

Métodos de análisis de las verbalizaciones: una contribución del análisis textual y análisis conceptual mediante el uso de software

Eduardo Escalante Gómez

Director del Instituto de Investigaciones, Facultad de Psicología, Universidad del Aconcagua, Mendoza, Argentina. Correo electrónico: ed.escalante@gmail.com

32

Recibido: mayo de 2009 / Aceptado: junio de 2009

PARA LOS INVESTIGADORES EN CIENCIAS HUMANAS Y CIENCIAS SOCIALES, el análisis e interpretación conceptual de los textos constituye una etapa importante de sus indagaciones científicas. El investigador en algún momento de la investigación se enfrenta a la interpretación conceptual de los textos.

El análisis de textos asistido por computadoras es un método que ayuda al análisis interpretativo de un concepto en un corpus textual, producido por un individuo o un grupo. La finalidad de este tipo de análisis es descubrir el repertorio semántico y categorial asociado con un determinado concepto, con segmentos de texto y con el corpus total.

Este tipo de análisis se funda en diferentes fuentes disciplinarias: lingüística, teorías de la enunciación y del discurso, análisis cualitativo y análisis estadístico. En el contexto de este escrito, se define el análisis conceptual como un método interpretativo para la exploración sistemática de las propiedades semánticas y de los elementos indiferenciados ensamblados en un texto o un discurso (Meunier, 1998; Bandom, 1994).

Palabras clave: categorización / co-ocurrencias / escalamiento multidimensional / análisis semántico / programas computacionales

1. Introducción¹

Este artículo presenta a los académicos en general, y a los investigadores en particular, una mirada sobre las discusiones y búsquedas que se realizan con instrumentos para procesar datos cualitativos en toda su complejidad. Muestra que el análisis textual, el análisis

lingüístico, y el análisis temático son procesos necesarios para elaborar interpretaciones teóricas de los datos. Son muchas las cuestiones que emergen en la aplicación de la metodología cualitativa. Los estudios cualitativos enfrentan al investigador a las grandes discusiones de las ciencias del hombre.

Hay cuestiones de corte epistemológico asociadas a la construcción del conocimiento científico: cuestiones metodológicas relativas a la validez de tratamiento estadístico cuando se trata de categorías y enfoques cualitativos, y cuestiones relativas al lenguaje y la validez de la aplicación del análisis matemático para estudiarlo². Inducción, deducción, abducción, inferencia, juicio, son algunos de los términos que requieren un acertado análisis.

A modo de precisión, es importante tener presente que lo que se expone en este artículo no apunta a una matematización de las palabras³, sino a la incesante búsqueda del sentido. No se trata de medir por medir, pues lo que se pretende es el entendimiento de la conciencia que las produjo y aclarar si pudieren ser también el resultado de algo meta-subjetivo o más allá del sujeto (ver Rastier, 1989; Reinert, 1993; y Tournier, 1980). Es decir, el alcance de este artículo es introducir al lector a una realidad de exploración lógico-matemática mediante herramientas computacionales, sin por ello postular una determinada manera de construcción científica, pero sí argumentando que el trabajo de investigador cualitativo requiere de un riguroso análisis epistemológico que clarifique por un lado la diferencia entre explicar, comprender e interpretar, y por el otro la diferencia entre comprender y querer⁴.

1.1. El escenario metodológico: una breve síntesis

La construcción científica procede por desplazamientos y reducción del conjunto de datos recopilados. El investigador cualitativo procede a discretizar⁵ un “corpus”, que significa clasificar los textos⁶ (automática o no automáticamente); identificar los temas (tópicos) o argumentos que se entretujan en los textos; usar diferentes estrategias donde el algoritmo significa: a) etiquetar manualmente los segmentos de textos (combinando un tratamiento cuantitativo, por ejemplo análisis de correspondencias, con una exploración cualitativa), y b) creación de repertorios de indización (que también combinan tratamientos cuantitativos con procedimientos cualitativos).

Las diferentes estrategias de reducción pueden significar desplazamientos entre la automatización (por ejemplo, se puede estar usando MaxQda) y procedimientos manuales o artesanales.

Las herramientas que el investigador puede utilizar son para: 1) realizar análisis multidimensionales (Hamlet II, Alceste); 2) realizar etiquetados manuales de segmentos de texto (MaxQda, Nvivo, Atlas.ti); y 3) generar o aplicar repertorios de categorías temáticas (TROPES). La intencionalidad es estudiar proximidades, concordancias, similitudes, distancias, equivalencias, sinonimia. Y el resultado es producir gráficos (no dirigidos) de clusters o conglomerados que sinteticen las categorías y permitan la interpretación del corpus que se esté analizando (especialidad del investigador). El investigador busca la “objetivación”. Más que con objetividad, se trata de poner, en este caso, los textos en perspectiva⁷.

La interpretación se puede sustentar, por ejemplo, en: 1) frecuencias de palabras; 2) co-ocurrencia, por ejemplo la investigación de motivos (expresión, argumentación); 3) asociación de categorías (hipotética o no); 4) forma del discurso: polo narrativo (descriptivo) o polo argumentativo (explicación). Este proceso se puede realizar a nivel local (texto) o a nivel global (corpus), o a nivel de la evolución (diacronía).

