de técnicas de análisis adecuadas, debiendo éstas hacer frente a numerosos desafíos derivados del tamaño de las muestras, la complejidad de los análisis, el refinamiento de los sistemas de clasificación, la flexibilidad que los procedimientos reclaman, etc. Las respuestas a algunos de estos problemas parecen guardar relación con la aplicación de los ordenadores al análisis de datos cualitativos (Richards y Richards, 1987). Esto plantea inevitablemente ciertas cuestiones: ¿puede contribuir el empleo del ordenador al desarrollo de los procedimientos de análisis empleados en la investigación cualitativa?, ¿es posible aplicar modelos computacionales que respeten las principales características del diseño cualitativo, tales como significación, flexibilidad, etc.?

Cuando se habla de ordenadores en investigación cualitativa en seguida pensamos en el gran cúmulo de datos procedentes de entrevistas, protocolos, notas de campo, etc. que podrían ser tratados. Es decir, se piensa en ellos como una solución cuantitativa más que cualitativa a los problemas metodológicos. Hasta ahora el ordenador se ha aplicado predominantemente a análisis «estructurales», por ejemplo, al análisis de contenido clásico (Berelson, 1971). Algunos de los primeros usos «cualitativos» se hicieron también a partir de procesadores de textos y programas de bases de datos. Nos encontramos ahora en una nueva fase, la fase de aplicación a análisis «interpretativos». En ella se busca el tratamiento de información simbólica no transformada en números, es decir, se establecen relaciones entre datos a partir de nexos (nodes, buttons), —palabras, frases o estructuras lingüísticas comunes a diferentes unidades de datos—.

Sabemos que el diseño cualitativo aconseja que las codificaciones se hagan después de un largo proceso de familiarización con los datos. Ello ha fomentado el desarrollo de una generación de programas «interactivos», alternativos a los de «sucesión lineal» (Miles y Huberman, 1984), que permiten modificar, ampliar o reducir categorías durante el proceso de análisis. Estos programas, distintos a los sistemas de gestión de base de datos, no necesariamente requieren partir los textos en segmentos a la hora de introducirlos en el ordenador, permitiendo posteriormente definir y redefinir con facilidad los segmentos.

Con este marco de referencia, presentamos a continuación un modelo de análisis de textos utilizando el SPAD (Lebart, 1982). Dichos textos han sido elaborados por alumnos de centros especiales y refleja, en algunos casos, lenguajes complejos y particulares, lo que plantea un reto adicional a este tipo de metodología.

METODOLOGÍA

Instrumento de recogida de datos

Los datos han sido recogidos mediante una redacción encabezada por el siguiente epígrafe: «ESCRIBE EN POCAS LÍNEAS, Y CON TUS PROPIAS PALABRAS, TU OPINIÓN SOBRE LA ESCUELA Y LAS RAZONES POR LAS QUE VIENES A ELLA». Esta redacción forma parte de un grupo de instrumentos dirigidos a valorar la percepción de sucesos escolares de una amplia muestra de alumnos de Enseñanza Primaria de Asturias (San Fabián et al., 1989).

Estos instrumentos, en su mayoría de respuesta abierta, nos permiten trabajar con la expresión libre del alumno y analizar la forma en que éste construye significados.

Respecto a la cohorte de Educación Especial incluida en este trabajo, cabe señalar que, dada la problemática motórica y sensorial de los alumnos, ha sido necesaria, en algunos casos, la colaboración de los tutores para poder lograr una mejor comprensión de los contenidos a modo de «traducción simultánea».

Muestra utilizada

La información procede de un grupo de 43 alumnos/as escolarizados en los centros de Educación Especial del Principado de Asturias. Estos sujetos muestran necesidades educativas especiales por poseer minusvalías de tipo motórico, sensorial o psíquico y se estima que su nivel intelectual y cognoscitivo se corresponde con el de ciclo medio de la enseñanza primaria, oscilando sus edades entre los 13 y los 34 años.

ANÁLISIS DE DATOS

La información procedente de las redacciones ha sido procesada mediante el paquete SPAD, disponible en la Universidad de Oviedo.

