

PROTOCOLO DE TRANSCRIPCIÓN PARA EL ANÁLISIS COMPUTARIZADO DEL CONTENIDO DE SESIONES DE PSICOTERAPIA, ENTREVISTAS Y OTROS TEXTOS DE INTERÉS CLÍNICO EN ESPAÑOL¹

E. Mergenthaler², S. Gril³, A. Ávila-Espada⁴ y J. J. Vidal⁵

Although use of psychotherapy transcripts and other clinical materials is becoming increasingly important in psychotherapy process research, large-scale collaborative work is hindered by lack of suitable transcription standards. New guidelines more detailed and complete derived from preceding researches (Mergenthaler y Stinson, 1992; Mergenthaler y Gril, 1996) are presented here for the transcription of Spanish discourse. These guidelines may be used not only for psychotherapy research but also for other purposes (such as education, linguistics, anthropology, among others). This protocol for transcripts arises here will be used for several tasks: a) It will be used for transcripts and readable by human judges; b) it will also be easily submitted for computer-aided text analysis, making easier to obtain quantitative indicators in the content analysis and the adequate preparation of texts for the ulterior qualitative analyses. It is presented a brief summary of the main computer-aided text analysis.

Key words: Psychotherapy Research, Transcription Rules, Text Analysis

RELEVANCIA DE LA INVESTIGACIÓN DEL ANÁLISIS DE CONTENIDO

La investigación actual en el campo de la psicoterapia y de la psicoterapia psicoanalítica está fundamentalmente basada en material primario (Luborsky y Spence, 1978) correspondiente a protocolos verbales que contienen transcripciones literales de sesiones de psicoterapia. Esto puede observarse en los principales encuentros científicos donde se presentan y discuten investigaciones, como los de la Sociedad de Investigación en Psicoterapia, donde cada vez más se presta mayor atención a ese material como base de la investigación empírica. Podemos dividir las distintas líneas de investigación entre aquellas que utilizan categorizaciones del material producidas por evaluadores o jueces humanos y aquellas que se basan en métodos computerizados. En función de la cuestión a investigar los investigadores

han desarrollado diferentes formatos especializados de transcripción. Pero a pesar de que un formato de transcripción puede resultar útil para un análisis específico, puede ser de poca utilidad o relevancia para otros estudios.

Esto causa problemas, especialmente en los estudios comparativos, colaborativos o convergentes. Se hace difícil realizar comparaciones aplicando métodos analíticos alternativos a las mismas sesiones, pero con diferentes normas de transcripción. El lector queda entonces abandonado a imaginar cómo fueron manejadas por el transcriptor las palabras interrumpidas por el tartamudeo, las alteraciones del discurso y muchas otras interferencias y producciones singulares.

Transcribir intercambios verbales de registros de audio y vídeo presenta innumerables, complejos y sutiles problemas para el investigador. Lo que hace que una transcripción sea adecuada para su utilización en los análisis mediante computadora puede hacerla indescifrable para un lector humano y viceversa.

Desarrollar reglas siempre es difícil. En tanto parece poco probable que exista una regla completa de transcripción que sea aceptada universalmente, es también obvio que desarrollar algunas aproximaciones estandarizadas resultarán de suma utilidad para la investigación.

El lenguaje, desde el punto de vista de la semiótica, es un sistema de símbolos cuya estructura está determinada por reglas que se basan en la relación entre forma y contenido. En consecuencia distinguimos mediciones formales, gramaticales y sustantivas. Dependiendo de si se trata del texto de un único hablante o del discurso completo de los participantes en un diálogo, se podrán introducir subdivisiones en cada una de estas mediciones. Este tipo de mediciones posibilita hablar tanto de valores monádicos o diádicos como emplearlas para detectar en el discurso de un paciente patrones repetitivos condicionados por su enfermedad.

Las posibilidades de las mediciones formales, gramaticales y sustantivas varían ampliamente. En el análisis formal, la más conocida es la frecuencia de aparición, que sirve de base para las proporciones y las distribuciones estadísticas. Uno de los ejemplos más antiguos es la variabilidad del discurso del paciente, que se mide por la razón tipo-signo, que se puede considerar como la capacidad de hacer uso de palabras nuevas o distintas en el transcurso de la sesión de psicoterapia o, incluso, en todo el tratamiento. Esta razón se calcula dividiendo el número de palabras diferentes (tipos, amplitud del vocabulario) entre el número total de palabras (signo, extensión del texto) que existen en un texto determinado. Chotlos (1944) demostró que esta medición no es independiente de la extensión del texto. Herdan propuso emplear el logaritmo del valor de la razón «tipo-signo», cifra que puede ser constante para muestras de textos de extensión variable. En lo sucesivo nos referiremos a esta medición como «Gamma». De acuerdo con Holsti (1969), «la hipótesis según la cual la variabilidad del discurso puede verse incrementada con un tratamiento adecuado ha sido generalmente confirmada» (p. 75). Si bien en las últimas décadas no se ha empleado demasiado la razón «tipo-signo» en la investi-

gación psicoterapéutica, existen algunos estudios literarios que son importantes. Baker (1988) ha encontrado que en las obras de Shakespeare y otros autores de renombre, el ritmo (su sinónimo para la proporción «tipo-signo») está en relación con la madurez y desarrollo de sus autores (p. 37). En el contexto de la investigación psicoterapéutica consideramos que el aumento en la versatilidad verbal del paciente, reflejado en la relación tipo-signo, es un signo de avance (hecho que suele estar relacionado, aunque no siempre, con la mejoría del paciente); teniendo en cuenta lo anterior, podemos considerarla como una medición objetiva del proceso psicoterapéutico tanto desde la perspectiva macroanalítica como desde la microanalítica. El concepto de «avance», lo empleamos aquí con el mismo sentido con que se emplea en el trabajo clínico. Lingüísticamente hace referencia al empleo de un mayor número y diversidad de palabras por parte del paciente para descubrir el mismo núcleo temático.