Por ejemplo, los métodos derivados del estudio estadístico de las frecuencias significan una etapa de clasificación de los documentos para la generación de tablas bidimensional o multidimensionales a las que se puede aplicar posteriormente análisis de correspondencias, análisis de componentes principales, análisis factorial, análisis de escalamiento multidimensional (en adelante EMD)⁸, entre otros.

1.2. Las herramientas

34

En la actualidad existe una serie de instrumentos informáticos que asisten al análisis interpretativo experto. Me refiero a herramientas tales como Atlas.ti (Barry, 1998), las herramientas de análisis de textos de inspiración estadística (Lebart & Salem, 1994) tales como SPAD, Alceste (Reinert, 1986), HAMLET II (Brier & Hopp, 2006), LIMC2007 (Pennebaker; Booth & Francis, 2007). Estos últimos programas emplean estrategias matemáticas ligeramente diferentes a las que usan programa como Atlas.ti, MaxQda, Nudist. Más abajo ilustraré sobre algunos de estos enfoques.

Todas estas herramientas asisten a la identificación de relaciones, temáticas, semánticas, de ensamblaje de palabras (simples o compuestas), de sinonimias de un determinado corpus. Este proceso permite que el analista experto dirija la atención sobre un concepto focalizado y pueda explorar su funcionamiento. El proceso se ve facilitado si se hace una exploración a partir de un *diccionario* dado, una clasificación dada o simplemente se construye dicho diccionario -asumiendo una postura más desde la teoría fundamentada.

1.3. La hipótesis

Este tipo de herramientas computacionales asisten el análisis de los textos y su análisis conceptual, respetando las exigencias del trabajo interpretativo del experto (por ejemplo procesos de validación). La expresión de un concepto presente en un texto se puede indagar a través de sus regularidades lingüísticas y es posible identificar de manera algorítmica relaciones entre ellas. El investigador puede generar los contextos de un concepto, también puede proceder a clasificar dichos contextos y además proceder a anotar las categorías de las clases.

El análisis conceptual se realiza mediante la exploración e interpretación de los contextos lingüísticos y semánticos de una expresión susceptible de expresar un concepto. Una de las técnicas clásicas que se usan es la *concordancia* y sus variantes. Su función es producir todos los contextos lingüísticos y semánticos –frase(s) o enunciado(s)– de un término. Este proceso de clasificación se puede realizar de manera automática.

El procedimiento metodológico puede significar varias etapas relativamente secuenciales: 1) selección de un corpus de textos y los conceptos que se quiere estudiar -la preparación del corpus; 2) extracción de los contextos lingüísticos y semánticos del o de los conceptos(s) seleccionados; 3) clasificación automática de los contextos; y 4) anotación categorial de estas clases.

2. Algunas herramientas de análisis: TROPES, Hamlet, LIWC2007

A continuación se expondrán tres de las herramientas de gran utilidad al momento de realizar los análisis de texto y el análisis conceptual.

2.1. TROPES

Diferentes programas computacionales pueden extraer palabras clave, segmentos de textos y diferentes elementos verbales. Ejecutan análisis morfo-sintácticos del corpus y construyen redes de frases sustantivas que pueden ser unidades terminológicas. Esto es parte de lo que hace TROPES, pero su especificidad es extraer las semánticas de los textos, esto es, analizar su contenido. El programa clasifica las palabras, detecta los contextos, filtra los tópicos según su relevancia, grafica permitiendo visualizar cada referencia detectada en el contexto discursivo, identifica el escenario semántico para la interpretación.

El trabajo en TROPES se puede dividir en los siguientes pasos. El *primer paso* que debe dar el analista o investigador es realizar una clasificación semántica del texto que se quiere analizar, por ejemplo, un texto de 1,500 palabras. Este proceso se realiza en tres diferentes niveles: *Universo 1*, *Universo 2* y *referencias utilizadas*, organizados en una estructura jerárquica por el programa o por el investigador. El primer nivel, el *Universo 1*, se compone de conceptos muy generales (cuestiones sociales, cuestiones de derecho y justicia, educación y enseñanza, entre otras. El segundo nivel, *Universo 2*, se compone de conceptos especializados del *Universo 1* (por ejemplo, alumnos y profesores, comportamientos positivos, identidad y familia, etc.). Este universo referenciado representa el concepto con 200 a 1000 clases “semánticas equivalentes” posibles. Esta clasificación permite identificar las palabras relevantes. El tercer nivel, *referencias utilizadas*, ejecuta el agrupamiento analógico. Se pueden organizar 10,000 clases equivalentes. Por ejemplo, la palabra “juez” puede ser asociada con la clase “abogado”, incluida en el *Universo 2*, y en el *Universo 1* con “ley” (ver cuadro 1). Es posible modificar o mejorar estas clasificaciones usando un escenario semántico, e incluso crear las propias clasificaciones.

En el *segundo paso*, se constata que entre las 1,500 palabras muchas son irrelevantes. Por lo tanto, se requiere un trabajo previo de selección antes de seguir adelante. Este proceso lo realiza TROPES eliminando las palabras etiquetadas como “conectores”, “modalizaciones” o “preposiciones”. Se generan, por ejemplo 400 nuevas palabras (incluyendo los términos conjugados).

