
TRABAJOS METODOLÓGICOS

ESTUDIO COMPARATIVO DE DOS PROGRAMAS INFORMÁTICOS DE ANÁLISIS DE DATOS CUALITATIVOS

por

M^a Paz Sandín Esteban

José Luis Medina Moya

Facultad de Pedagogía. Universidad de Barcelona

RESUMEN

En este artículo se presenta un estudio comparativo de dos programas de análisis de datos cualitativos (Aquad y Ethnograph).

Se examinan las características de ambos programas, describiendo sus ventajas e inconvenientes y los procedimientos específicos que cada uno de ellos utiliza para el análisis de este tipo de datos.

Asimismo, se presentan algunos ejemplos de cómo los programas de análisis de datos cualitativos pueden ayudarnos en la tarea de organizar e interpretar los datos provenientes de transcripciones de entrevistas, notas de campo, diarios, etc.

ABSTRACT

This paper shows a comparative study about two programs (Aquad and Ethnograph).

The characteristics of both are showed, discovering its advantages and objections and the specific proceedings that they use for the analysis of qualitative data.

Likewise, some examples are introduced to show how the qualitative analysis programs can help us in the labour of organizing and interpreting the data proceeding from different founts.

1. INTRODUCCIÓN

Todavía es reciente la época en la que el investigador que utilizaba técnicas de recogida de datos como la observación participante o la entrevista en profundidad, debía enfrentarse al análisis y valoración de una cantidad ingente de datos: hojas y hojas obtenidas a partir de la transcripción de entrevistas, numerosas notas de campo pertenecientes a sesiones de observación, diarios, extractos de periódicos, etc. En algunos casos, parecía inabordable la posibilidad de organizar tan múltiples y variados materiales y, en no menos ocasiones, el investigador se desalentaba ante el maremágnum de información. En esta situación, los investigadores utilizaban su imaginación, creatividad y experiencia para organizar los datos que habían obtenido: tijeras y cartulinas, fichas y colores, constituían las únicas herramientas de reducción y análisis de datos en las investigaciones de tipo cualitativo. Así pues, se hacía imprescindible dedicar una gran cantidad de tiempo a la organización de los materiales antes de poder iniciar cualquier análisis.

En la actualidad, existen programas informáticos concebidos para auxiliar al investigador en la tediosa tarea de organizar sus datos. De esta manera, al pasar los aspectos mecánicos del trabajo al ordenador, el investigador puede invertir más tiempo y atención en los aspectos de interpretación, análisis y valoración crítica de los resultados.

El presente trabajo pretende ofrecer una aproximación a dos programas informáticos de análisis de datos cualitativos (Ethnograph y Aquad) valorando los aspectos positivos y negativos de cada uno de ellos. Debido a que ya existe en la literatura escritos referentes al programa Aquad¹, nuestro interés se centrará de forma especial en dar a conocer las posibilidades que Ethnograph nos aporta en este tipo de análisis.

No es el objetivo del estudio profundizar en las técnicas de análisis cualitativo de datos, ni ser un manual técnico que recoja todos los pormenores del funcionamiento de los programas. Sin embargo, sí que nos gustaría que este trabajo sirviera de introducción a aquellas personas que quieren realizar un análisis de este tipo y no conocen cuáles son las ventajas, utilidades y posibilidades que ofrecen los programas informatizados en este campo.

2. BREVE HISTORIA DE ETHNOGRAPH

El programa de análisis de datos cualitativos, Ethnograph, *nació* cuando dos estudiantes, Jack Clark y John Seidel, estaban colaborando en un estudio etnográfico sobre el tratamiento de pacientes que sufrían síndrome de dolor crónico (Universidad de Colorado, 1980). Mientras trabajaban en este proyecto, Jack y John pensa-

1 GIL, J. et al. *Formulación y comprobación de hipótesis con ayuda del programa AQUAD*. Comunicación presentada al VI Seminario de Modelos de Investigación Educativa. Madrid.

ban cómo los ordenadores podían ayudarles en sus necesidades como analistas de datos cualitativos.