Existen otros aspectos formales del discurso que pueden ser analizados desde la perspectiva de la teoría de la información. A partir del trabajo de Shannon (1948), han comenzado a adquirir importancia los conceptos de entropía y redundancia. Spence (1968) fue quien introdujo estos conceptos en la investigación psicoterapéutica: «Se define la redundancia como el complemento de la aleatoriedad; si existe un conjunto de sucesos altamente redundantes, la aparición de los primeros permite predecir la aparición posterior de los demás con un alto grado de fiabilidad. Sin embargo, una característica distintiva del material que se maneja en la clínica psicoterapéutica es el alto grado de incertidumbre. Debido a ello, por definición, la redundancia es baja» (p. 354). Spence (1968), en su análisis considera que en la medida que se vaya avanzando en la terapia, cabría esperar que la redundancia fuera incrementándose gradualmente. «En la medida que la adquisición gradual de insight por parte del paciente se vaya acompañando de una resolución adecuada de sus conflictos éste podrá al final del tratamiento, alcanzar una posición que le permita comprender una mayor variedad de comportamientos que hasta ahora le resultaban confusos. El paciente será capaz de ver qué comportamientos aparentemente aleatorios forman parte de un patrón mayor y reconocer las repeticiones de un tema subyacente, lo que le permitirá ser más redundante que lo que era anteriormente» (p. 358). Nosotros interpretamos el incremento en la redundancia del paciente que aparece en el transcurso de la terapia como la adquisición de una mayor «libertad interna» por parte de éste, hecho que contrasta con momentos en la psicoterapia en los que predomina la resistencia y la redundancia será por ende, menor.

Spence y Owens (1990) introdujeron recientemente los conceptos de co-ocurrencia léxica y fuerza de asociación como aplicaciones de una medición formal. Esta medición, que en un estudio referido a la monitorización de la superficie analítica fue aplicada a los patrones existentes en el discurso del paciente en el contexto de su análisis, era empleada en función de la tasa de co-ocurrencia de los

pronombres de transferencia que son los pares «tú / a mí, me», «a mí / yo», «tú / yo». Los autores concluyeron que las manifestaciones concretas hechas por el paciente eran repetidas sistemáticamente en la intervención del terapeuta» (p 43).

Como ejemplo de las mediciones gramaticales tenemos el empleo de las partes del discurso tal y como se comentó anteriormente. Sabemos, por estudios recientes, que los pacientes difieren significativamente entre sí en cuanto al uso que hacen de las categorías sintácticas; así, por ejemplo, los pacientes neuróticos tienden a emplear más verbos que los pacientes psicóticos. La llamada «distancia del discurso en partes» se refiere al uso distintivo que hacen el terapeuta y el paciente de las partes del discurso, (en la sección de métodos se hablará del cálculo de esta medición). La distancia del discurso en partes es una medida diádica que generalmente está más allá del control de los que participan en el discurso, y nos da una idea del grado de interacción que puede existir entre el paciente y el terapeuta. Consideramos que si existe una buena «alianza de trabajo» existirá una menor separación y una conducta más colaboradora entre ambos participantes. Esto significa que ambas partes cambiarán el uso de las clases de distancia del discurso, existiendo así una proporción más parecida de ambas clases en ambos participantes. Existen estudios referidos a la construcción interactiva de frases que aportan evidencia lingüística que sirve de apoyo para la interpretación de procesos tales como la «alianza de trabajo».

La Abstracción es otro de los ejemplos que se incluyen dentro de las mediciones gramaticales y que se refiere a los aspectos morfológicos; es un constructo orientado al desarrollo de la comprensión y de la percepción, teniendo además efectos lingüísticos evidentes. El lenguaje, además de ser una fuente rica de sustantivos abstractos, permite construir nuevos vocablos abstractos adicionales. Así, partiendo de conceptos concretos se puede realizar una transformación morfológica de las palabras; por ejemplo, a partir de «ser tierno» construimos fácilmente la palabra ternura. La abstracción es una herramienta lingüística necesaria para la reflexión, discusión y descripción de interacciones complejas.

La expresión «análisis de contenido» se refiere generalmente a mediciones sustantivas. Los diccionarios de análisis del contenido han sido empleados en diversas áreas entre las que se incluyen el lenguaje de la afectividad, temas relacionados con la ansiedad, imágenes de regresión y el tono emocional. Estos diccionarios, que se han empleado tanto en estudios longitudinales como transversales, son útiles a los efectos de la clasificación cuando se calcula la distribución de frecuencias de estas categorías o sub-categorías. También sirven como instrumentos de detección selectiva para localizar momentos claves de una sesión terapéutica o para localizar sesiones importantes a lo largo de una serie de transcripciones que sean más tarde susceptibles de análisis más detallados. Además pueden ser fácilmente distribuidos entre los investigadores, con lo que se facilita los esfuerzos conjuntos.

Si bien las mediciones formales son independientes del lenguaje concreto en que se hable, en las mediciones gramaticales como por ejemplo la distancia del discurso en partes y en las técnicas de análisis del contenido sustantivo será necesario recurrir a enfoques que sean específicos del lenguaje. Existe un número creciente de diccionarios, disponibles en inglés principalmente, francés, alemán y español, que incluyen procedimientos para identificar categorías sintácticas y palabras abstractas. Algunos de estos diccionarios analíticos del contenido han sido desarrollados en un idioma y luego traducidos y adaptados a otro. Un ejemplo de ello es el *Diccionario de Imágenes regresivas* de Martindale (1975) que fue desarrollado en inglés y luego traducido al alemán y al francés. Otros han sido desarrollados independientemente en varias lenguas; este es el caso de los *Diccionarios de Emoción y Abstracción* de Mergenthaler (aquellos que estén interesados pueden tener una copia escribiendo al Banco de Casos de Ulm: utb@sip. medizin. uni-ulm. de, o como parte del software CM cuando son diccionarios *inbuilt*, esto es, incorporados de forma automática en el programa), que fueron desarrollados primero en inglés, y más tarde en alemán, español, holandés, italiano y francés. En algunos casos se dispone solamente de la versión en uno de los idiomas. Entre estos figuran el *Diccionario de Temas de Ansiedad* publicado en alemán y la *Actividad Referencial* (E. Mergenthaler y W. Bucci) manuscrito aún sin publicar. Las mediciones y diccionarios empleados en este estudio son o bien independientes del idioma o bien están disponibles en versión inglesa.