En el *tercer paso*, con este nuevo grupo de palabras se ejecuta manualmente una poda. Al final de esta etapa, para un texto se seleccionan cuatro a cinco conceptos del *Universo 1* (por ejemplo, ley, familia, acuerdo, persona), aproximadamente diez conceptos en el

Cuadro 2. Muestra de indicadores del lenguaje utilizados por **TROPES** (tomada de Ghiglione & Bromberg, 1998)

Indicadores	Categorías	Ejemplos
Verbos	Factivos (verbos de acción) Estativos (verbos de estado) ⁹ Declarativos (sobre un estado) ¹⁰ Actuación	hacer, ensayar, dar ser, tener, existir, parecer decir, pensar, querer ordenar, declarar, querer, prometer
Conectores	Condición Causa Oposición Comparación Tiempo Lugar	si, entonces, en función de parece que, debido al hecho que pero, por el contrario como cuando, en seguida donde
Adjetivos	Objetivos Subjetivos Numéricos	antiguo, nuevo, normal, correcto, real primero, mil
Modalizaciones (adverbios y locuciones adverbiales)	Tiempo Lugar Manera Afirmación Duda Negación	después, actualmente aquí, atrás habitualmente, relativamente, bien, mal evidentemente, realmente puede ser jamás

En **TROPES** se pueden identificar cuatro estilos distintos de discursos: argumentativo, enunciativo, narrativo, y descriptivo (ver cuadro 3). Este cuadro no constituye una lista exhaustiva, sino una tendencia o un referente.

Cuadro 3. Estilos de discursos distinguidos por **TROPES**, indicadores idiomáticos principales de utilización para las distinciones y explicaciones (adaptación de Ghiglione & Bromberg, 1998)

Estilos	Principales indicadores idiomáticos	Explicación
Argumentativo	verbos estativos modalizaciones de negación modalizaciones de intensidad conectores de pero conectores de adición	El entrevistado argumenta, explica...
Enunciativo	verbos declarativos modalizaciones de afirmación modalizaciones de intensidad conectores de pero conectores de adición conectores de causa	El entrevistado revela su punto de vista...
Narrativo	verbos factivos modalización del tiempo modalización del lugar modalización de la manera modalización de la afirmación conectores de adición conectores de disyunción conectores de comparación	El entrevistado expone una sucesión de sucesos que se desarrollan en un momento dado, en un lugar dado.
Descriptivo	verbos factivos modalización del tiempo modalización del lugar conectores de tiempo conectores de lugar adjetivos objetivos	El entrevistado identifica, describe o clasifica las cosas o las personas.

38

El investigador puede utilizar las clasificaciones semántica propuesta por **TROPES** o elaborarlas en él, como Piolat (Piolat & Bannour, mimeo), quien elaboró un lexicón denominado **EMOTAIX-TROPES** sobre la influencia de la ansiedad en la escritura expresiva. Se puede iniciar el proceso investigativo organizando los discursos de las personas entrevistadas respecto de un escenario en común (esto es, un agrupamiento de universos de referencia de los entrevistados que describen una determinada situación).

Con los resultados obtenidos en **TROPES** se puede realizar análisis multivariados en programas como **SPAD** y **SPSS**, por ejemplo, análisis de EMD, análisis de componentes principales, análisis de correspondencia, entre otros. Para cada variable analizada se puede usar una prueba U de Mann-Whitney a efecto de verificar la diferencia del universo de referencia y el indicador idiomático para poder determinar entre los discursos de los entrevistados, por ejemplo, el estilo de discurso utilizado: las modalizaciones de intensidad, los verbos afectivos, estativos y declarativos.

The screenshot displays the TROPES software interface. On the left, a tree view under 'Resultados' shows a hierarchical structure of grammatical categories. The 'Verbos' category is expanded, showing sub-categories like 'Factivo' (63.9%), 'Estativo' (27.8%), 'Declarativo' (8.3%), and 'Performativo' (0.0%). Other categories include 'Conectores', 'Modalizaciones', and 'Adjetivos'. On the right, the 'Secuencias' panel shows a list of text excerpts with various grammatical features highlighted in blue, such as 'me llevo bien en el laburo', 'así. Va bien asícuando era chico', 'y vas saliendo adelante', 'y cambié una banda', 'y como ven que le pongo onda', 'Por ahí si no hay en la pintura', 'repartir comida', 'repartir. Ahora se me ha complicado', 'con la tarjeta de mi jefe me lo saco', 'También estoy armando mi pieza', 'que termine el colegio voy a poder tener un estudio completo', 'hacer una carrera y voy a ganar más en otro trabajo', 'si me pongo las pilas la mayoría de mis amigos están yendo a la facultad', 'Por ahora tengo que terminar en el cens de noche la secundaria y después ir a la facultad', 'En el trabajo podría meter horas extras pero no podría ir a la escuela', 'entonces de a poco lo voy haciendo', 'pero me he puesto las pilas', 'como que el resultado fue mejor que agarrarte a piña', 'joya por ahí se me pasa de todo', 'Si, como que he cambiado en la calle', 'y me dicen vos te pones las pilas', 'yo he visto otros chabones dice', 'yo confío en pocos porque a los nuevos los tenés que contar con los dedos de las manos', 'que confía en mí por ahí hay chabones que se me ponen a hablar', 'y el chabón me cuenta como va', 'Vas creciendo y vas viendo otras cosas', 'que no les contás porque les hablas y lo cuentan', and 'se entera todo el mundo. cuentan todo'.

