

Sociolingüística cualitativa y lingüística informática

ANXO M. LORENZO SUÁREZ
Universidade de Vigo

Publicado en *Panorama de la investigación en lingüística informática* (editado por J. Gómez Guinovart, A.M. Lorenzo Suárez, J. Pérez Guerra & A. Álvarez Lugrís), Logroño, Revista Española de Lingüística Aplicada, 1999, pp. 247-61.

Resumen

El análisis de los datos sociolingüísticos cualitativos puede beneficiarse del uso de diferentes recursos informáticos que son de interés para cuatro tareas principales de investigación: la organización del material cualitativo, su descripción y etiquetado, la interpretación de los datos, y la presentación gráfica de los resultados. Para ello disponemos de programas genéricos y de programas específicos para el análisis cualitativo, los cuales, a pesar de no ser diseñados específicamente para el tratamiento de datos sociolingüísticos, pueden ser empleados para las tareas de análisis propias de ese objeto de estudio.

Abstract

The analysis of sociolinguistic qualitative data can benefit from the use of different computer resources which are useful for four main investigation tasks: the organization of qualitative material, their description and labelling, interpretation of data and display of results. For that purpose general and specific software can be used for qualitative data management. Even though these sorts of software have not been specifically designed for sociolinguistic data, they can be applied to sociolinguistic analysis tasks.

Palabras clave: sociolingüística, análisis de datos cualitativos, lingüística informática.

Keywords: sociolinguistics, QDA (Qualitative Data Analysis), computational linguistics.

1.- INTRODUCCIÓN

Pocas dudas nos planteamos aquellos que nos dedicamos a la investigación de las realidades sociales acerca del interés y la rentabilidad que nos presentan las aplicaciones informáticas en determinadas fases de un protocolo experimental de investigación. Si en nuestras investigaciones empleamos algún tipo de datos, sean estos cualitativos o cuantitativos, su ordenación, manejo y análisis se convierten en una tarea adecuada para recibir algún tipo de implementación informática.

Una disciplina como la sociolingüística, al igual que buena parte de aquéllas englobadas tras la etiqueta de lingüística aplicada, se beneficia de los avances en las nuevas tecnologías. Si en el plano teórico y epistemológico la sociolingüística se nutre de experiencias y de desarrollos conceptuales generados por la lingüística, la sociología, la psicología o la antropología, en el plano metodológico la sociolingüística se beneficia de los avances producidos principalmente en la sociología y la psicología social –en lo relativo a los métodos de muestreo y en los instrumentos de medida de las realidades sociales–, así como en la estadística, y, por supuesto, hoy en día, en la informática.

Teniendo en cuenta esa situación, consideramos que el papel de la informática en la investigación sociolingüística debe centrarse en servir de soporte en determinadas fases de investigación. La informática tiene una amplia rentabilidad metodológica para el tratamiento del material sociolingüístico en las tres etapas de cualquier diseño de investigación: la preparación del material, la ordenación de los datos conseguidos, y, finalmente, el análisis de tales datos. De esta forma, la ordenación, codificación, manejo y análisis del material sociolingüístico cualitativo y cuantitativo, generado en un determinado protocolo de investigación, se presentan como las tareas más óptimas para ser implementadas mediante diversos recursos informáticos, como analizaremos en los próximos apartados.

Teniendo en cuenta esa situación, el objetivo de este trabajo se centra en la consideración de los distintos elementos de análisis que la informática puede aportar en los diseños sociolingüísticos de investigación de carácter cualitativo. Para ello dividiremos en tres partes el contenido de este capítulo. En primer lugar, pondremos de relieve la naturaleza de los datos cualitativos en un campo tan diverso como el sociolingüístico. En segundo lugar, haremos referencia a las tareas de investigación que se pueden implementar electrónicamente contando con datos de tipo cualitativo. Y en tercer lugar analizaremos el contenido de los diferentes tipos de programas informáticos disponibles.

2.- CONTENIDO Y ALCANCE DEL ANÁLISIS DE DATOS SOCIOLINGÜÍSTICOS CUALITATIVOS

Las técnicas de análisis cualitativo están formadas por los métodos empleados para el estudio de la “información ordinaria”, es decir, la amplia gama de textos orales (Bryman y Burgess 1994; Denzin y Lincoln 1994; Kelle 1995; Miles y Huberman 1994; Sarabia y Zarco 1997). Los métodos cualitativos se están propagando con celeridad a todos los campos del saber, debido a que permiten manejar, describir y explorar datos complejos, poco o nada estructurados, a la vez que permiten crear y probar nuevas teorías e ideas.