CONSIDERACIONES TÉCNICAS SOBRE LA GRABACIÓN Y LA TRANSCRIPCIÓN

Sin la formulación de reglas definidas para la grabación y transcripción, los materiales que se obtengan en distintos lugares no podrían ser comparados, o generarían un volumen de problemas que dificultaría la investigación. Sistematizar las reglas de transcripción no sólo beneficia a los métodos de análisis de textos asistidos por computadora, sino que dichas reglas pueden ayudar a sincronizar investigaciones multidisciplinarias convergentes abarcando registros de audio y vídeo, parámetros fisiológicos, escalas de valoración (tipo Likert u otras), tanto previas como posteriores a la sesión, así como otros datos. Además los textos transcritos no deberán presentar incompatibilidades para su tratamiento informático a través de los principales sistemas de análisis de textos (TAS/C, USALc, Code-A-Text, Generador de Concordancias, Útiles de Texto)

El sistema de reglas para la transcripción, que se pretende independiente de los objetivos específicos para la investigación, se organiza en base a seis principios, resumidos a continuación:

1. Preservar la naturaleza morfológica de la transcripción. La presentación gráfemica de la forma de las palabras, la forma de los comentarios y el uso

de la puntuación deben ser lo más semejante posible a la presentación y el uso generalmente aceptado en textos escritos.

2. Preservar la naturaleza de la estructura de la transcripción. El formato impreso debe ser lo más semejante posible a lo que es generalmente aceptado, como las versiones impresas de obras de radio o guiones cinematográficos. El texto debe estar claramente estructurado por indicadores del discurso.
3. La transcripción debe ser una reproducción exacta. La pérdida de información resultante de la transcripción de registros visuales o acústicos a registros escritos de la entrevista debe ser la mínima posible. Una transcripción no debe ser reducida prematuramente y debe permanecer en forma de datos “directos”, es decir, tal cual se produjeron.
4. Las reglas de transcripción deben ser universales. Deben posibilitar en la mayor medida posible que la transcripción sea adecuada tanto para uso humano como para la utilización de computadoras, y los diversos procedimientos de análisis de contenido.
5. Las reglas de transcripción deben ser completas. Debe ser posible para quien realiza la transcripción prepararla utilizando sólo esas reglas, basadas en su lenguaje diario. No se requieren conocimientos específicos como p. e. “codificadores” provenientes de teorías lingüísticas. Debe ser posible transcribir distintos tipos de discurso terapéutico y clínico con el mismo grupo de reglas. Estas reglas deben ser independientes de quien realiza la transcripción, comprensibles y aplicables tanto por personal administrativo como por científicos.
6. Las reglas de transcripción deben ser intelectualmente elegantes. Deben ser elegantes, limitadas en su número, simples y fáciles de aprender.

De acuerdo a la experiencia con estas reglas se estima que un mecanógrafo/a hábil requerirá aproximadamente 8 horas para realizar el primer paso de una transcripción de una sesión de psicoterapia cuya duración esté en torno a los 50 minutos, si la fluidez verbal es normal. En un segundo paso, se necesitan otras 4 horas para la verificación y corrección. La transcripción y la verificación puede requerir un 50% más de tiempo para el discurso que sea grabado con dificultades técnicas o para discursos poco claros, así como en discursos con una acentuación muy marcada, discurso idiosincrásico, o discurso poco fluido con tartamudeo. Incluir la codificación del tiempo, si es necesario, puede requerir una hora o más. Los verificadores necesitan tener buenas habilidades de lectura y de escucha, con cierto entrenamiento previo; la tarea puede ser llevada a cabo por mecanógrafos, secretarios, alumnos de postgrado, ayudantes de investigación o los propios científicos.

Los transcripores de los materiales deben ser individuos atentos a los detalles, motivados, interesados en el trabajo, conscientes del valor que tiene para la

investigación. No deben tener límites rígidos de tiempo para completar las transcripciones, permitiéndoseles trabajar sin interrupciones innecesarias. Existen diferencias en el tiempo que lleva transcribir diferentes sesiones y el transcriptor necesitará familiarizarse con las características del discurso de cada nuevo hablante. Es asimismo necesario que los transcriptores y los verificadores estén alertas frente a la posibilidad de experimentar sentimientos similares a los expresados en la psicoterapia, especialmente depresión. Deben tener la oportunidad de tomarse descansos así como de simultanearlos con otros materiales, menos cargados de emociones. Por último hemos de resaltar el especial celo que debe tenerse en salvaguardar las garantías éticas de respeto a los sujetos implicados en la investigación, de acuerdo a las normas y convenios al respecto, y que en el caso de la transcripción se concreta principalmente en la salvaguardia del anonimato, y la codificación-criptación de todos los nombres propios, geográficos u otros que pudieran conducir a la identificación de las personas. Un especial celo en el cuidado de las normas éticas es responsabilidad tanto de los investigadores como transcriptores, así como una custodia confidencial de los materiales, cuya difusión estará restringida según los acuerdos a que en cada caso se haya llegado.