Ilustración 1. Procesamiento en TROPES

El uso de esta herramienta puede resultar muy interesante para investigar los ensayos escritos por alumnos de cursos de posgrado. Por ejemplo, para identificar sus estilos narrativos (factivo, estativo, declarativo, preformativo) y su relación con las formas argumentales requeridas por el programa de posgrado. Esta es un área de desarrollo incipiente a nivel internacional.

2.2. Hamlet II

El programa HAMLET II permite examinar diferentes opciones de contexto para contar la *co-ocurrencia* de las palabras según el nivel de análisis requerido por el investigador. Después de estandarizar las frecuencias brutas empatadas (joint frequencies) se puede considerar el resultado como una *matriz de similitudes* entre un grupo de palabras usadas y sometidas a un *análisis categorial* para hacer las relaciones más visibles. En esta misma tradición se encuentra el análisis factorial, pero en el análisis textual el modelo de análisis multidimensional no métrico es más apropiado y no requiere los supuestos que se requieren cuando se hacen mediciones con variables continuas. El modelo EMD busca reproducir solamente el orden de las distancias entre puntos, representando las variables en la matriz mediante la reducción de la dimensionalidad del espacio en el cual se representan. Además, se pueden visualizar en no más de tres dimensiones, en cambio el análisis factorial requiere un número mayor. En términos de análisis de datos cualitativos, el enfoque depende de la formulación de una lista de palabras teóricamente relevantes *-diccionario-*, de las cuales se puede definir la co-ocurrencia respecto de una unidad de contexto significativa, en los textos sometidos al análisis. Una vez que la lista de palabras ha sido definida por el investigador, se puede hacer una serie de comparaciones, por ejemplo, entre textos de diferentes entrevistas con preguntas abiertas o en profundidad.

Análisis contextual preliminar¹¹

Es conveniente realizar análisis exploratorios preliminares de análisis contextual usando la rutina “Keyword-in-Context Soutine: KWIC” (palabra clave en contextos). Se pueden emplear diferentes combinaciones de palabras indicando sus sinónimos o palabras equivalentes (opción disponible en el programa). Para ilustrar la potencialidad del programa se ha tomado una entrevista recopilada de una investigación organizacional. El análisis contextual exploratorio permitió seleccionar diez palabras contextualizadas -esto es, referidas a temas de gestión y formación- con sus correspondientes sinónimos. Para tal efecto, además de seleccionar el texto de la entrevista, se creó el archivo con el vocabulario.

Cuadro 4. Frecuencias de las diez palabras

CATEGORY/WORD COUNTS				
VOC.LST.	FREQUENCY	% VOC.LST.	% TEXT	CONTEXT UNITS
conocimientos	9	8.18	0.34	3
creo	19	17.27	0.72	11
Director	6	5.45	0.23	4
experiencia	4	3.64	0.15	3
formación	14	12.73	0.53	6
objetivo	10	9.09	0.38	4
orden	5	4.55	0.19	3
perfil	12	10.91	0.45	4
persona	26	23.64	0.98	11
presión	5	4.55	0.19	1

2640 words were read from the text file.
110 entries were found in the search list, and
22 context-units were counted, of which
18 contained at least one item in the search list.

Aplicando **HAMLET II** con este vocabulario se obtiene la estadística individual y frecuencias empatadas de ocurrencias en las frases. Las frecuencias de palabras individuales se cuentan con las frecuencias empatadas para todo par posible de palabras y se calculan las frecuencias empatadas estandarizadas.

Cuadro 5. Valores estandarizados

STANDARDISED JOINT INDEX VALUES

Jaccard coefficient - ignores joint non-occurrence

	i	1	2	3	4	5	6	7	8
		9							
conocimientos	1								
creo	2	0.08							
Director	3	0	0.25						
experiencia	4	0	0.17	0.75					
formación	5	0.13	0.13	0.43	0.50				
objetivo	6	0	0.15	0.14	0.17	0.11			
orden	7	0	0.17	0	0	0.13	0.17		
perfil	8	0.75	0.07	0	0	0.25	0	0	
persona	9	0.27	0.29	0.15	0.17	0.42	0.07	0.17	0.36
presión	10	0	0.09	0.25	0	0	0	0	0

Para obtener un panorama claro de la estructura total, **HAMLET II** ofrece dos formas gráficas de representación de la información en esta matriz. Puesto que las palabras listadas ocurren con diferentes frecuencias individuales, es apropiado convertir las frecuencias brutas en un coeficiente de concurrencia que considere esto. La medida elegida es el coeficiente de similitud de *Jaccard* -medida de similitud- para datos dicotómicos.

HAMLET II ofrece dos métodos alternativos de generación de clusters o conglomerados jerárquicos que se pueden aplicar en este caso para asignar las entradas individuales del vocabulario a los clusters: el método de “vecino más cercano” o “vínculo simple”, busca la similitud mayor entre un ítem no asignado y aquellos que hay en los clusters existentes; el método del “vínculo completo” o del “vecino más lejano” define la similitud entre los grupos como la similitud entre el par menos similar de los ítems individuales.