Hoy en día, cualquier objeto de estudio puede ser analizado con el análisis cualitativo. Apoyada en métodos rigurosos, la investigación cualitativa dispone de técnicas innovadoras y poderosas para la exploración de la información no estructurada sin que la simplifiquemos o sin que perdamos su complejidad y su contexto de ocurrencia. Las técnicas del análisis de datos cualitativos se emplean en investigaciones conducidas por sociólogos, historiadores, antropólogos, críticos literarios, investigadores de la salud, analistas de mercado, abogados, etc. Por otra parte, las técnicas para la obtención de datos cualitativos son, por ejemplo, las entrevistas abiertas, el análisis documental, los grupos de discusión, las historias de vida, etc. De esta forma, la base inicial para los métodos cualitativos podrá estar constituida por transcripciones de entrevistas o de documentos, por correspondencia comercial, por notas de campo, por documentos literarios o históricos, entre otros.

Como ejemplo de las potencialidades de este tipo de técnicas de análisis, veamos cuáles son las utilidades que podemos desarrollar con el paquete informático NUD·IST (véase el Cuadro 7), uno de los programas más completos disponibles en el mercado: recolectar, comparar y organizar documentos; codificarlos y explorarlos con la finalidad de poner de relieve patrones estructurados; buscar secuencias textuales relevantes (palabras, frases, etc.) en el texto; explorar casos individuales; grabar ideas y teorías emergentes; relacionar teorías y

datos, etc.

La sociolingüística se presenta como una disciplina que alberga un objeto de estudio extremadamente amplio, circunstancia derivada de la diversidad de enfoques teóricos y de metodologías de investigación igualmente diversas que intervienen en su configuración. Si intentamos poner de manifiesto las características más relevantes de su contenido específico, la sociolingüística podría ser definida como todo hecho lingüístico observado en su contexto social, y más concretamente, el estudio de la lengua analizada en su contexto social, con el objetivo de explicar la variabilidad lingüística y la interrelación entre los factores sociales y los fenómenos lingüísticos (Fasold 1984, 1990).

Teniendo en cuenta el panorama actual de la investigación sociolingüística, y aún a riesgo de reducir en exceso las orientaciones de estudio que se desarrollan en la actualidad, tres son las líneas de investigación que predominan. Por un lado, la sociolingüística variacionista, iniciada por William Labov y centrada en el análisis de la variación y del cambio lingüístico (Labov 1972, 1994; Sankoff 1978; Silva-Corvalán 1988). En segundo lugar, la sociología del lenguaje, que suscita un amplio espectro de temas de investigación – desde el análisis del bilingüismo, pasando por el estudio de los procesos de mantenimiento, sustitución y extinción lingüística, y llegando hasta la planificación lingüística, etc.–, todos ellos con una alta autonomía disciplinar (Fishman 1972; Dorian 1989; Romaine 1989; Seliger y Vago 1991). En tercer lugar, la sociolingüística interaccional, vinculada metodológica y conceptualmente con la etnografía del habla, la etnometodología y la interacción simbólica, ha influido poderosamente en el desarrollo de centros de interés más generales como la conversación o el análisis del discurso, y más concretos como, por ejemplo, la alternancia lingüística (Gumperz y Hymes 1972; Hymes 1974; Bauman y Scherzer 1974; Gumperz 1982).

Con respecto a los aspectos metodológicos de la investigación sociolingüística (Moreno 1990; Silva-Corvalán, 1988; Milroy 1987), la diversidad teórica, la variedad de temas de investigación y de realidades a investigar, y la amplitud de fuentes en las que surtirse ha suscitado la presencia de materiales sociolingüísticos dispares para el investigador. Este hecho puede suponer, cuando menos, un problema metodológico que es preciso resolver desde un principio, teniendo en cuenta que la metodología de investigación debe poner en relación tres aspectos: a) la orientación teórica; b) el objeto de estudio; y c) la finalidad de la investigación. Todo ello va a condicionar poderosamente el tipo de material sociolingüístico que recojamos y que tengamos que analizar. Queda claro que no es igual obtener datos sobre variación lingüística que sobre actos de habla; ni es semejante analizar las consecuencias lingüísticas del contacto que estudiar los resultados del aprendizaje de una segunda lengua. Cada tipo de estudio, definido por unos objetivos, un objeto de estudio y una finalidad de investigación, genera un determinado tipo de material sociolingüístico y, casi implícitamente, genera una determinada estrategia de investigación.