MÉTODOS DE ANÁLISIS DE CONTENIDO POR COMPUTADORA

El análisis de contenido asistido por ordenador se puede considerar como una variante mecanizada del análisis clásico de contenido. El método usa listas de palabras ordenadas temáticamente en categorías. Estas listas se llaman diccionarios las cuales se comparan con el texto analizado, y se contabiliza a continuación la frecuencia de ocurrencia de cada palabra única, dando como resultado una distribución de frecuencias para las categorías asociadas. Como asunción básica del análisis de contenido, esperamos que cuanto más prominente sea un tema en un texto, más referencias podamos encontrar en él. Berelson (1952) y Kracauer (1952) objetaron esta asunción a partir de un estudio que llevaron a cabo en el que se centraron en los aspectos del análisis de contenido cualitativo vs. cuantitativo. Aunque esta controversia nunca fue resuelta, su importancia decreció en los años sucesivos, por lo se adoptó una postura pragmática. Kracauer (1972) y más tarde Howe (1988) señalaron que las dos aproximaciones coincidían en parte con análisis cuantitativos que conducían a consideraciones cualitativas y con análisis cualitativos que con frecuencia requerían una cuantificación. También esta controversia llegó a ser menos crítica a medida que se produjeron más resultados por medio de los métodos cuantitativos, particularmente los análisis de contenido asistidos por ordenador que se dirigían a temas cualitativos (Oxman et al. 1988a, 1988b; Rosenberg, Scnurr & Oxman, 1990). No se trata de cuantificar los diversos significados o los contenidos de un texto sino más bien, identificar aspectos temáticos más generales. Una entrada

del diccionario que se empareje con una palabra del texto analizado es considerada un «marcador» el cual indicará la presencia de un constructo temático. Según esta aproximación, la controversia sobre los aspectos cualitativos y cuantitativos se vuelve de nuevo obsoleta al interpretar los hallazgos temáticamente cualitativos con respecto a su significado. Los análisis de contenido asistidos por ordenador tal y como se proponen aquí, proporcionan la siguiente información: ¿están en el texto los fenómenos de interés? y si es así, ¿dónde se localizan?

Atendiendo a los métodos computacionales con los que este protocolo de transcripción se puede utilizar, podemos iniciar nuestra exposición por el TAS/C© (Mergenthaler, 1993). Este software fue específicamente concebido para aplicarse al campo de la investigación psicoterapéutica y ofrece muchas maneras diversas de investigar transcripciones literales. Así, se pueden hacer análisis separados del discurso del paciente y del terapeuta aunque también se puede analizar un texto completo sin hacer esta distinción. Además se pueden llevar a cabo análisis para cualquiera de los segmentos marcados en un texto, como por ejemplo, las unidades dadas para los marcadores del tiempo, los bloques de un número equivalente de palabras o la identificación de los sueños, recuerdos tempranos u otros episodios. El TAS/C© está diseñado para manejar grandes cantidades de texto. Así por ejemplo permite analizar todas las transcripciones de un tratamiento completo de una sola vez. El software del TAS/C© requiere por ahora, de una unidad de entorno. Se está desarrollando una versión para Macintosh© y para ordenadores personales que usan Windows©. El software del TAS/C© utiliza una interface de datos que permite transferir sus resultados a otros programas entre los que se incluyen SYSTAT© para Macintosh© de cara a los análisis estadísticos y Delta Graph Pro© para Macintosh©, con el fin de representar gráficamente los hallazgos.

Para analizar sesiones de psicoterapia teniendo como marco teórico el Modelo del Ciclo Terapéutico (Mergenthaler, 1996) se puede utilizar el software CM©. El CM© combina el Modelo del Ciclo Terapéutico (Mergenthaler, 1996) con el Ciclo Referencial (Bucci, 1997) y proporciona una representación gráfica de las transcripciones de las sesiones psicoterapéuticas. En concreto, este programa usa como input una transcripción realizada de acuerdo a las reglas anteriormente señaladas y produce como output una representación gráfica de la sesión basada en bloques de palabras que usan la Emoción (positiva y negativa), la Abstracción y la Actividad Referencial Computerizada (Mergenthaler y Bucci, 1999) como medidas del lenguaje. Opcionalmente el CM© proporciona otros muchos archivos de datos para análisis estadísticos. El programa está disponible gratuitamente en la siguiente dirección de internet: <http://inf.medizin.uni-ulm.de>

Un método innovador y desarrollado recientemente es USALc© (Sotelo y Ávila, 2000) una herramienta pensada y creada para realizar estudios en los que el propio lenguaje, la expresión verbal, la sintaxis, etc., se consideran y tratan como

causa y efecto de cambios en la forma y estilo con que un paciente determinado refleja su realidad, sus conflictos, su percepción del mundo y la de los demás a través de sus expresiones habladas. USALc© está formado por una serie de utilidades conectadas y fuertemente compenetradas que favorecen el trabajo de análisis y de clasificación. Utiliza como material de trabajo sesiones de Psicoterapia grabadas y pasadas a texto según un protocolo de transcripción estandarizado. Estos textos han de formatearse según ciertas convenciones a través de una serie de sencillos pasos con el fin de regularizar las entradas de texto en el corpus de análisis principal de USALc©.