El proceso de análisis incluye los siguientes pasos:

1. Creación de un fichero de textos procurando mantener en la medida de lo posible la literalidad de los mismos.
2. Selección de palabras utilizando dos criterios: a) aparecer al menos en tres ocasiones y b) tener cuatro o más letras. Con ello se trata de eliminar nexos del tipo «con», «de», «los», etc. y de evitar una excesiva dispersión de palabras. El número de palabras seleccionadas mediante este procedimiento ha sido de 48.
3. Revisión de aquellas palabras cuyo significado depende del contexto en que aparecen, por ejemplo, «todos», «tengo», «estoy», etc. Además, ciertas «frases hechas» o expresiones tópicas son representadas por la palabra más característica, acompañada de uno o más asteriscos según el número de acepciones asignadas. Así, por ejemplo, en las frases donde aparecen referencias al colegio tales como «está bien», «lo paso bien», «me gusta» (el colegio), han sido sustituidas por «bien*», «paso*» y «gusta*» respectivamente.
4. Elaboración de la tabla de palabras por alumnos, en la que cada individuo es caracterizado por la frecuencia con que utiliza cada una de las palabras seleccionadas.

Estos cuatro primeros pasos son, en lo esencial, equivalentes a una codificación manual del texto, pero con las siguientes ventajas: a) la selección básica de las palabras se realiza de forma idéntica para todos los alumnos; b) la mayor parte del trabajo lo efectúa el ordenador de forma automática; c) es posible revisar la codificación de forma iterativa y d) se puede repetir el proceso con distintos criterios, tantas veces como se crea conveniente, sin que esto suponga un incremento grande de trabajo.

5. Realización de un análisis de correspondencias que da como resultado la selección de 10 ejes factoriales que explican el 75% de la variación global.
6. Clasificación jerárquica ascendente sobre los 10 ejes factoriales de la que se obtienen cinco clases.

7. Interpretación de cada una de estas clases en función de las palabras que mejor las caracterizan, tanto por su presencia, como por su ausencia.
8. Caracterización de los grupos obtenidos por sus variables demográficas. En el ejemplo que presentamos este paso no resulta de interés al tratarse de una muestra reducida.

RESULTADOS

La siguiente es una selección de cinco textos representativos de la expresión y contenido reflejados por los alumnos.

«ME GUSTA EL COLEGIO PORQUE TIENE TALLERES Y PISCINAS. EL COLEGIO ESTÁ BIEN».

«YO TENGO 4 AÑOS EMPIEZA EL COLEGIO SORDO. YA TIENES 13 AÑOS, PRONTO UN AÑO ACABO AL COLEGIO».

«PORQUE ME GUSTA VENIR Y TABIEN ESTOY MUY CONTENTO EN EL COLEGIO CON LOS PROFESORES Y YO NO ME QUISIERA MARCHAR DE AQUI PORQUE ESTOY MUY CONTENTA AUNQUE ME LO DIJERAN QUE SI QUISIERA SALIR YO DIRÍA QUE NO POR QUISIERA SEGRÍ EN EL COLEGIO Y ALLUDAR POR LO QUE AGA FALTA».

«VENGO A HACER COSAS Y A ESCRIBIR Y A ESTUDIAR LO PASO MUY BIEN. SALIMOS AL RECREO, VAMOS A LA PLAYA Y JUGAR SUSA NOS TOCA LA GUITARRA LEEMOS EL PERIÓDICO».

«VENGO AL COLEGIO PORQUE OBLIGAN PERO DE TODOS MODOS ME LO PASO BIEN LO QUE MENOS HECHAMOS ES AL MONO BUBU QUE SE MURIÓ LO QUE MAS ME GUSTA ES JUGAR POR EL PATIO JUGAR A JUEGOS DE MESA ESTUDIAR SALIMOS AL RECREO Y VEMOS A VECES LA TELEVISIÓN Y SE HACE GIMNASIA NOS DAN REGALOS LOS MAESTROS ESTAMOS EN CLASE UNA HORA Y SALIMOS AL RECREO DESPUÉS ENTRAMOS OTRA VEZ SE HACE A VECES EXCURSIONES FIESTAS SE COME EN EL COLEGIO A VECES HAY TEATRO HAY A VECES MÚSICA».

El número total de palabras utilizadas es de 661, de las que 268 son distintas entre sí. Después de realizar lo indicado en el paso 2 del análisis se obtienen 409 de las que sólo 48 son distintas. Como resultado del paso 3 se seleccionan 30 palabras, las cuales han sido utilizadas por los alumnos en 257 ocasiones.

Las palabras seleccionadas con su frecuencia son las siguientes:

TABLA 1

0 AÑOS	5
0 LEER	3
0 BIEN*	6
0 COSAS	26
0 PASO*	4
0 GENTE	3
0 GUAPO	3
0 PLAYA	3
0 HACER	7
0 JUGAR	7
0 GUSTA	4
0 LUGAR	4
0 MENOS	3
0 GUSTA*	12
0 FUTURO	13
0 PASO**	4
0 AMIGOS	9
0 NUEVAS	3
0 GRANDE	3
0 COLEGIO	48
0 ENSEÑAN	3
0 OBLIGAN	6
0 FIESTAS	3
0 RECREO*	7
0 ESTUDIAR	13
0 APRENDER	34
0 CONTENTO	4
0 TRABAJAR	9
0 ESCRIBIR	4
0 EXCURSIONES	4
total de palabras	257

El procedimiento de clasificación utilizado ha dado como resultado la elección de cinco clases, cuyas inercias inter e intra nos indican una suficiente diferenciación entre sí.