Pero para complicar si cabe un poco más el panorama metodológico, en la investigación sociolingüística, al igual que en muchas disciplinas que se preocupan de investigar las

realidades sociales, se actualiza la distinción entre metodología cualitativa y cuantitativa¹. La metodología de investigación cuantitativa se basa en la obtención de datos reducibles a valores numéricos, mientras que la investigación cualitativa nos conduce a la obtención de datos no cuantificables, es decir, datos que no pueden ser sometidos a un tratamiento numérico y que, en cambio, poseen un valor propiamente contextual. Decidirnos por una u otra metodología dependerá, según los casos, bien de la elección del investigador – condicionado por unas determinadas expectativas teóricas–, bien del objeto de estudio analizado. En otras palabras: hay áreas de investigación y temas específicos que requieren una orientación cuantitativa y otros que necesitan una perspectiva cualitativa.

Cualquiera que sea el caso, no estamos ante una elección sin consecuencias, aún más si consideramos la implementación informática del material resultante de la investigación. Como veremos posteriormente, disponemos hoy en día de recursos informáticos para el análisis de cualquier tipo de material sociolingüístico. Por lo tanto, la decisión de emplear una u otra metodología repercutirá sobre el soporte informático más útil y conveniente para la etapa de análisis y tratamiento de los datos.

3.- HERRAMIENTAS INFORMÁTICAS EN LA SOCIOLINGÜÍSTICA CUALITATIVA

Antes de comenzar la descripción de las tareas que son susceptibles de tratamiento informático, debemos aclarar que los datos sociolingüísticos cualitativos proceden de diseños de investigación de carácter experimental (Seliger y Shohamy 1989; Gilbert 1993; Lorenzo 1994: 103-6). En ese marco de referencia, tales datos sociolingüísticos cualitativos constituirán el objeto de estudio a partir del cual procederemos a confirmar o a refutar las hipótesis de trabajo previamente definidas. Es por ello que los datos cualitativos tienen el mismo status experimental que los datos de tipo cuantitativo.

En primer lugar, los datos para el análisis sociolingüístico cualitativo pueden proceder de una de las cuatro fuentes siguientes: encuestas y entrevistas abiertas, grupos de discusión, notas de campo, o observaciones personales. Tal como comentamos en el apartado anterior, los datos cualitativos, a diferencia de los cuantitativos, carecen de un valor definido, teniendo en cuenta que no son reducibles a valores puramente numéricos. En ese sentido, los datos cualitativos representan categorías que adquieren su valor según la posición contextual que ocupan.

En segundo lugar, en consonancia con esas características propias de los datos cualitativos, las necesidades informáticas para el tratamiento del material cualitativo se sitúan en varias tareas de análisis: almacenamiento y ordenación de los datos, edición del material recogido, búsqueda y recuperación de datos en su contexto de ocurrencia, y, finalmente, siempre que sea posible, la generación y el contraste de hipótesis a partir de aquellos datos.

Con el fin de sistematizar el contenido de las actividades de análisis que se pueden

¹ Como indica Moreno (1990), es preciso evitar la confusión entre la cualificación de los datos obtenidos –que consiste en la interpretación de datos cuantitativos– y el análisis cualitativo –proceso de investigación con resultados no cuantificables–.

realizar, describiremos las que constituyen cuatro grandes tipos de tareas en el tratamiento informático del material cualitativo, tal como exponemos en el cuadro siguiente.

-
- Organización
 - Descripción y etiquetado
 - Interpretación
 - Presentación
-

Cuadro 1. Tareas realizadas electrónicamente

En principio, son etapas sucesivas, de manera que cada una debe preceder a la siguiente. En cualquier caso, es posible que, en función del objeto de estudio y del objetivo de investigación, debamos contentarnos con emplear los recursos informáticos propios de las dos primeras tareas.

3.1.- Tareas de organización

Este tipo de actividades están centradas en la transferencia del material textual a un formato organizado y comprensible para los programas informáticos que se vayan a emplear. Y, además, será preciso abordar la codificación del material textual mediante la inserción de palabras clave en aquellos segmentos de texto susceptibles de posteriores búsquedas y recuperaciones.

No es necesario insistir en la importancia de esta etapa para el éxito de las tareas siguientes: cualquier fragmento deficientemente transcrito o pobremente codificado tendrá un valor muy limitado.