Estudiaron y aprendieron etnografía con los profesores Edward Rose y Rolf Kjolseth. Sobrevivieron a las transcripciones de los datos mediante máquina de escribir y su posterior tratamiento de “cortar y pegar” a través de múltiples fotocopias de los datos. Durante esta experiencia ambos llegaron a la conclusión de que tenía que existir una forma mejor de hacerlo. Ese interés por la potencialidad que podía ofrecer el ordenador en este tipo de tareas tuvo como resultado el desarrollo, por parte de científicos sociales, del programa de análisis de datos cualitativos, Ethnograph.

Ethnograph está diseñado para **ayudar** a los investigadores en algunos de los **aspectos mecánicos** del análisis de datos. Siguiendo la clasificación de Tesch (1990), el programa Ethnograph pertenecería al grupo de programas informáticos encaminados al *análisis interpretativo* de datos, mientras que el programa Aquad pertenecería al grupo que se ha denominado de *construcción de teorías*.

3. FASES EN EL PROCESO DE TRATAMIENTO DE DATOS

(utilizando un programa informático)

A continuación vamos a mostrar cuáles son los pasos o etapas que debemos superar desde la elección de la técnica de recogida de datos y su aplicación, pasando por el tratamiento informático, hasta la generación de resultados e interpretación.