TABLA 2

	INERCIAS	TAMAÑO	PRECUEEN	DISTAN
INERCIA INTER-CLASES	1.5600			
INERCIAS INTRA-CLASE				
CLASE 1 / 5	.8407	15	65.00	.9244
CLASE 2 / 5	.6743	20	152.00	.2346
CLASE 3 / 5	.0757	2	6.00	16.1780
CLASE 4 / 5	.3936	4	23.00	3.9946
CLASE 5 / 5	.0046	2	11.00	10.5664
INERCIA TOTAL	3.5489			

La interpretación del significado de cada clase se realiza mediante aquellas palabras que son utilizadas con frecuencias significativamente distintas a las empleadas en el resto de las clases ($p < .05$).

TABLA 3

NUM . PALABRA	IDEN	PORCENTAJES			VAL. TEST	PROBA
		CLA/FRE	FRE/CLA	GLOBAL		
CLASE 1 / 5				25.29		
3 . BIEN*	BIEN	100.00	9.23	2.33	3.52	.0002
14 . GUSTA*	GUST	75.00	13.85	4.67	3.43	.0003
21 . ENSEÑAN	ENSE	100.00	4.62	1.17	2.15	.0156
17 . AMIGOS	AMIG	55.56	7.69	3.50	1.66	.0481
28 . TRABAJAR	TRAB	.00	.00	3.50	-1.48	.0691

La *primera clase* se caracteriza porque señalan que el colegio «está bien», les «gusta», les «enseñan» y «tienen amigos». Es, por tanto, una opinión global y positiva sobre la escuela. Este grupo lo forman 15 alumnos.

TABLA 4

NUM . PALABRA	IDEN	PORCENTAJES			VAL. TEST	PROBA
		CLA/FRE	FRE/CLA	GLOBAL		
CLASE 2 / 5				59.14		
26 . APRENDER	APRE	79.41	17.76	13.23	2.46	.0069
16 . PASO**	PASO	.00	.00	1.56	-1.93	.0269
1 . AÑOS	AÑOS	.00	.00	1.95	-2.30	.0107
3 . BIEN*	BIEN	.00	.00	2.33	-2.63	.0043
14 . GUSTA*	GUST	16.67	1.32	4.67	-2.77	.0028

La *segunda clase* es la más numerosa, 20 alumnos, y se caracteriza, fundamentalmente, porque dicen que van al colegio «a aprender» y no hacen ninguna referencia afectiva respecto a «pasar bien» o «pasar muy bien», lo que puede considerarse como una percepción centrada en lo académico.

TABLA 5

NUM . PALABRA	IDEN	PORCENTAJES			VAL. TEST	PROBA
		CLA/FRE	FRE/CLA	GLOBAL		
CLASE 3 / 5				2.33		
27 . CONTENTO	CONT	75.00	50.00	1.56	4.03	.0000

La *tercera clase* se distingue porque indican, exclusivamente, que «van contentos al colegio», lo cual nos permite identificarlos como grupo en el que predomina una opinión de tipo afectivo.

TABLA 6

NUM . PALABRA	IDEN	PORCENTAJES			VAL.TEST	PROBA
		CLA/PRE	PRE/CLA	GLOBAL		
CLASE 4 / 5		C4			8.95	
8 . PLAYA	PLAY	100.00	13.04	1.17	3.22	.0006
9 . HACER	HACE	57.14	17.39	2.72	2.98	.0014
2 . LEER	LEER	66.67	8.70	1.17	2.02	.0218
16 . PASO**	PASO	50.00	8.70	1.56	1.74	.0412
26 . APRENDER	APRE	.00	.00	13.23	-1.84	.0326

La característica de la *cuarta clase* es la utilización, casi unánime, de la palabra «playa». También aparecen las palabras «hacer», «leer», «escribir» y «pasarlo muy bien» y no utilizan nunca la palabra «aprender», al igual que sucedía en la clase anterior. Los alumnos que integran esta clase se diferencian del resto por su referencia al enclave geográfico-marítimo en que está situado el centro y por las actividades recreativas que realizan en él. La suya es una visión distendida y concreta de la escuela.