3.2.- Tareas de descripción y etiquetado

Las tareas que se contienen en esta etapa serán, en primer lugar, las correspondientes a la búsqueda y recuperación de material: la localización de segmentos relevantes y su preparación para ser analizados. En segundo lugar, será preciso vincular los datos mediante la conexión lógica entre segmentos de textos que mantengan relaciones relevantes, formando categorías, conglomerados (*clusters*) o redes de información. En tercer lugar, el análisis del contenido será otra tarea propia de esta etapa de investigación: contar frecuencias o secuencias de texto, localizaciones de palabras y frases. Algunos ejemplos de este tipo de actividades serían las siguientes: la localización de frases y de palabras relevantes en el texto; la creación de listas de palabras y el conteo de ocurrencias; o la generación de índices y la creación de concordancias contextuales. Finalmente, puede ser muy interesante mostrar los datos, situando aquellos datos seleccionados en un formato organizado de matrices, de redes, etc.

3.3.- Tareas de interpretación de los datos

Una tercera fase será aquella encargada de permitir la falsación de algunas hipótesis de investigación. Así, mediante la extracción y unión de segmentos relacionados temática o formalmente, o mediante la realización de aseveraciones sobre la estructura de los textos y sobre la relación entre distintos conceptos, podremos estar en disposición de intentar dar un paso más y adentrarnos en el contraste y comprobación de hipótesis sobre los datos cualitativos.

Más concretamente, en la interpretación de los datos encontramos dos grandes tipos de actividades: por un lado, el desarrollo de explicaciones sistemáticas y coherentes sobre los datos, mediante la extracción y unión de segmentos relacionados temática o formalmente, las aseveraciones sobre la estructura de los textos, y la relación entre conceptos; por otro lado, la comprobación –bien para confirmar bien para refutar– de las hipótesis de trabajo.

3.4.- Tareas de presentación de los datos y de los resultados

El objetivo de esta última etapa se centra en la representación gráfica de todo tipo de información recolectada en las etapas examinadas anteriormente. Por otra parte, dado que la descripción de la investigación y de sus resultados debe ser preparada para ser difundida entre los especialistas, será preciso contar con herramientas para la preparación de informes finales y de documentos de trabajo.

4.- PROGRAMAS PARA EL TRATAMIENTO DEL MATERIAL CUALITATIVO

Una vez que hemos examinado las tareas que se deben abordar con el material sociolingüístico cualitativo, es el momento de presentar y de comentar los diferentes programas informáticos que, desde principios de los años ochenta, están en disposición de permitirnos la realización de las actividades consideradas en el apartado anterior. La cantidad de programas que permiten el tratamiento informático del material cualitativo aumenta día a día, a medida que se desarrollan nuevas utilidades de análisis y nuevos campos de aplicación². Hay que señalar que no conocemos ningún programa cualitativo que haya sido implementado explícitamente para el análisis de material sociolingüístico cualitativo. Este hecho no supone una carencia importante por cuanto el funcionamiento de los programas diseñados inicialmente para historiadores, antropólogos, economistas o pedagogos pueden ser perfectamente aprovechados por el sociolingüista cualitativo. En cualquier caso, esa realidad

² Para una perspectiva global, véase Fielding y Lee (1993); Richards y Richards (1994); Richards (1997). Una de las descripciones más actualizada de los programas cualitativos la podemos encontrar en Weitzman y Miles (1995). Para consultar un resumen de los programas existentes, véase Miles y Weitzman (1994). Por otra parte, el proyecto CAQDAS (*Computer Assisted Qualitative Data Analysis Systems*), desarrollado en Gran Bretaña, organiza seminarios, coordina un grupo de discusión en Internet (*qual-software*) y dispone de una página web (<http://www.soc.surrey.ac.uk/caqdas/>) sobre todo tipo de programas para el análisis de material cualitativo.

sí refleja en alguna medida la poca preocupación mostrada en la simplificación de algunas tareas de investigación y en la puesta a punto de mecanismos de análisis complejos.

En el Cuadro 2 presentamos los seis tipos de programas de los que hoy en días podemos disponer para el tratamiento del material cualitativo. En primer lugar, distinguimos entre los programas genéricos y los programas diseñados específicamente para el tratamiento de material cualitativo. Con respecto a los primeros, tienen como característica principal que son programas no pensados específicamente para analizar datos cualitativos. Distinguimos tres tipos de programas genéricos: los procesadores de texto, los programas recuperadores de material textual, y los organizadores de texto.