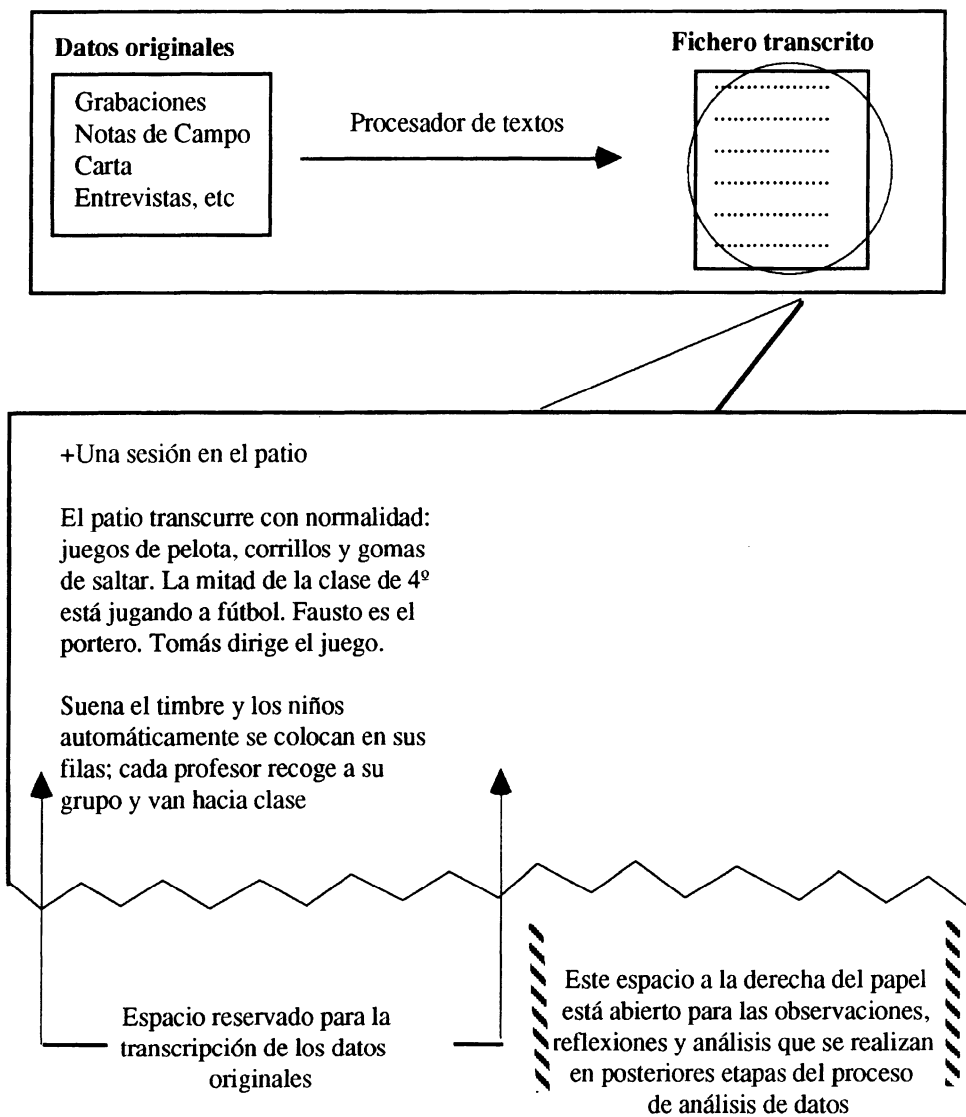
Recogida de datos

En primer lugar deberemos obtener los datos que posteriormente queremos analizar. Éstos, pueden haberse obtenido a través de diferentes estrategias y, por tanto, podremos trabajar con notas de campo, entrevistas, abstracts, diarios, periódicos, historias orales, memorias, cartas, grabaciones,

Transcripción

Antes de que cualquier programa de tratamiento de datos pueda ayudarnos a poner en orden nuestros materiales, debemos introducirlos en el ordenador, transcribiendo las cintas grabadas o reproduciendo los datos obtenidos en sesiones de observación en un procesador de textos, siguiendo unas mínimas reglas de transcripción. Si se trabaja con entrevistas o sesiones de observación recogidas en diferentes días, se recomienda que cada entrevista o sesión se introduzca en ficheros independientes para facilitar el tratamiento posterior de los datos.

Transcripción de datos



Conversión

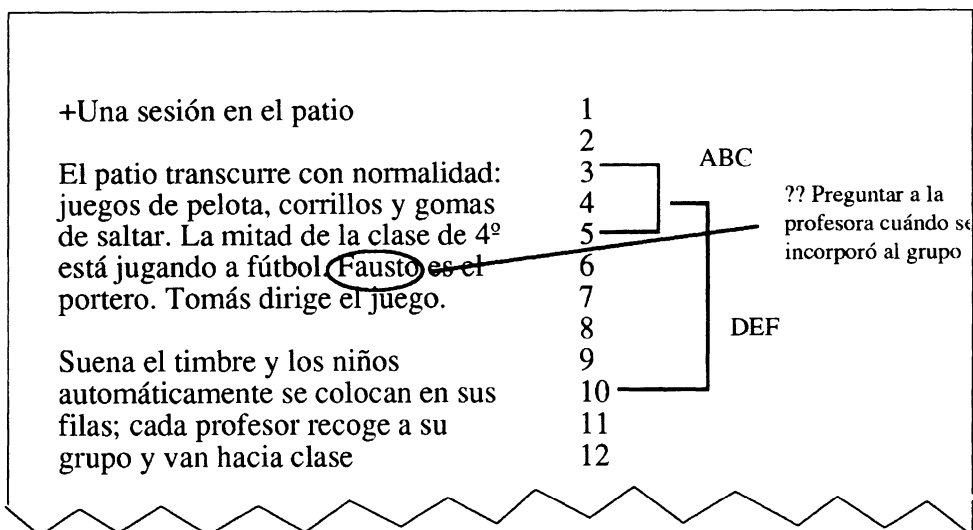
El siguiente paso consiste en convertir el fichero transcrito mediante un procesador de textos (Word Perfect u otros) en un fichero standard tipo ASCII, puesto que es el único formato que *entiende* el programa.

Numerar e imprimir

Una vez hemos obtenido el fichero ASCII ya podemos trabajar con el programa informático de tratamiento de datos: numeraremos las líneas del texto y lo imprimiremos

Asignación de códigos

Una vez se ha impreso el documento numerado, podemos reflejar libremente en él nuestros comentarios y reflexiones, codificar y analizar el texto. El texto numerado es como una hoja borrador en la que podemos subrayar, tachar, resaltar, ... Tanto si pensamos en números, como palabras o dibujos, podemos usar la parte derecha de la página como un lugar de trabajo para plasmar nuestras impresiones.