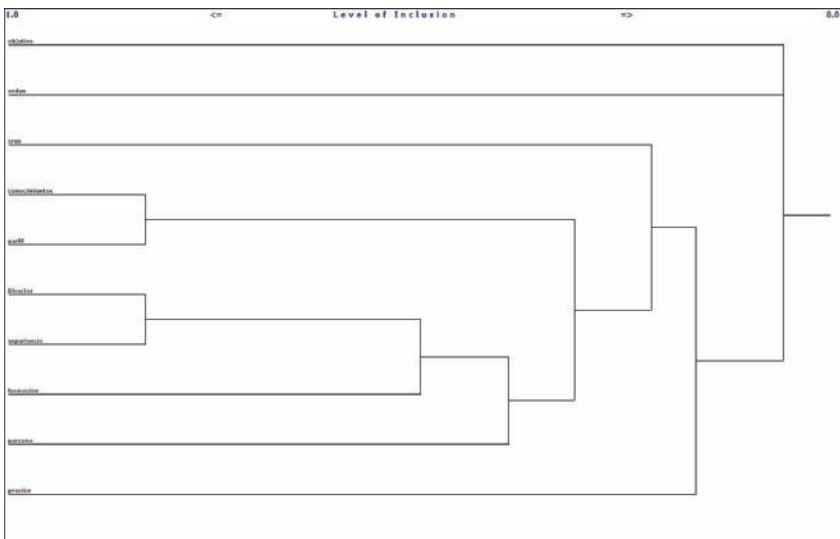


Ilustración 2. Dendograma

El interés primario de **HAMLET II** es aplicar el EMD a la matriz de ocurrencias conjuntas. El procedimiento usado es una versión adaptada de la rutina Michigan-Nijmegen Smallest Space Analysis (MINISSA). Se considera que este procedimiento es conveniente y robusto para reducir la información que contiene la matriz en una dimensionalidad más pequeña, preservando las magnitudes relativas de la similaridad entre los ítems principales del vocabulario que se emplea en el análisis. Representan valores de nivel ordinal.

A partir de las matrices de ocurrencias empataadas se pueden aplicar los métodos de EMD. Para esto es preciso considerar dos materias complejas y que están relacionadas entre sí. A saber, la reducción de dimensiones que permitan explicar los resultados en un porcentaje significativo de certeza y la generación de un gráfico de coordenadas. Conceptualmente las dimensiones de análisis sólo pueden ser representadas gráficamente de dos formas: un plano bidimensional (dos dimensiones) o un volumen (tres dimensiones). En este artículo el modelo de sistema se deducirá desde un análisis bidimensional. El EMD permite contrastar hipótesis y describir la estructura de una matriz de datos. En lo fundamental, se trata de reproducir en un mapa perceptual determinados juicios en relación a determinados objetos bajo ciertas consideraciones. Se trata con datos que no suponen un análisis métrico, dado el tipo de datos, en escala nominal y ordinal.

42

Los resultados obtenidos, como los del análisis de correspondencias, permiten que interpretemos las relaciones entre los “objetos” y las dimensiones que organizan el espacio en el cual se representan. El grado de correspondencia entre las distancias, entre los puntos obtenidos por el mapa de MDS y aquellos de la matriz de entrada, es medido (inverso) por una función de stress. A menor valor del stress (por ejemplo, < 0.10), mayor es la calidad del ajuste obtenido.

HAMLET II permite rotar el diagrama y mover la configuración en torno a los ejes, lo que es arbitrario en estos resultados, pero algún grado de claridad se obtiene.

HAMLET II : display [C:\Windows\Temp\ham(005).tmp]

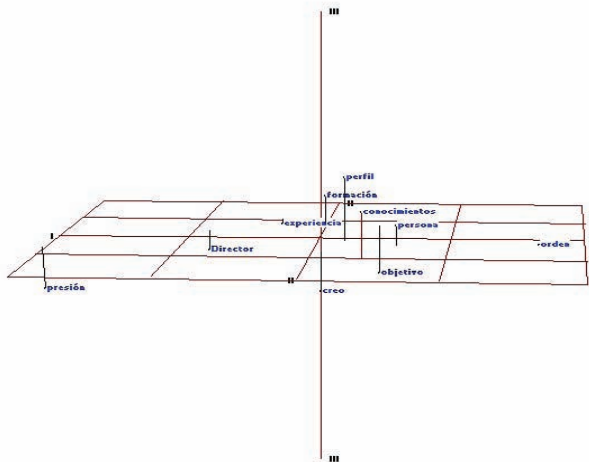


Ilustración 3. Escalamiento multidimensional (EMD)

La siguiente ilustración muestra los mismos datos reducidos a una configuración bidimensional con sus respectivos clusters.