I. Programas genéricos:

1.1 Procesadores de texto.

1.2 Recuperadores de texto.

1.3 Organizadores de texto (*textbase managers*).

II. Programas diseñados para el análisis de datos cualitativos:

2.1 Programas de codificación y recuperación.

2.2 Programas para la generación / comprobación de hipótesis.

2.3 Programas para la generación de redes conceptuales.

Cuadro 2. Tipos de programas para el tratamiento de datos cualitativos

Con respecto a los *procesadores de texto*, la amplitud de recursos que ofrecen las versiones actuales de los principales procesadores de texto existentes en el mercado (Microsoft Word, WordPro, Corel WordPerfect, etc.) permite que haya un conjunto de actividades de organización y de manejo de datos cualitativos que se pueden realizar con este tipo de soportes, actividades como las siguientes: la producción y revisión de los textos; la realización de búsquedas con diferentes niveles de complejidad; la realización de conexiones hipertextuales, o la generación de macros para automatizar rutinas de búsquedas y recuperaciones.

El segundo tipo de programas genéricos son los que hemos denominado como *recuperadores de texto*. Se trata de programas de búsqueda textual muy sofisticados, especializados en la localización de todo tipo de secuencias de caracteres, de combinaciones de caracteres, y todo ello realizado en uno o en varios documentos a la vez. En algunos programas de recuperación textual también es posible realizar determinados tipos de análisis de contenido: contar apariciones de elementos en el texto, mostrar palabras en contexto, crear listas de palabras y concordancias, etc.

El tercer tipo de programas genéricos son los *organizadores de texto* (denominados originalmente en inglés *textbase managers*). Aunque se asemejan a los recuperadores de texto en las actividades que realizan y en las utilidades principales que albergan, estos programas presentan una organización del material cualitativo en formato de base de datos. Este hecho permite que sean muy adecuados para profundizar en la organización, ordenación y formalización de distintos subconjuntos de textos.

Pasemos ahora a analizar el segundo gran tipo de programas cualitativos, aquéllos que hemos definido como programas diseñados específicamente para el análisis de datos cualitativos³. Estamos ante un conjunto de programas que han sido preparados específicamente para las necesidades y los intereses del análisis cualitativo, además de ser implementados por investigadores procedentes de diferentes áreas concretas de conocimiento. Además, a diferencia de los programas anteriormente mencionados, suele tratarse de programas nacidos de prototipos experimentales no comerciales, algunos de los cuales mantienen aún ese status mientras que otros se han convertido en programas plenamente comerciales distribuidos por empresas importantes.

Hemos distinguido tres tipos de programas en función del alcance de los recursos que contienen: en primer lugar, programas de codificación y recuperación; en segundo lugar, programas para la generación y comprobación de hipótesis; y en tercer lugar, programas para la generación de redes conceptuales.

Con respecto a los programas de *codificación y recuperación*, su contenido consiste en realizar, electrónica y ordenadamente, las labores de organizar los datos (ordenar, cortar y pegar). Por otro lado, entre las tareas que realizan encontramos las siguientes: dividir los textos en segmentos independientes, adjuntarle a cada segmento información de codificación, cortar, ordenar, y reorganizar todos los segmentos según los intereses y las necesidades del investigador, etc.

En relación con los programas para la *generación y comprobación de hipótesis*, además de incorporar las prestaciones que constatamos en el párrafo anterior para los programas de codificación y recuperación, permiten analizar las conexiones entre códigos basada en un sistema de reglas y en la lógica formal, con las siguientes finalidades: para el establecimiento de clasificaciones de generalización alta; para formular aseveraciones; para probar el grado de aplicación de las aseveraciones en los datos disponibles.

Finalmente, con respecto a los programas para la *generación de redes conceptuales*, estamos ante las utilidades más complejas implementadas hasta el momento. A partir de los datos, de su codificación y de las relaciones entre conceptos que el investigador establece, estos programas permiten formular agrupamientos entre segmentos de texto en virtud de afinidades semánticas. Para representar tales relaciones, las variables se muestran como nodos vinculados a otros nodos mediante líneas o flechas que representan relaciones específicas. El objetivo es construir y probar una hipótesis sobre los datos mediante redes (gráficas) de ideas.

Tras establecer una tipología de programas cualitativos, vamos a mostrar a continuación un análisis comparativo entre algunos de los programas más destacados en cada categoría⁴. Los rasgos que comparamos están relacionados con las utilidades y el potencial de

³ En muchos casos es posible acceder a una versión gratuita de demostración de un determinado programa cualitativo, versión que contendrá limitaciones en el tiempo de uso o restricciones en su contenido. Para encontrar información sobre las versiones de demostración accesibles en Internet, pueden consultarse las páginas siguientes: <http://www.soc.surrey.ac.uk/caqdas/packages.htm>; <ftp://ifsws.soziologie.uni-jena.de/pub/text-analysis>.