Introducción de códigos en el programa

Los programas de análisis cualitativo permiten introducir, modificar e imprimir, con gran facilidad, los esquemas de codificación que hemos hecho sobre los textos. Cada programa posee una forma particular de trabajar con los códigos y, por ello, se remite al lector/a a que consulte en los próximos apartados las características concretas de cada uno de los programas que aquí se analizan.

Búsqueda de códigos y análisis de datos

Es en este momento, cuando el programa de análisis de datos nos muestra toda su potencia a la hora de seleccionar los fragmentos de texto que habíamos categori-

zados según determinados códigos. Se nos permitirá hacer búsquedas simples (un sólo código) y búsquedas múltiples (dos o más códigos).

También en este apartado difieren los programas en cuanto a funcionamiento y posibilidades que ofrecen para el análisis de datos.

Interpretación

Es en este apartado donde el ordenador deja de «ayudarnos» puesto que es el investigador el que a través de los datos que ha obtenido mediante la búsqueda y combinación de códigos y, a la luz de sus perspectivas teóricas, quien debe ofrecer una interpretación y/o significación de los resultados obtenidos así como aventurar hipótesis explicativas.

4. CARACTERÍSTICAS Y FUNCIONAMIENTO DE LOS PROGRAMAS

Características técnicas

Aquad:

Versión: 3.0

Ordenador: IBM-PC's, XT's, AT's

Para su instalación es indispensable la disponibilidad de un disco duro con alrededor de 1 Mb de memoria libre. El programa instalado ocupa aproximadamente 717 Kb.

El programa funciona según un sistema de ventanas.

Ethnograph

Versión: 3.0

Ordenador: IBM-PC's, XT's, AT's compatibles con DOS versión 2.0 o superior.

Requiere un mínimo de 256 Kb de memoria RAM. Se instala en disco duro ocupando un espacio aproximado de 313 Kb.

El programa está construido en su totalidad según el principio de ventanas.

Normas de transcripción

Aquad:

El programa Aquad admite un máximo de 48 caracteres por línea. Al final de cada fichero de datos es obligatorio escribir: ^Z. En cuanto a la transcripción de datos, no presenta ningún tipo de limitación o normativa explícita.

Ethnograph

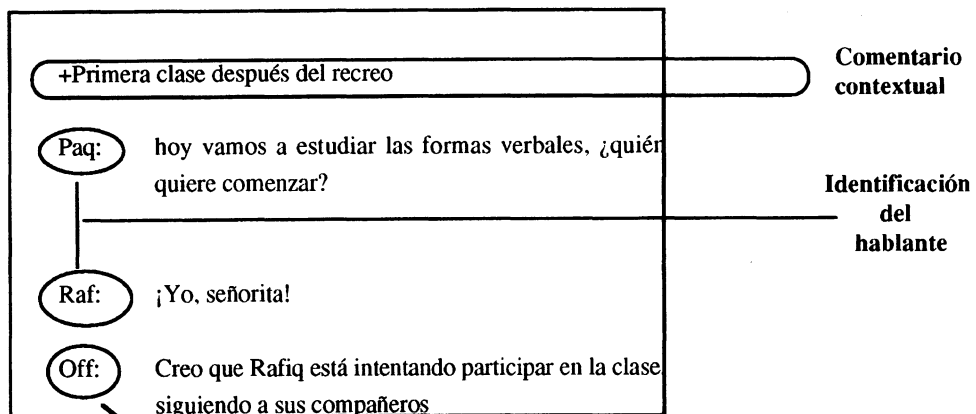
Admite un máximo de 40 caracteres por línea. Vamos a dedicar un breve espacio a la introducción de datos en este programa puesto que presenta algunas normas que

son imprescindibles cumplir para el buen funcionamiento posterior de programa:

Restricciones en la columna 1

Si observamos la siguiente figura veremos que la **identificación del hablante** empieza en la columna 1 y que las líneas del texto empiezan algunas columnas más hacia la derecha. También se puede observar una línea de **comentario contextual** que comienza con el símbolo “+” en la columna 1

La identificación del hablante y la línea de comentario contextual *son los únicos tipos de información* que pueden empezar en la columna 1



(Puede utilizarse una identificación determinada, por ejemplo "Off", que recoja pensamientos, sensaciones, comentarios, ...)
Posteriormente, podremos obtener un listado de todas las apreciaciones personales que hemos efectuado durante la sesión de observación o la entrevista

Comentarios contextuales (CC):

- Sólo pueden introducirse cuando se están preparando los datos. No se pueden entrar después de haber utilizado el procedimiento de numeración de líneas .
- Un comentario contextual se reconoce por el símbolo “+” en la columna 1.
- No puede tener más de 35 letras, espacios o caracteres.

Utilidad de los comentarios contextuales:

1. Un comentario contextual ayuda a identificar el contexto general de un segmento cuando es extraído del fichero de datos.
2. Se pueden utilizar varios comentarios contextuales a lo largo del documento.

3. El uso de un nuevo comentario contextual, marca el fin del anterior CC y tendrá efecto hasta que sea introducido un nuevo CC.

4. Un comentario contextual no es un tipo de código que se pueda utilizar en la selección de códigos. Simplemente es una nota informativa que se introduce en los datos.

Ejemplo:

Imagine una entrevista que se desarrolla entre Vd. y la persona informante. Durante la entrevista una tercera persona, por ejemplo, la esposa del entrevistado, se “añade” inesperadamente a su conversación. La presencia de esta tercera persona puede influir en los comentarios que el entrevistado hará y Vd. quiere tener constancia de ello cuando analice sus datos. Un comentario contextual (+La esposa del Sr. López ha entrado en el despacho) puede ser insertado en la transcripción de datos para señalar el momento en que la esposa entra en escena. Este comentario precederá cada segmento codificado extraído del fichero de datos que pertenezca a ese periodo de tiempo (el comentario contextual continuará vigente hasta que se introduzca otro comentario contextual que indique un cambio de situación, de actividad)

Numeración del fichero original

Aquad

Presenta la numeración de líneas a la izquierda del texto. También numera las líneas en blanco.

Ethnograph

El fichero numerado presenta las líneas a la derecha del texto. Las líneas en blanco también «cuentan» pero no se numeran.

Ventajas: el hecho de aparecer las líneas a la derecha facilita la definición de códigos, puesto que es la derecha del texto original, en el espacio abierto a las deliberaciones posteriores a la toma de datos, donde se suele hacer «a mano» la primera codificación.

Cabeceras

Veamos cuál es el formato de presentación de los listados en ambos programas

Aquad

Se presenta un listado con el texto original, numeradas las líneas, pero no hay ninguna indicación de lo que acaba de salir por la impresora

Ethnograph

Aparece el texto numerado y una cabecera en cada página indicanco qué tipo de información aparece en el listado: si es una búsqueda de un código simple o múltiple, si es un listado de conteo de frecuencias de un determinado código, hora, día y número de página. La ventaja de este formato de presentación es que estas notas informativas facilitan el trabajo cuando se empieza a manejar gran cantidad de información impresa y necesitamos saber a qué tipo de listado pertenecen.

Vamos a utilizar las notas de campo de una sesión de observación llevada a cabo en una escuela, para ejemplificar las características que hemos comentado anteriormente.

En primer lugar, se presenta el listado que ofrece el programa Aquad cuando le pedimos que nos liste todos aquellos fragmentos categorizados con el código IP (Idioma empleado por la profesora), en una sesión de observación.