Cuando en una investigación se recogen datos verbales (u otros) a partir de técnicas de observación nos encontramos ante la gran problemática de su codificación. En estos casos no resulta tan fácil obtener datos cuantitativos directamente, por lo que se suele recurrir a utilizar un método de análisis cualitativamente flexible pero adecuado. Nos referimos al Análisis de Contenido, en este caso de Textos sacados de sesiones de Psicoterapia. Un caso simple de este tipo de análisis sería el conteo de palabras de una determinada población de textos (p. e. artículos de revistas) formulando las hipótesis adecuadas. En este caso se obtendrían las frecuencias de aparición de una serie de palabras que pueden después utilizarse en el análisis tanto como variables independientes como dependientes. Podemos decir que esta técnica de trabajo esta abierta tanto a la creatividad del investigador como a su experiencia. A pesar de su flexibilidad, este tipo de análisis precisa de una cierta y mínima estructuración. Distinguir lo latente de lo manifiesto no sólo sería útil a nivel metodológico y práctico, también lo sería a nivel lógico y conceptual. De esta organización primaria puede depender en gran medida si se obtienen resultados satisfactorios o no. Para el tipo de tarea que realizaremos con USALc© se distingue entre lo que el sujeto dice (lo manifiesto) y lo que el sujeto quiere decir o lo que dice entre líneas, (conscientemente en el primer caso y de forma inconsciente o no totalmente consciente en el segundo. En ambos casos estaremos ante un análisis latente). En definitiva, USALc es una potente aplicación con múltiples funciones para gestionar y analizar textos de diverso tipo y procedencia. Está especialmente orientado al análisis de diálogos, más orientado al análisis de entrevistas, sesiones de terapia, etc. Una de sus grandes ventajas es que los objetos en los que se almacena la información de cada análisis son bases de datos con un 100% de compatibilidad con Access. Al ser máxima la compatibilidad el usuario puede abrir y trabajar las bases de datos creadas por USALc en y desde Access. También cuenta con la posibilidad de trabajar con audio como complemento y varias utilidades adicionales (Doctor USALc, Contrastes «a posteriori», etc) que confieren mayor funcionalidad a esta aplicación.

Otra herramienta codificadora de reciente difusión es el Code-A-Text©, (Cartwright, 1996) un programa de software diseñado para fomentar la investiga-

ción y la formación en psicoterapia ya que es un instrumento que facilita el análisis en profundidad de los textos psicoterapéuticos. Además de para analizar los diálogos de la terapia, puede utilizarse también para almacenar y codificar anotaciones del proceso, anotaciones del caso e informes y datos cualitativos obtenidos a través de cuestionarios de respuesta abierta. Las grabaciones de los diálogos pueden analizarse por separado o junto con la transcripción de los textos.

A lo largo de la conversación que el terapeuta y el paciente mantienen durante la sesión terapéutica, cada interlocutor tiene un turno de palabra. En el Code-A-Text, las palabras habladas en cada turno constituyen una Unidad del Discurso, que puede subdividirse en segmentos más pequeños. Éstos a su vez pueden ser codificados en función de las palabras contenidas en cada turno (Códigos de Contenido); en función de las interpretaciones vinculadas a los segmentos (Códigos de Interpretaciones) o también por último, se pueden aplicar los códigos a las escalas desarrolladas con fines de investigación (Códigos de Escalas). Utilizando estos códigos, el investigador puede examinar las relaciones entre los segmentos y usar los procedimientos de clasificación que el Code-A-Text proporciona, para comparar o contrastar los segmentos.

Cada proyecto o estudio que utiliza el Code-A-Text usa un único Marco de Codificación, consistente en una base de datos de carácter relacional que contiene todos los elementos requeridos para el análisis. Cada marco de codificación consta de:

- Un Archivo: un diccionario de todas las palabras del proyecto.
- Un Manual de Codificación: que contiene todas las escalas, los códigos y las instrucciones de codificación.
- Textos: las transcripciones de los diálogos.
- Códigos: los cuales se han aplicado a los segmentos del texto.

Las palabras y las frases del texto así como las interpretaciones del investigador pueden ser activas (de interés para el estudio) o pasivas. Cada palabra activa se representa por medio de una escala dicotómica, la cual tiene dos valores: verdadero o falso. Si el Code-A-Text encuentra una *palabra* activa en un segmento, entonces ese segmento es codificado como verdadero. Esto se hace de forma automática y toda la información es almacenada en el archivo. En cuanto a las *interpretaciones*, el Code-A-Text realiza la misma función, sólo que no lo hace de forma automática y además no almacena la información. En lo que respecta a las *escalas*, no hay limitaciones numéricas referidas a la cantidad de las mismas que se pueden utilizar.

El Code-A-Text ofrece un rango de búsqueda y una serie de funciones analíticas que ayudan al investigador a identificar los patrones subyacentes a un texto o textos de marco de codificación. Estas funciones permiten establecer complejos criterios de búsqueda y sus resultados nos servirán para la creación de nuevos textos y escalas de codificación, además de proporcionarnos datos útiles para programas estadísticos.

Hay otras herramientas complementarias, de especial interés, como el Generador de Concordancias (CONC©) (Thompson, 1990) es un programa diseñado para facilitar el estudio intensivo de un texto mediante la producción de una lista de todas las palabras contenidas en él, incluyendo una breve referencia del contexto anterior y posterior a la ocurrencia de cada palabra. En muchas áreas de estudio dicha lista se llama Concordancia, y así la vamos a denominar en la exposición que sigue. La concordancia se realiza en tres columnas. La del margen izquierdo señala el lugar en que la palabra aparece en el texto. La columna central indica el contexto inmediatamente anterior a la ocurrencia de la palabra. La columna del margen derecho contiene una lista de las palabras incluidas en el texto por orden alfabético. Cada palabra es seguida de una pequeña referencia del contexto posterior a la ocurrencia de la misma. Entre las varias aplicaciones o utilidades que podemos dar a CONC, destacamos que permite seleccionar las palabras que se incluyen en la concordancia; bien haciendo una lista de ellas individualmente o bien usando un patrón para describirlas. Además da la posibilidad de escoger u omitir palabras cortas o largas así como palabras que ocurren frecuente o infrecuentemente. También puede detectar solamente, si así conviene, múltiples ocurrencias de palabras cuando son clasificadas conjuntamente. Asimismo ofrece la posibilidad de realizar en el texto estudios estadísticos simples como por ejemplo, contar el número de ocurrencia de las palabras que cumplen un determinado patrón. También se puede con este análisis establecer otras concordancias con archivos de textos simples y también con archivos de textos interlineares. Un texto interlineal consta de un texto de línea base más otra u otras líneas de anotaciones que están ordenadas verticalmente.