⁴ La mayor parte de la información mostrada en los cuadros siguientes está contenida en Miles y Weitzman (1994). Véase también Weitzman y Miles (1995: 310-339). Para obtener información específica de cada

cada programa. En primer lugar, informamos sobre la versión del programa que evaluamos y el sistema operativo en el que se puede implementar. A continuación, señalamos si dispone de una utilidad de codificación incorporada, si permite realizar las operaciones de buscar y recuperar material textual, si dispone de una opción que integre una base de datos, y si el programa cuenta con una herramienta interna de edición de notas, memos, etc. Además, en algunos programas, como hemos visto anteriormente, existe la posibilidad de establecer conexiones entre las diferentes partes del texto. Por otra parte, un objetivo que persiguen algunos programas consiste en ayudar a aclarar las ideas mediante mecanismos lógicos o basados en reglas, o mediante el establecimiento de redes semánticas. Para dar cuenta de ello, incluiremos en los cuadros como características de evaluación los rasgos de vinculación de datos, de generación de matrices, de mostrar datos en forma de red, y la posibilidad de generar hipótesis. Finalmente, incluimos una casilla en la cual informamos sobre la facilidad de empleo del programa en general por parte de un usuario medio⁵.

Para realizar tal comparación emplearemos una doble notación de símbolos: por un lado, emplearemos una notación de carácter informativo; por otro, usaremos una notación para calificar la potencia de la característica en cuestión.

-
- Notación *informativa*
 - Diseñado para esa finalidad.
 - No diseñado para esa finalidad, pero puede desarrollarla limitadamente.
 - Característica ausente.
 - Notación *calificativa*
 - Característica Fuerte.
 - Normal.
 - Débil.
 - Característica ausente.
-

Cuadro 3. Contenido de las notaciones empleadas

En cuanto a los programas de recuperación textual, observamos en el cuadro que la mayoría se centran en el desarrollo de utilidades de búsqueda y de recuperación de datos, al tiempo que pueden ser empleados para establecer distintas estructuras de bases de datos.

producto informático, pueden consultarse las siguientes páginas web, en las cuales encontramos información conjunta de diferentes programas cualitativos con sus correspondientes enlaces a las páginas individuales: <http://www.soziologie.uni-jena.de/home/klein/textnae.htm>; <http://www.unige.ch/ses/sococ/qual/qual.htm>; <http://www.ualberta.ca/~jrmorris/qda.html>.

⁵ Antes de exponer los resultados de la muestra comparativa, queremos indicar que la lista de programas que empleamos en cada categoría no es todo lo exhaustiva que pudiese haber sido teniendo en cuenta la cantidad de programas y de aplicaciones informáticas que se pueden encontrar en el mercado. En cualquier caso, hemos contado con una muestra que consideramos representativa de los productos informáticos más interesantes de cada tipo de programa cualitativo.

PROGRAMA	Versión	Sistema Operativo	Codificación	Buscary Recup.	Base de Datos	Notas	Vinculación Datos	Generar Matrices	Mostrar Red	Generar Hipótesis	Facilidad Usuario
Metamorph	4.0	Mc D U	—	●	□	■	■	—	—	—	●
Orbis	1.0	D W	□	●	■	□	—	●	—	—	●
Sonar Professional	8.5	Mc W	—	●	□	■	■	—	—	□	●
The Text Collector	1.7	D	—	●	□	—	—	—	—	—	●
WordCruncher	4.5	D	—	●	□	—	■	—	—	—	○

Cuadro 4. Rasgos de los programas de recuperación de textos

Los programas de organización de textos, tal como comprobamos en el Cuadro 5, presentan una mayor amplitud de funciones en comparación con los programas anteriores. Así, no sólo las tareas de codificación y de búsqueda y recuperación constituyen los puntos fuertes de estos programas, sino que se generaliza la utilidad de base de datos y se van ampliando las posibilidades para la vinculación entre los datos, la edición de notas y la generación de matrices.

PROGRAMA	Versión	Sistema Operativo	Codificación	Buscary Recup.	Base de Datos	Notas	Vinculación Datos	Generar Matrices	Mostrar Red	Generar Hipótesis	Facilidad Usuario
askSam	5.1	D W	—	●	■	□	■	—	—	—	●
Folio VIEWS	3.1	D W	■	●	■	■	■	○	—	—	●
Tabletop	1.x	Mc	—	●	■	■	—	●	—	■	●
MAX	3.5	D	■	○	■	—	■	—	—	—	●

Cuadro 5. Rasgos de los programas de organización de textos

Estos son los programas diseñados especialmente para rotular y etiquetar, y para llevar a cabo todo tipo de búsquedas. A pesar de estar diseñados con esos objetivos, no todos los programas parecen responder satisfactoriamente a las características propuestas, tal como podemos contemplar en el cuadro siguiente.