43	Abrimos el libro de lenguaje (la profesora es catalana y se dirige a los niños, en esta clase, en castellano)
44	
45	
46	
99	La fecha en la pizarra está escrita en castellano
100	
216	Ahora habla en catalán, supongo que ha comenzado la clase de catalán
217	
404	Las preguntas están en catalán
<hr/> Código IP se utilizó 4 veces en texto 001	

Veamos, ahora, cuál es el formato de presentación que nos ofrece Ethnograph. Nótese la información que aportan los comentarios contextuales (precedidos del símbolo «+») para recontextualizar el segmento extraído. También se informa convenientemente si el fragmento extraído, está codificado con uno o más códigos o pertenece a otro fragmento.

SORTED OUTPUT FOR FILE 1ASCII 7/22/1993 19:38		Page 1
SC+BIG PICTURE:+IP		
1ASCII	+Primera observación (mañana)	
SC:	+IP	
#-IP	#-PC1	
:	Abrimos el libro de lenguaje (la	43 -#
:	profesora es catalana y se	44 #
:	dirige a los niños, en esta	45 #
:	clase, en castellano)	46 #
1ASCII	+Miércoles 21, sesión mañana	
SC:	+IP	
#-IP		
:	La fecha en la pizarra está escrita	99-#
:	en castellano	100-#
1ASCII	+Corrección 2º ejercicio de lengua	
SC:	+IP	
#-IP		
:	Ahora habla en catalán, supongo que	216 -#
:	ha comenzado la clase de catalán	217 -#
1ASCII	+Examen de sociales	
SC:	+IP	
#-IP		
:	Las preguntas están en catalán	404-!

Inserción de códigos

Aquad

Características de los códigos

Pueden introducirse hasta un máximo de diez códigos en una misma línea. La longitud de un código, por defecto, es de tres caracteres aunque esta característica se puede modificar.

Niveles de solapamiento o de contención de códigos

No se especifican límites

Formato pantalla de inserción

La inserción de códigos en el programa Aquad puede realizarse de dos formas:

a) Inserción de códigos sobre el fichero original: aparece en pantalla el fichero original y se codifica directamente sobre los datos. Un inconveniente de este sistema, es que si de un código al siguiente existe mucho espacio, debemos pasar irremediablemente *por encima* de todo el texto.

b) A través de la pantalla de inserción: si utilizamos este sistema, en la pantalla no nos aparecerá el texto original, sino un módulo especial del programa en el que se introduce:

1º Línea de inicio de código

2º Línea de finalización de código

3º Código

La ventaja de este sistema es la rapidez. Si ya se han trabajado los textos numerados y se han realizado las codificaciones sobre el papel, no es necesario volver a los datos, sino simplemente dar al programa la información sobre el tipo de código, línea de inicio y finalización.

Ethnograph

Características de los códigos

Pueden introducirse hasta un máximo de doce códigos que empiecen en la misma línea y que definan un fragmento. El código puede tener una longitud de hasta diez caracteres sin necesidad de modificar ningún parámetro del programa.

Niveles de solapamiento o de contención de códigos

Ethnograph admite un máximo de siete niveles de solapamiento o anidación. Ninguna línea del texto puede pertenecer a más de siete segmentos.

Inserción de códigos

No contempla la opción de trabajar sobre el fichero original. Los códigos se introducen a través de un módulo especial del programa. Este sistema, permite una

rápida codificación de los códigos, incluyendo la posibilidad de visualizarlos con el fin de confirmar que el código que se ha introducido y las líneas inicial y final que definen el fragmento son correctas.

Búsqueda de segmentos

La búsqueda de segmentos codificados es el corazón de un programa de análisis de datos cualitativos. Es en este momento cuando se seleccionan los fragmentos que habían sido codificados en los datos originales y se trasladan a un nuevo contexto: el contexto de las *cosas similares*.