TABLA 7

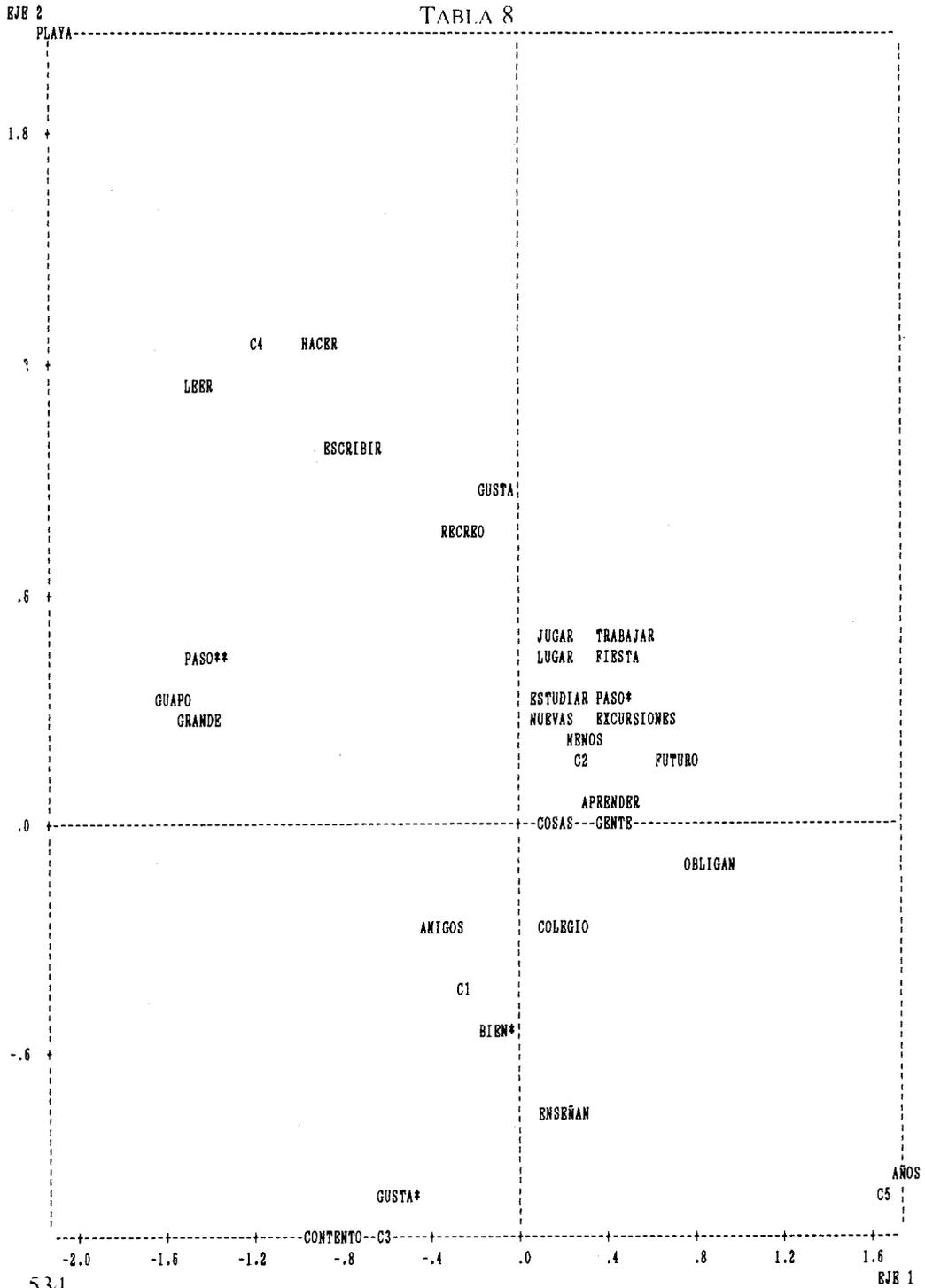
NUM . PALABRA	IDEN	PORCENTAJES			VAL.TEST	PROBA
		CLA/PRE	PRE/CLA	GLOBAL		
CLASE 5 / 5		C5			4.28	
1 . AÑOS	AÑOS	100.00	45.45	1.95	5.32	.0000

La *clase quinta* se caracteriza por una referencia a los «años» que han estado, o que estarán, escolarizados.

Del análisis efectuado se desprende que las clases tres y cinco están definidas por el empleo de un menor número de palabras —una o dos básicamente— que las utilizadas en las clases restantes (en éstas todas las palabras resultaron significativas); además, estas dos clases agrupan un mayor número de alumnos, por lo que han resultado más homogéneas que las clases uno, dos y cuatro.