Correspondence Analysis - categories

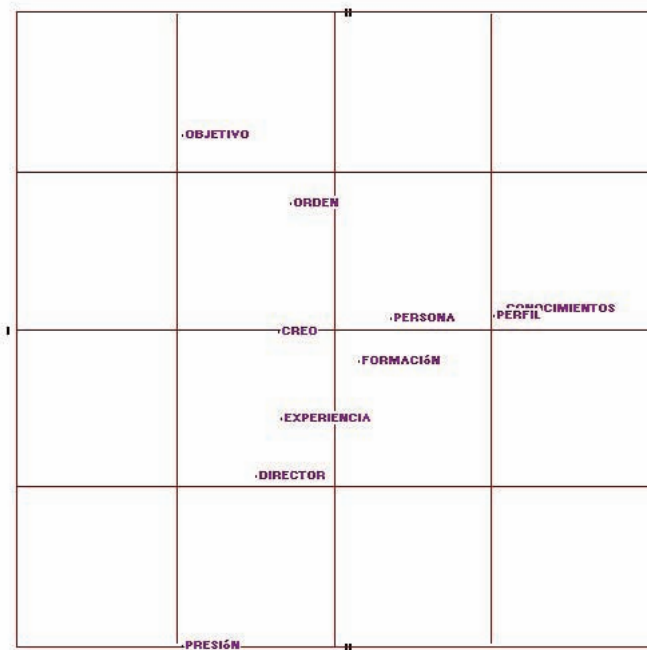


Ilustración 4. Plano bidimensional

Los ítems de mayor preocupación tienden a aparecer en el centro del diagrama. En este caso las creencias (creo), persona, formación. Esto no sorprende por la naturaleza de la entrevista realizada. Preocupan el impacto de las innovaciones tecnológicas en las personas, proceso que puede no estar libre de tensiones. Por otra parte, las cuestiones sobre los procesos de formación adquiridos también constituyen preocupación, en especial en un medio que puede privilegiar la antigüedad o el esfuerzo¹².

2.3. LIWC2007

Una de las técnicas que ha demostrado ser valiosa para estudiar el contenido de las narraciones escritas es el LIWC, Linguistic Inquiry and Word Count (Pennebaker, Francis & Booth, 2001), un programa informático de análisis de textos ampliamente utilizado en psicología. Según sus autores, la manera como los individuos hablan y escriben provee ventanas respecto de sus mundos emocionales y cognitivos. El programa LIWC, basado en una estrategia de recuento de palabras, procesa los textos categorizando las palabras en función de diversas dimensiones: procesos psicológicos, asuntos personales, dimensiones lingüísticas, etc. A partir de este marco conceptual y metodológico se estudian determinados grupos sociales, alumnos, adolescentes en situación de riesgo, etc.

A continuación se presenta un cuadro abreviado de las dimensiones y ejemplos de palabras incluidas en el diccionario validado por los autores y que LIWC2007 utiliza en sus análisis, aclarando que el investigador puede definir su propio diccionario de categorías para someter sus textos a análisis.

Cuadro 6. Muestra de dimensiones y palabras

Dimensión	Ejemplos
I. Dimensiones lingüísticas estándar	
Pronombres personales	yo, ellos, uno mismo
II. Procesos psicológicos	
Procesos sociales	amigos
Amigos	compañero, par
Familia	madre, hermano
Humanos	niño, grupo
Procesos afectivos	feliz, feo
Emociones positivas	alegre, bueno
Emociones negativas	aversión, sin valor
Procesos cognitivos	causa, debería
Discrepancia	podría, debería
Certidumbre	siempre, nunca
Inhibición	bloqueo, obstáculo
III. Preocupaciones personales	
Trabajo	jefe, trabajo
Logro	meta, intento

44

Hay varios estudios interesantes en campos diferentes al de la psicología. Uno de éstos se refiere a analizar cómo los alumnos entienden el estilo de las normas APA¹³. Un segundo estudio alude al análisis de las reflexiones sobre las prácticas en equipos de trabajo¹⁴.

Para ilustrar cómo opera este programa se ha seleccionado una entrevista de un adolescente con problemas con la ley¹⁵. Se encuentra finalizando el periodo de la mediana adolescencia. Se ingresó el texto de la entrevista realizada por la psicóloga (1,004 palabras). Los resultados se presentan en el cuadro 7 que resume los indicadores lingüísticos y sus respectivas dimensiones. Todas las variables reflejan el porcentaje del total de palabras.

Cuadro 7. Procesamiento de la entrevista

Categorías LIWC	Porcentaje
Auto-referencias	5.18
Afecto	
Emoción positiva	2.79
Sentimiento positivo	0.60
Optimismo	0.60
Emoción negativa	0.70
Procesos cognitivos	
Causas	1.49
Insight	1.89
Inclusión	6.27
Exclusión	1.29
Procesos sociales	
Referencia al otro	5.08
Amigos	0.80
Comunidad	2.79
Preocupaciones personales	
Ocupación	4.58
Trabajo	2.19
Logro	1.49
Escuela	1.89

Se trata de un adolescente que ha vivido relaciones interpersonales, y ante situaciones traumáticas, es la relación con los otros un punto de referencia importante, acercándose e interesándose por el otro y con el otro “por ahí me sé quedar en la plaza solo... pensando en la situación...y después te encontrás con un amigo y te da consejos... por ahí te ven solo y bueno, te hacen el aguante...”. Esto queda en evidencia al examinar el cuadro 7 en la dimensión procesos sociales. Muestra interés por las relaciones humanas porque visualiza el diálogo como una alternativa a la resolución de problemas a nivel de interacción “tenés que hablar, lo hablas y podes hacer una amistad, no tenés que pelear. Te enojas y lo hablas...es mejor si lo hablas, hablas y solucionas las cosas”. Los resultados respecto de los procesos cognitivos confirman que a partir de actividades de tipo cognitivo y conductual se pueden registrar cambios en su sentir y pensar: “pensar distinto, sentirme distinto. Por ahí sí, me ayuda pensar las cosas, por ahí se siente distinto”. La dimensión preocupaciones personales revela su interés y acercamiento afectivo hacia amigos, compañeros de trabajo y de la escuela (*dimensión vincularidad*). Recurre a otras personas como referentes ante problemas (amigos, compañeros de trabajo) y toma en consideración la evaluación que de él hagan por sobre la personal, pudiendo en la actualidad adoptar una actitud optimista para con sus circunstancias personales pasadas y actuales. Presenta cierto grado de fortaleza psíquica que le permite evaluar las situaciones previo a actuar, solicitando ayuda cuando las circunstancias lo requieren con capacidad para organizar proyectos personales y económicos de subsistencia propia y familiar (*dimensión funcionamiento psíquico*).