Posibilidades de búsqueda

Aunque cada programa tiene su forma particular de funcionar, ambos ofrecen las siguientes posibilidades de búsqueda:

1. Se pueden efectuar búsquedas basadas en un sólo código. El criterio de localización de un segmento es un único código.
2. Podemos buscar fragmentos etiquetados con más de un código:
Ethnograph permite búsquedas combinando hasta 5 códigos mediante los operadores «and» y «not». Por ejemplo, podrían efectuarse las siguientes búsquedas:
 - CódigoA y CódigoB
 - CódigoA *pero no* CódigoB
 - CódigoA y CódigoB *pero no* CódigoC

A modo de ejemplo se muestra a continuación el listado que ofrece el programa Ethnograph cuando solicitamos una búsqueda de todos aquellos fragmentos que estén codificados con los códigos PMF (dificultades observadas en los niños/as) y PR (rendimiento observado):

SORTED OUTPUT FOR FILE 1ASCII 7/22/1993 19:38 Page 2		
SC-BIG PICTURE: +PMF +PR		
1ASCII	+Actividad con el libro de lenguaje	
SC: +PMF +PR		
#-PR	#-PMF	
:	del verbo cantar. El siguiente	51 -#
:	en ser preguntado es Rafiq	52 #
:	que no sabe contestar a la pregunta	53 -#
1ASCII	+Corrección oral de ejercicios	
SC: +PMF +PR		
#-PR	#-PMF	
:	Rafiq tiene dificultades para	117 -#
:	encontrar las formas	
	118 #	
:	verbales adecuadas	119 -#
1ASCII	+Clase de sociales	
SC: +PMF +PR		
#-PR	#-PMF	
:	Fausto se equivoca varias	117 -#
:	veces (me sorprende, porque en su	118 #
:	primera intervención leyó de	119 #
:	forma más fluida)	120 -#

3. Los listados de códigos pueden obtenerse en el orden en que han sido codificados los datos o en orden alfabético

4. Podemos obtener listados con la frecuencia de aparición de códigos en un fichero determinado. Este tipo de listados, aunque muy sencillos, pueden aportar gran cantidad de información. A continuación, se muestra a modo de ejemplo, los códigos que se utilizaron en tres sesiones de observación realizadas en una escuela, ordenados según la frecuencia de aparición (Ethnograph)

FREQUENCY List of codewors used in coding 1ASCII 9/25/1993 22:06							
N	CODEWORD	N	CODEWORD	N	CODEWORD	N	CODEWORD
17	A3	13	PR	12	PMF	10	RAI
6	A1	5	A4	4	IP	4	PMT
4	PMI	3	RB	2	DI	1	ROI
FREQUENCY List of codewors used in coding 2ASCII 9/25/1993 22:06							
N	CODEWORD	N	CODEWORD	N	CODEWORD	N	CODEWORD
13	A3	10	PMI	8	PR	5	A4
4	I	4	PMF	3	DN	3	IP
2	PMR	2	DC	1	A	1	RAI
FREQUENCY List of codewors used in coding 3ASCII 9/25/1993 22:06							
N	CODEWORD	N	CODEWORD	N	CODEWORD	N	CODEWORD
5	A3	4	A4	3	PR	3	PMI
1	RPI	1	RAI	1	A8	1	DN

Analizando estos listados, podemos observar cuáles han sido los códigos que tienen una frecuencia mayor y que, por tanto, han acaparado el foco de atención del observador. Por ejemplo, el código A3 se utiliza en 17 ocasiones en la primera sesión de observación y va disminuyendo su frecuencia de aparición paulatinamente a lo largo de las sesiones. En cambio, el código RAI que en la primera sesión aparece diez veces, sólo lo hace en una ocasión en las siguientes sesiones. También, podemos obtener información de qué códigos aparecen en determinadas observaciones y no lo hacen en otras. ¿Por qué no aparecen? Quizá, han pasado desapercibidos para el observador o realmente no se han producido. Todas estas valoraciones sería muy difícil realizarlas en una lectura lineal de las notas de campo, mientras que el resumen de estos listados nos aporta nuevas visiones y perspectivas que, de otro modo, quedarían diluidas en los textos originales.

5. UTILIZACIÓN DE VALORES DE BASES DE DATOS

Este tipo de búsqueda es muy interesante y la encontramos en ambos programas. Consiste en caracterizar los ficheros como si fueran registros de una base de datos.