Además de la concordancia de palabras o de morfemas, CONC también puede producir una concordancia de todas las letras (caracteres o fonemas) de un texto. La concordancia de una letra puede limitarse justo a aquellas letras que ocurren en un entorno particular. Esto puede resultar muy útil para hacer análisis fonológicos. También añadir que las concordancias de letras pueden hacerse sobre textos simples o interlineares. En la actualidad, CONC puede sólo tratar la relación entre dos campos. Uno de ellos es el llamado primario; contiene las palabras que aparecen en el contexto de la concordancia. El otro, llamado secundario, contiene anotaciones que CONC usará para modificar la secuencia de clasificación o los patrones de inclusión de palabras que podrían utilizarse si hiciésemos simplemente una concordancia del campo primario.

Por último reseñaremos una herramienta sencilla, Útiles de Texto©, una aplicación de Macintosh diseñada por los miembros del Instituto de Verano de Lingüística (S. L. I., Dallas). Permite evaluar diversas características de un texto aportando varios tipos de información:

- 1- Lista de palabras: Proporciona una lista de todos los tipos de palabras diferentes que aparecen en el texto con un recuento de la frecuencia de ocurrencia de cada palabra.
- 2- Recuento de carácter: Muestra los distintos caracteres ortográficos contenidos en el texto con un recuento de la frecuencia de ocurrencia de cada carácter utilizado.
- 3- Recuento diacrítico: Ofrece una lista de todos los caracteres acentuados que aparecen en el texto con un recuento de la frecuencia de ocurrencia de cada carácter.
- 4- Formato estandar de marcadores: Se trata de un código que marca el inicio y el final de un texto determinado. Está basado en estándares usados por el Instituto de Verano de Lingüística.
- 5- Apuntes del proceso: Informe sobre los errores contenidos en el texto. Este procedimiento nos va a permitir por lo tanto hacer un análisis exploratorio inicial de la estructura de los textos dialogados, que complementa análisis realizados por los otros sistemas mencionados: TAS/C, USALc, Code-A-Text y CONC.

MÉTODO DE TRABAJO Y PROTOCOLO DE TRANSCRIPCIÓN PROPUESTO

A continuación y como apéndice, se incluyen los documentos producidos que incluyen: a) Instrucciones generales sobre la transcripción; y b) Reglas ejemplificadas del Protocolo de Transcripción de Sesiones de Psicoterapia en Español. Aunque este método y sus dificultades viene siendo contrastado desde hace varios años, se prevé efectuar revisiones del mismo, recogiendo las sugerencias de los investigadores que lo utilicen.

INSTRUCCIONES DE TRANSCRIPCIÓN DE SESIONES DE PSICOTERAPIA PARA EL ANÁLISIS DE CONTENIDO¹

Material Clínico:

* La(s) cintas entregadas, las cuales serán custodiadas y tratadas de acuerdo a las condiciones establecidas en el *Contrato de Confidencialidad*.

Equipamiento necesario:

* Magnetófono de cassettes (y auriculares)

* Protocolo de Transcripción de Sesiones.

* Listado de claves.

* Ordenador con procesador de textos (compatible con Word 6.0, Word 97). Si no se dispone del mismo, se utilizarán en los horarios apropiados el aula de informática y/o los de la Unidad de Investigación.

Procedimiento a seguir:

* Primera escucha y transcripción palabra por palabra, según las reglas del Protocolo de Transcripción

* Si la transcripción no ha sido hecha directamente al procesador de textos, ésta se hará antes de la revisión.

Segunda escucha, revisión de la transcripción y toma de tiempos.

[Se recomienda que las tareas de escucha, transcripción y revisión sean recíprocas; En cada equipo, integrado por dos personas, se sugiere que se alternen en dichas tareas, a fin de reducir errores y vicios de transcripción]

* Las cintas y su transcripción en papel y diskette serán entregados en las fechas acordadas.

Instrucciones especiales:

* Si no se indica lo contrario, todo nombre propio de persona o lugar se pondrá la primera en mayúsculas precedida de un asterisco y si no aparece en el listado de claves, el nombre completo irá en mayúsculas. De esta manera se facilita su posterior encriptación.

TAREAS COMPLEMENTARIAS

Material BASE:

* La transcripción (en papel y diskette), una vez revisada por un miembro del Equipo de la Unidad de Investigación, será completada respecto de:

a) Corrección de errores y erratas

b) Encriptación de nombres propios y de lugar, generando un listado de equivalencias.

Y de nuevo entregar en diskette la versión corregida con el listado de claves al personal de la Unidad de Investigación.

ANÁLISIS DE CONTENIDO:

Se procederá con los métodos indicados, de entre los siguientes: TAS/C, USALc, Code-A-text, CONC, Útiles de Texto.

¹ *Unidad de Investigación en Psicología Clínica y Psicoterapia* de la Universidad de Salamanca, España. Dirección postal: Avenida de la Merced, 109-131 37005-Salamanca (España) Fax +34-923294607; E-Mail: avil@gugu.usal.es. En este estudio han colaborado, además de los autores, los becarios de investigación de la Universidad de Salamanca Silvia Rodríguez Oliva y Luis Angel Saul Gutiérrez.