Los resultados comentados se resumen en la siguiente gráfica.

TABLA 8



CONCLUSIÓN

Aunque el análisis debería ser guiado por el marco teórico (explícito) del investigador y por los propios datos, de hecho, las metodologías cualitativas, al requerir un alto grado de flexibilidad, pueden hacer que elementos subjetivos produzcan una imagen desfigurada de la realidad. Precisamente el uso de ordenadores permite reducir sensiblemente la posibilidad de que el investigador realice una sobrevaloración de algunos aspectos no apoyada en los datos. De esto no se deriva subordinación o rigidez mecánica, puesto que el ordenador nos permite explorar y ensayar sin alejarnos de la evidencia de los datos. Como señala Huber (1988:6), la utilización de los ordenadores en la investigación cualitativa fomenta «la comunicación, la reconstrucción y el control de los procesos y sus resultados».

El programa que hemos utilizado posee algunas características que lo sitúan a la altura de otros más utilizados en el análisis cualitativo (e.g. el HIPERCARD, el NUDIST, etc.). No presenta limitaciones en el tamaño de los registros ni en el número de categorías, tampoco exige determinar previamente las hipótesis. Permite trabajar con palabras, textos y/o ficheros. Se pueden tratar múltiples casos en un mismo fichero y considerar después cada caso como un segmento. Permite combinar las características de los textos con cualquier otra información de los sujetos —por ejemplo, variables demográficas—, haciendo posible una aplicación directa de paquetes estadísticos típicos. Todo esto hace al SPAD una herramienta de gran utilidad para analizar datos procedentes de entrevistas no estructuradas, notas de observación participante, protocolos, documentos, preguntas abiertas de cuestionarios, etc.

En relación con el ejemplo presentado sobre las opiniones escolares del alumnado con necesidades especiales, los análisis directos sobre los textos permiten apreciar un predominio de expresiones tópicas y un lenguaje rígido y redundante. Los resultados obtenidos detectan la presencia de dos grandes grupos que se diferencian por su alusión al colegio como un lugar para «aprender» o bien como algo que les «gusta».

Pensamos que el programa utilizado ha sido útil en nuestro análisis y, si bien habrá de perfeccionarse en sucesivos trabajos, su eficacia se hará más patente a la hora de analizar grandes muestras de datos.

Mediante el SPAD podemos clasificar y analizar fragmentos de textos, pero ¿ello nos da un conocimiento de la totalidad de los datos?, ¿cómo podemos trascender los datos parciales? En adelante necesitamos buscar sistemas que nos permitan realizar un análisis más holístico. Nos encontramos aún en la fase de reducción de datos. El SPAD nos puede ayudar a codificar y categorizar información. Pero, ¿puede el SPAD ayudarnos a descubrir temas, a generar y contrastar hipótesis, a construir teorías? Existen todavía importantes dilemas: ¿segmentación previa de los textos frente a segmentación emergente?, ¿búsqueda mecánica frente a posibilidades de exploración teórica?, ¿potencia frente a flexibilidad?, ¿características generales frente a análisis detallado? En definitiva, se trata de una nueva formulación del debate cuantitativo-cualitativo.

BIBLIOGRAFÍA

- BERELSON, B. (1971): *Content Analysis in Communication Research*. New York, Hafner Pub. Co.
- HUBER, G. L. (1988): Análisis de Datos Cualitativos: La aportación del Ordenador. Sevilla, *Reunión Científica sobre Avances en el Estudio del Pensamiento del Profesor*, págs: 1-16.
- LEBART, L. (1982): *SPAD. System portable pour l'analyse des données*. Paris, Cesisio.
- MILES, M. Y HUBERMAN, A. (1984): *Qualitative Data Analysis. A Sourcebook of new methods*. California, Beverly Hills, Sage Publications.
- RICHARDS, L. y RICHARDS, T. (1987): Qualitative Data Analysis: Can Computers Do It? *Australian and New Zealand Journal of Sociology*, 23, 1:23-35.
- SÁNCHEZ VALLE, I. (1989): La metodología del análisis de contenido aplicada a las ideas expresadas en los prólogos de diferentes textos de pedagogía general. *Bordon*, 41, 1:127-152.
- SAN FABIÁN, J. L., CORRAL, N., FERNÁNDEZ, S. Y MARTÍNEZ, R-A. (1989). *Percepción y representaciones de la realidad escolar en el alumnado del Ciclo Superior de E.G.B.* Madrid, CIDE.