PROGRAMA	Versión	Sistema Operativo	Codificación	Buscary Recup.	Base de Datos	Notas	Vinculación Datos	Generar Matrices	Mostrar Red	Generar Hipótesis	Facilidad Usuario
HyperQual2	1.0	Mc	■	○	■	■	■	○	—	□	●
Kwalitan	4.0	D	■	●	□	■	■	—	—	□	●
Martin	2.0	W	■	○	□	■	□	—	—	—	●
QUALPRO	4.0	D	■	○	□	□	—	—	—	—	●
The Ethnograph	4.0	D	■	●	□	■	—	—	—	—	●

Cuadro 6. Rasgos de los programas de codificación y de recuperación textual

Los programas para la generación de hipótesis son los que presentan una mayor amplitud de tareas implementadas, por cuanto la mayoría cubren el conjunto de características desde la codificación hasta la generación de hipótesis. Quizás por ello son hoy en día los programas más difundidos y con más aceptación general entre los investigadores cualitativos..

PROGRAMA	Versión	Sistema Operativo	Codificación	Buscary Recup.	Base de Datos	Notas	Vinculación Datos	Generar Matrices	Mostrar Red	Generar Hipótesis	Facilidad Usuario
AQUAD	4.0	D	■	○	□	■	□	●	—	■	●
ATLAS/ti	1.1	D	■	●	□	■	■	●	●	■	●
HyperRESEARCH	1.55	McW	■	●	□	—	—	○	—	■	●
NUD-IST	3.0	McW	■	●	■	■	□	●	●	■	●
QCA	3.0	D	—	○	—	□	—	○	—	■	○

Cuadro 7. Rasgos de los programas para la generación de hipótesis

Finalmente, los programas para la generación de redes conceptuales centran su actividad en mostrar la relación entre conceptos, tarea para la cual se generalizan las utilidades de vinculación de datos y de mostrar la red de relaciones, en tanto que quedan en un segundo plano otro tipo de utilidades relacionadas con la búsqueda, recuperación y codificación del material cualitativo.

PROGRAMA	Versión	Sistema Operativo	Codificación	Buscar y Recup.	Base de Datos	Notas	Vinculación Datos	Generar Matrices	Mostrar Red	Generar Hipótesis	Facilidad Usuario
Inspiration	4.0	McW	—	●	—	■	■	—	●	—	●
MECA	1.0	McDU	—	—	—	—	■	—	●	■	○
MetaDesign	4.0	McW	—	○	—	□	■	—	●	—	●
SemNet	1.1beta	Mc	—	●	—	■	■	—	●	■	●

Cuadro 8. Rasgos de los programas para la generación de redes conceptuales

La selección de un tipo determinado de programa dependerá de las necesidades de investigación y dependerá de lo lejos que queramos llegar en la explotación de las posibilidades de computación con los programas cualitativos. En muchos casos tan sólo precisaremos un programa que nos permita codificar nuestros textos y buscar y recuperar ciertos fragmentos en sus contextos de ocurrencia. En otros casos, y a medida que aumentamos el horizonte de nuestras expectativas, podemos acercarnos a los programas que nos permiten realizar generalizaciones teóricas a partir de nuestros datos y contrastar ciertas hipótesis de trabajo.

En lo relativo a la elección de un programa concreto dentro de cada categoría, la elección de uno concreto puede verse facilitada por las características expuestas en los cuadros precedentes. La posibilidad, en bastantes casos, de disponer de una versión de demostración puede ayudar a decantarse por aquellos que sean más adecuados a nuestros objetivos particulares de investigación.

5.- FINAL

El análisis de datos cualitativos ha experimentado un salto considerable en los últimos años, gracias en parte al creciente desarrollo de herramientas informáticas que posibilitan un análisis variado y complejo del material textual recogido. El progresivo refinamiento de las utilidades contenidas en los paquetes informáticos está ya permitiendo el desarrollo de herramientas de análisis que hasta ahora parecían formar parte exclusiva de los programas destinados al análisis de datos cuantitativos, como son el contraste y la comprobación de hipótesis.