REGLAS EJEMPLIFICADAS DEL PROTOCOLO DE TRANSCRIPCIÓN DE SESIONES DE PSICOTERAPIA EN ESPAÑOL

Niveles y contenidos de la transcripción	Regla	Ejemplo	Observaciones
ASPECTOS FORMALES Y ESTRUCTURALES			
1) Encabezamiento de la transcripción	La transcripción debe tener un encabezamiento.	(Caso: 0002 Etiqueta y datos: PUBLICITARIA, mujer, 22 años * María (P) Sesión: 093 Fecha: 20-10-98 Observaciones: transcripción literal de la grabación magnetofónica de la sesión según el protocolo MSGAD).	Esta información debe ser colocada entre paréntesis, en texto normal, sin mayúsculas, ni negrita, ni cursiva.
2) Códigos para indicar quien habla	Cada giro del discurso comienza en una nueva línea y va precedido de un código que indica quién es el hablante. En sesiones de caso único el terapeuta se simboliza T y el paciente P. En caso de que haya más de un paciente o terapeuta los códigos son del tipo X_p , X indica el rol del hablante y n es un dígito opcional (si hay más de un hablante). Si la n es omitida se asume la utilización del dígito 1. Si en el transcurso de la sesión hay alguna interrupción que afecte a la terapia, se transcribirá todo su contenido entre paréntesis, sin asignar códigos a los sujetos. Este formato puede manejarse en monólogos, diálogos, terapia individual, terapia de grupo, con un terapeuta o coterapeutas. Un comentario después del encabezamiento de la transcripción debe ser utilizado para clarificar el rol de los hablantes, por ejemplo: Deben agregarse la fecha, el momento del día, así como las pausas, utilizando formas especiales de codificación.	T: qué le hizo sentir eso? P1: me sentí confundido y enojado. P2: tú nunca dijiste que eso te enojara. (P1 = Hijo, P2= Madre, P3= Padre, T1= Terapeuta, T2= Coterapeuta).	El primer hablante T es el terapeuta y P1 y P2 son dos pacientes. Su dígito implícito es el 1, si hubiera dos terapeutas se anotaría T1 y T2.
3) Fecha y señalización del tiempo.			

Niveles y contenidos de la transcripción	Regla	Ejemplo	Observaciones
4) Fecha de la sesión.	<p>La fecha se introduce bajo la forma "DIA.MES.AÑO". Son dos dígitos para la representación del día del mes, una abreviatura de tres letras mayúsculas para el mes, y cuatro dígitos para representar el año, todas separadas por puntos, sin espacios.</p> <p>La fecha de la sesión debe ser colocada en el comienzo de la transcripción de la sesión, tras el encabezamiento. Si no se conoce la fecha exacta, la información desconocida debe ser sustituida por ceros.</p>	<p>(d: 10.ENE.1986).</p>	<p>La d: indica que es la fecha de la sesión. El ejemplo se lee, diez de enero de 1986.</p> <p>Este formato permite su utilización con sistemas de computación.</p>
5) Hora del día	<p>El comienzo de una sesión es indicado por un código horario.</p>	<p>(t:10:02:15)</p>	<p>La t: indica la hora de comienzo de la sesión, si está disponible. Todos estos códigos se realizan en el siguiente formato "HH:MM:SS" (dos dígitos representando la hora, los minutos, y los segundos, separados por dos puntos). Es preferible utilizar el sistema horario de 24 horas. La hora de la sesión debe ser colocada al comienzo de la transcripción de la sesión en la línea siguiente a la de la fecha. Si no se tiene información exacta, la información desconocida se sustituirá por ceros.</p>
6) Estructura formal del texto.	<p>Se utilizará letra tipo: Times New Roman de tamaño 12; sin negrita, ni cursiva, ni subrayados, con interlineado sencillo. Además la forma de la hoja será siempre vertical y con el texto en formato justificado, sin columnas, ni encabezados, ni pies de página. Tampoco se usarán tabulaciones, ni sangrías.</p>	<p>T: m-hm P: que sin ellas estás arruinado y con ellas eres rico - - entonces claro, cualquier cosa se puede retrasar hasta el viernes, sabes? estamos a martes, son días, o no? - - - - - pues claro - - - T: cómo te sientes?</p>	

Niveles y contenidos de la transcripción	Regla	Ejemplo	Observaciones
7) Mayúsculas.	<p>Todo el texto irá en minúsculas inclusive después de punto. Los nombres propios y lugares geográficos, así como los nombres de empresas y marcas comerciales se transcribirán atendiendo al listado de claves que se entregue a los transcripores, si es así el caso. Según la regla siguiente:</p> <ul style="list-style-type: none"> - El nombre real que se escucha en la cinta se sustituirá por el correspondiente en el listado de claves, anteponiendo un asterisco y utilizando la mayúscula, exclusivamente en la primera letra. - En caso de que algún nombre no aparezca en el listado de claves se transcribirá en mayúsculas, sin asterisco. - Si no se da listado de claves, se pondrán todos los nombres de personas, lugares, empresas y marcas comerciales en mayúsculas sin asterisco. 	<p>Ej: Cinta: Rocio Clave: María Transcripción: *María.</p> <p>Ej: Cinta: Marisa. Clave: el nombre no aparece. Transcripción: M.A.R.I.S.A.</p>	<p>Esto permite el uso de los programas más simples para el recuento de palabras.</p>
8) Simultaneidad	<p>Cuando hay dos hablantes esto puede ser fácilmente manejado agregando un signo de suma + cuando comienza un discurso simultáneo (superposición de discursos), continuando la superposición del primer hablante hasta que dicha superposición finaliza. Eso se continúa por el otro discurso superpuesto correspondiente al segundo hablante que se finalizará por otro signo +. El resto del discurso no simultáneo se transcribe en su orden natural.</p>	<p>En el siguiente ejemplo las palabras "se negó nuevamente" y "sí, usted" fueron verbalizadas a la vez.</p> <p>P: le iba a dar la carta a *Pepe pero él + se negó nuevamente T: sí, usted + me contó eso anteriormente.</p> <p>Segundo ejemplo: las palabras "no lo hacías", y "sí es que no" fueron verbalizadas a la vez y "yo nunca" y "m-hm" también fueron verbalizadas simultáneamente.</p>	<p>La transcripción de discursos sobrepuestos es más fácil si el diálogo puede ser grabado en estéreo con micrófonos diferentes para el paciente y para el terapeuta.</p> <p>Estos signos + van siempre precedidos y seguidos de un único espacio.</p> <p>En los casos en los que se sucedan dos o más simultaneidades, se transcribirán según la regla citada pero cada simultaneidad en un párrafo diferente.</p>
9) Neologismos	<p>Los neologismos deben ser deletreados de la mejor manera posible. Las palabras creadas entlazando otras palabras deben transcribirse unidas con guiones.</p>	<p>T: hacías lo contrario + no lo hacías P: sí es que no + era ni siquiera eso sí es que no sé es que + yo nunca T: m-hm! +</p> <p>P: soy un muerto-vivo.</p>	