El paso siguiente sería comparar los textos de las entrevistas restantes, verificar cómo el uso del lenguaje puede diferir de entrevistado a entrevistado, y de contexto a contexto. Además se puede proceder a análisis como el de componentes principales y análisis factorial.

Un estudio interesante es *Investigating Moral Events* (2009)¹⁶, que analiza el contenido negativo y positivo de las narraciones y eventos morales. La autora emplea análisis multidimensionales en el tratamiento de los datos mediante el uso del programa STATISTICA, por ejemplo el análisis de cluster y análisis factorial. La tesis es un excelente itinerario metodológico para quienes quieren estudiar las narrativas de determinados actores sociales.

3. Conclusiones

46

Es fundamental para el investigador cualitativo disponer de procedimientos y herramientas analíticas para sus análisis textuales y análisis de contenido. Esto le permite explorar datos y validar sus hallazgos. Trabajar con un corpus textual no es fácil, en especial por la heterogeneidad que puede resultar de los datos recopilados a través de técnicas cualitativas. A través del proceso de análisis van surgiendo regularidades que van siendo conceptualizadas.

El enfoque de este artículo ha sido mostrar la complementariedad de procedimientos para analizar los textos que aborda el investigador cualitativo, pero también el uso de herramientas para integrar análisis cualitativos con análisis cuantitativos multidimensionales.

Notas

- 1 Hay distintas tradiciones para analizar los textos. La tradición lingüística trata el texto como un objeto de análisis por sí mismo; la tradición sociológica/antropológica lo trata como una apertura a la experiencia humana; la tradición psicoanalítica es el estudio de lo no dicho. También hay distintas posiciones entre teóricos, por ejemplo Saussure y Ricoeur. Este artículo se enmarca en el objetivo de buscar la coherencia y el sentido en los datos, en la comprensión del texto como un todo y la interpretación de las partes, y la preocupación por el significado presente en el texto. Aborda la cuestión básica: ¿cuál es el significado de este texto? También es preciso señalar que hay diferentes formas de analizar los textos: análisis de contenido, análisis temático, análisis de la conversación, análisis del discurso, narrativa y metáfora, cada una con sus propias estrategias al momento de recoger y procesar los datos.
- 2 Un artículo interesante al respecto es Shank (1995).
- 3 Se sugiere la lectura del enfoque francés de la estadística textual como herramienta exploratoria en ciencias sociales. Ver Benzécri *et al.* (1981).
- 4 El autor agradece los comentarios de la evaluación de este artículo, y la mejor manera de hacerlo, es incorporando algunas de las ideas del informe recibido por su precisión y rigor.
- 5 Se trata de un neologismo que no aparece en el diccionario de la Real Academia de la Lengua Española, pero el término se puede inferir desde los análisis numéricos y en este texto se entenderá por "individualizar", "separar".
- 6 El discurso da cuenta más bien de la relación entre las intenciones del emisor (destinador) -hablante o escritor- y las del receptor (destinado) -oyente o lector-, implica los procesos correlativos de producción y comprensión y es objeto de una hermenéutica interpretativa. El texto, por el contrario, remite a la emisión misma, el mensaje, en cuanto producto sensible -oral o escrito- y se convierte en objeto del análisis textual. Lo que da vida al texto es la dinámica discursiva que ponen en juego tanto el autor en el proceso de su producción, como el destinatario en el de su interpretación.
- 7 Un excelente artículo sobre este concepto es el de Roberto Follari, *La ciencia como real maravilloso*. El autor argumenta: "La ciencia puede así ser advertida en lo que tiene de "objetivación", más que de "objetividad"; es el fruto de