Hemos visto en las páginas precedentes que el material cualitativo generado por una investigación sociolingüística puede ser analizado mediante recursos lógicos informáticos. A pesar de no disponer de herramientas pensadas específicamente para las labores sociolingüísticas, el amplio abanico de programas diseñados inicialmente para historiadores, pedagogos, antropólogos o economistas encaja con las necesidades y con el material recogido por los sociolingüistas.

BIBLIOGRAFÍA

- Ammon, U. et al. eds. 1987. *Sociolinguistics: an international handbook of the science of language and society*. Berlin: Mouton.
- Appel, R. y P. Muysken 1987. *Language contact and bilingualism*. Londres: Edward Arnold. [Traducción al castellano en Ariel, Barcelona, 1996]
- Bauman, R. y J. Scherzer eds. 1974. *Explorations in the ethnography of speaking*. New York: Academic Press.
- Bryman, A. y R. Burgess eds. 1994. *Analyzing qualitative data*. Londres: Routledge.
- Burgess, R. 1984. *In the field. An introduction to field research*. Londres: George Allen & Unwin.
- Denzin, N. K. y Y. S. Lincoln eds. 1994. *Handbook of qualitative research*. Londres: Sage.
- Fasold, R. 1984. *The sociolinguistic of society*. Oxford: Basil Blackwell.
- Fasold, R. 1990. *The sociolinguistics of language*. Oxford: Basil Blackwell.
- Fielding, J. 1993. "Coding and managing data". Ed. Gilbert, N. 218-238.
- Fielding, N. y R. Lee 1991. *Using computers in qualitative research*. Londres: Sage.
- Fishman, J. 1972. *The sociology of language: An interdisciplinary approach to language in society*. Rowley: Newbury House.
- Gilbert, N. ed. 1993. *Researching social life*. Londres: Sage.
- Gumperz, J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.
- Gumperz, J. y D. Hymes eds. 1972. *Directions in sociolinguistics*. New York: Holt.
- Hymes, D. 1974. *Foundations in sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press.
- Kelle, U. ed. 1995. *Computer-aided qualitative data analysis: theory, methods and practice*. Londres: Sage.
- Labov, W. 1972a. *Sociolinguistics patterns*. Philadelphia: University of Pennsylvania Press. [Versión en castellano: *Modelos sociolingüísticos*, Madrid, Editorial Cátedra, 1983].
- Labov, W. 1972b. *Language in the inner city: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
- Labov, W. 1994. *Principles of linguistic change. Internal factors*. Oxford: Basil Blackwell. [Versión española de Pedro Martín Butrageño, Madrid, Gredos, 2 vol., 1996].
- Lorenzo, A. M. 1994. "Tratamiento informático do material sociolingüístico: algunhas suxestións a partir dos protocolos de investigación". *Aplicaciones lingüísticas de la informática*. Ed. Javier Gómez Guinovart. Santiago de Compostela: Edicións Tórculo. 97-112
- Miles M. B. y E. A. Weitzman 1994. "Choosing computer programs for qualitative data analysis". Eds. Miles, M. B. y A. M. Huberman. 311-317.

- Miles, M. B. y A. M. Huberman 1994. *Qualitative data analysis. An expanded sourcebook*. Thousand Oaks: Sage.
- Milroy, L. 1987. *Observing and analysing natural language*. Oxford: Basil Blackwell.
- Moreno, F. 1990. *Metodología sociolingüística*. Madrid: Gredos.
- Procter, M. 1993. "Analysing survey data". Ed. N. Gilbert. 239-254.
- Richards, T. y L. Richards 1994. "Using Computers in qualitative analysis". *Handbook of Qualitative Research*. Eds. N. K. Denzin y Y. S. Lincoln. Thousand Oaks: Sage. 445-462.
- Richards, L. 1997. "Computers and Qualitative Analysis". *The International Encyclopedia of Education*. Oxford: Elsevier Science.
- Romaine, S. 1989. *Bilingualism*. Oxford: Basil Blackwell.
- Sankoff, D. ed. 1978. *Linguistic variation, models and methods*. New York: Academic Press.
- Sarabia, B. y J. Zarco 1997. *Metodología cualitativa en España*. Madrid: Centro de Investigaciones Sociológicas.
- Silva-Corvalán, C. 1988. *Sociolingüística: teoría y análisis*. Madrid. Alhambra.
- Seliger, H. y E. Shohamy 1989. *Second language research methods*. Oxford: Oxford University Press.
- Weitzman, E. A. y M. B. Miles 1995. *Computers programs for qualitative data analysis. A software sourcebook*. Thousand Oaks: Sage.