Cada fichero puede contener, además del texto correspondiente, por ejemplo, a la sesión de observación o la entrevista, datos referentes a la persona que ha proporcionado esa información (edad, sexo, profesión, etnia...)

Esta opción podría ser útil en el análisis de entrevistas donde tendríamos la posibilidad de, por ejemplo, pedir al programa:

«lista todos los segmentos caracterizados con el código “relación escuela-familia” de los alumnos entre 5 y 12 años, de sexo femenino y pertenecientes a la etnia pakistaníes»

5. VALORACIÓN GLOBAL DE LOS PROGRAMAS

Aunque ambos programas pueden ofrecernos resultados similares en cuanto al tratamiento de los datos, existen determinados aspectos en los que uno y otro deberían mejorar para ofrecer al usuario mayor versatilidad y facilidad de manejo. A continuación se ofrece una valoración global de los programas Aquad y Ethnograph. Después de haber leído este análisis comparativo queda a elección del lector cuál puede ser el más adecuado para el tratamiento de sus datos teniendo en cuenta cuáles son sus objetivos, intenciones de análisis, etc.

AQUAD

Recordemos que este programa pertenece a la categoría de programas dirigidos a la construcción de teorías (theory-building). Para ello, Aquad dispone de un módulo especial (Ethnograph no lo posee) denominado «Hipótesis», mediante el cual podemos plantear *relaciones hipotéticas* que queremos contrastar utilizando los datos obtenidos en entrevistas o sesiones de observación. Para ello, se formula una relación sospechada entre las unidades de significado de un texto como premisa, dejando posteriormente al ordenador que compruebe en los ficheros, si se cumplen las relaciones establecidas en la afirmación. Todas las hipótesis se basan en relaciones según una determinada *distancia* entre los códigos.

El manual del funcionamiento del programa está disponible en castellano, lo que facilita el trabajo de introducción del usuario. Sin embargo, la estructura del programa es muy rígida y el paso obligado por determinados menús enlentece el trabajo. También, hemos podido constatar que en determinadas ocasiones (no sabemos porqué) el programa se *bloquea* y obliga al usuario a apagar el ordenador lo que, una vez más, enlentece el proceso.

Los listados que ofrece este programa (búsqueda de códigos, frecuencias de aparición de códigos, segmentos pertenecientes a más de un código, etc.) están excesivamente simplificados, es decir, se ofrece la información que se ha solicitado, pero se olvidan algunos datos que pueden enriquecer en gran medida la interpreta-

ción de los resultados obtenidos (códigos solapados, comentarios contextuales, identificación del hablante, etc.)

ETHNOGRAPH

El programa Ethnograph no posee los procedimientos especiales de *generación de teorías* que presenta Aquad. En realidad, no es el objetivo de este programa la contrastación o verificación de relaciones hipotéticas, sino que ha sido diseñado específicamente para realizar análisis de tipo descriptivo/interpretativo (de-contextualización y re-contextualización, utilizando los términos de los creadores de Ethnograph).

El manual de este programa se presenta en inglés. Este manual constituye un material didáctico perfectamente elaborado y pensado para que el usuario, al mismo tiempo que se inicia en el manejo del programa, vaya descubriendo, a través de ejemplos, las posibilidades y potencia que el programa le ofrece para el análisis de datos cualitativos.

En cuanto al funcionamiento general del programa, cabe destacar la sencillez con la que se presentan los diferentes menús que en todo momento guían al usuario en la utilización de los diferentes procedimientos.

Asimismo, cabe resaltar la excelente presentación de los listados en los que se ofrece gran cantidad de información adicional.

BIBLIOGRAFÍA

- HUBER, G. L. (1991). *AQUAD. Análisis de datos cualitativos con ordenadores*. Sevilla, Carlos Marcelo
- SEIDEL, J. V., KJOLSETH, R. y SEYMOUR, E. (1988). *The Ethnograph: A User's Guide*. Qualis Research Associates.
- TESCH, R. (1990). *Qualitative Research: Analysis and Software Tools*. Bristol, The Falmer Press.