- una cierta forma de poner los objetos en perspectiva, de captar sus aspectos legaliformes y repetibles, de modo de hacer desaparecer de la percepción aquello que -visto como desordenado- queda fuera de dicho campo de estipulación previa.” Disponible en: <http://www.redcientifica.com/doc/doc200111120001.html>.
- 8 Conjunto de técnicas estadísticas que permiten analizar matrices de semejanza para proporcionar una representación visual de las relaciones entre los datos dentro de un espacio de dimensiones reducidas. Las tablas de input se constituyen de matrices cuadradas que contienen los valores de proximidad (disemejanzas) derivados del cálculo de un índice de asociación.
- 9 Se distinguen por su permanencia, por implicar ausencia de cambio. Son verbos que describen un evento que es un estado que no deja de darse, un estado no delimitable, un estado continuo que no expresa ni cambio ni progreso.
- 10 Transportan la presuposición de que lo representado por la oración completa es verdadero. Un excelente diccionario a consultar sobre verbos estativos y factivos se encuentra disponible en: <http://culturitalia.uibk.ac.at/hispanoteca/lexikon%20der%20linguistik/f/FAKTITIVE%20VERBEN%20%20%20Verbos%20factivos.htm>
- 11 Un excelente diccionario es: *Diccionario de Organización y Representación del Conocimiento. Clasificación, Indización, Terminología*. En el artículo se utilizarán algunas de las definiciones de este diccionario. Disponible en: http://www.eubca.edu.uy/diccionario/letra_a.htm. Consultado el: 5 de mayo de 2009.
- 12 Un estudio referido a la gestión, empleando la herramienta **TROPES** está disponible en: http://www.cavi.univ-paris3.fr/lexicometrical/jadt/jadt2004/pdf/jadt_104.pdf. Consultado el: 20 de abril de 2009.
- 13 Las normas APA constituyen un estilo de referenciación bibliográfica. El estudio está disponible en: <http://minds.wisconsin.edu/handle/1793/23115>. Consultado el: 22 de abril de 2009.
- 14 Disponible en: <http://portal.acm.org/citation.cfm?id=1316624.1316655>
- 15 Esta entrevista es parte de la tesis de Maestría en Criminología de Sorrentino (2008), de la cual fui director de tesis.
- 16 Tesis doctoral de Jessica R. Escobedo, disponible en: <http://etd.caltech.edu/etd/available/etd-11112008-122002/unrestricted/CompleteDissertationJEscobedo.pdf>. Consultado el: 20 de abril de 2009.

Referencias bibliográficas

- BARRY, C. A. (1998). “Choosing Qualitative Data Analysis Software: Atlas.ti and Nudist Compared”. En *Sociological Research Online*, Vol.3 No.3. Disponible en: <http://www.socresonline.org.uk/socresonline/3/3/4.html>. Consultado: 30 de enero de 2009.
- BENZÉCRI, J. P. et al. (1981). *Pratique de l'analyse des données, Linguistique et lexicologie*, Paris: Dunod.
- BRANDON, R. B. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.
- BRIER, A. & HOPP, B. (2006). *Software Manual and User Guide*. “Hamlet II New Generation. A Multidimensional Scaling Approach to Textual Analysis”. Disponible en: <http://www.apb.cwc.net/homepage.htm>.
- ESCOBEDO, J. R., (2009). *Investigating Moral Events: Characterization and Structure of Autobiographical Moral Memories*. Tesis doctoral, California Institute of Technology. Disponible en: <http://etd.caltech.edu/etd/available/etd-11112008-122002/unrestricted/CompleteDissertationJEscobedo.pdf>. Consultado: 20 de abril de 2009.
- FOLLARI, R. (s.f.). *La ciencia como real maravilloso*. Disponible en: <http://www.redcientifica.com/doc/doc200111120001.html>. Consultado: 26 de abril de 2009.
- GHIGHLIONE, R. & BROMBERG, M. (1998). *Discourse Politique et Television. La verite de l'heure*. Paris: Presses Universitaires de France.
- LEBART, L. & SALEM, A. (1994). *Statistique Textuelle*. Paris: Dunod.
- MEUNIER, F. (1998). “Computer Tools for Interlanguage Analysis: A Critical Approach”. En Granger, S. (ed.) *Learner English on Computer*. London & New York: Addison Wesley

Longman, pp. 19-37.

PENNEBAKER, J. W., BOOTH, R. J., & FRANCIS, M. E. (2007). *Linguistic Inquiry and Word Count (LIWC2007): A text analysis program*. Austin, TX: LIWC.net.

PENNEBAKER, J. W., FRANCIS, M. E. & BOOTH, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC)*: LIWC . Mahwah, NJ: Erlbaum Publishers.

PIOLAT, A., & BANNOUR, R. (s.f.). "EMOTAIX : Un Scénario de TROPES pour l'identification automatisée du lexique émotionnel et affectif". *L'Année Psychologique*. Disponible en: <http://sites.univ-provence.fr/wpsycle/membres/enseignants/apiolat.html>. Consultado: 24 de abril de 2009. mimeo.

RASTIER, F. (1989). *Sens et textualité*, Paris: Hachette.

REINERT, M. (1993). « Les «mondes lexicaux» et leur logique ». En *Langage et société*, Paris, Maison des Sciences de l'Homme, No.66:5-39.

REINERT, M. (1986). Un logiciel d'analyse lexicale : ALCESTE. En *Cahiers de l'analyse des Données*, 4, 471-484.

SHANK, G. (1995). "Semiotics, Qualitative Research in Education: The Third Crossroad". En *The Qualitative Report*, Vol.2, No.3, diciembre. Disponible en: <http://www.nova.edu/ssss/QR/QR2-3/shank.html>. Consultado: 30 de agosto de 2008.

48

SORRENTINO, L. (2008). *Resiliencia y factores de riesgo en varones de 14 a 18 años con conductas transgresoras de la ley penal*. Tesis de Maestría en Criminología. Facultad de Psicología, Universidad del Aconcagua, Mendoza, Argentina.

TOURNIER, M. (1980). « D'où viennent les fréquences de vocabulaire ? », En *Mots*, Paris: Presses de la fondation nationale des sciences politiques, No.1:189-209.