Niveles y contenidos de la transcripción

Regla	Ejemplo	Observaciones
<p>Si se va a realizar análisis del texto a través de computadora, las palabras no deben ser divididas al llegar al margen derecho de la hoja utilizando guiones (esto crearía problemas para la utilización de ciertas herramientas de análisis de texto a través de computadoras); la palabra debe ser escrita completamente en la línea siguiente.</p>	<p>P: y le dije a VE P: ayer vi un película de terror T: mm-hm P: era una típica película de te uve.</p>	
<p>No se utilizan abreviaturas, a no ser que el hablante verbalmente deletree una.</p>	<p>T: dónde dijo que estaba? P: en la calle *Sierpes, ese i e erre pe i e ese.</p>	
<p>No se utiliza un punto a continuación de la abreviatura, se debe usar en su lugar un espacio.</p>	<p>P: en 1981 vi los primeros dos-tercios de la película de cuatro bodas y un funeral, era en función de madrugada, a las doce-treinta de la noche y recuerdo que la entrada me costó trescientas pesetas.</p>	
<p>Si el paciente la usa pero no la deletrea, se transcriba la expresión fonética.</p>	<p>P: yo dije 'nunca me ha pasado, sentirme tan error se transcriben tal cual, completamente.</p>	
<p>Si el hablante deletrea expresamente una palabra, debe anotarse tal cual es verbalizada.</p>	<p>P: el gritó 'no disparas hasta que veas sus pupilas'.</p>	<p>En ningún caso se usarán puntos suspensivos dentro de la transcripción. Poner especial énfasis en seguir las reglas gramaticales ortográficas. (ERROR: osea, vamos haber, estube CORRECTO: o sea, vamos a ver, estuve)</p>
<p>Se escriben de la forma más completa posible. Solo los años, se transcriben como números. El resto de referencias numéricas se escriben siempre con letras.</p>	<p>La escritura debe seguir las normas del <i>Diccionario de la Real Academia Española</i>. Cuando se aplican diferentes reglas, es necesario incluirlas a todas en secuencia, todos los signos deben incluirse, si hay entre ellos un punto o un signo de interrogación, deben ir al final.</p>	
<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	
<p>La escritura debe seguir las normas del <i>Diccionario de la Real Academia Española</i>. Cuando se aplican diferentes reglas, es necesario incluirlas a todas en secuencia, todos los signos deben incluirse, si hay entre ellos un punto o un signo de interrogación, deben ir al final.</p>	<p>La escritura debe seguir las normas del <i>Diccionario de la Real Academia Española</i>. Cuando se aplican diferentes reglas, es necesario incluirlas a todas en secuencia, todos los signos deben incluirse, si hay entre ellos un punto o un signo de interrogación, deben ir al final.</p>	
<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	
<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	
<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	
<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	<p>Los errores lingüísticos, así como cualquier otro error se transcriben tal cual, completamente.</p>	

Niveles y contenidos de la transcripción

16) Ambigüedad

Regla

Algunos enunciados pueden resultar ambiguos al escuchar las grabaciones. Es una ventaja la combinación del uso de la computadora con recursos humanos en tanto permite convertir esas declaraciones ambiguas en no ambiguas. Una palabra alternativa que se incluye con la intención de clarificar debe ser colocada detrás de una barra (/). Asimismo, un número colocado inmediatamente después de la barra puede usarse como señal indicadora del significado de una palabra en un diccionario de análisis de contenido. En el caso de pronombres ambiguos, es posible nombrar el antecedente detrás de la barra o incluir varias palabras conectadas por guiones.

Ejemplo

P: nosotros /decimos decidimos que/ *Luisa- Lucas no debe ir.

Observaciones

Esta regla es básicamente utilizada durante las fases de anotación científica y verificación de las transcripciones.

17) Demarcación de segmentos

Pueden llevarse a cabo varias segmentaciones de la transcripción utilizando señalizaciones que indiquen el comienzo de un segmento (s: CÓDIGO) y el final de un segmento (e: CÓDIGO).

Se permite sobreponer distintos tipos de segmentos. Este enfoque puede utilizarse para distintos tipos de segmentaciones. Puede usarse para señalar episodios relacionales (RE), y sueños como en el siguiente ejemplo:

P: (s: RE) cuando le conté a *Jaime que la otra noche soñé (s:SUEÑO) que era una mariposa (e: SUEÑO) él se río (e:RE).

Por ejemplo, un tipo de código puede ser "SUEÑO". Cualquiera que sea la palabra que sustituya la palabra "CÓDIGO" debe ser anotada exactamente igual tanto al comienzo como al final del segmento comentado.

¿QUÉ TRANSCRIBIR?

1) Declaraciones verbales

Todas las palabras verbalizadas como palabras completas o partes de palabras son reproducidas de acuerdo a sus reglas ortográficas. Los dialectos y palabras pronunciadas en otra lengua se transcriben, siempre que sea posible, de acuerdo a sus correspondientes normas ortográficas.

Por ejemplo, si el discurso de una persona, hispanohablante, proveniente del medio rural, suena de la siguiente manera:

P: y después me fui pa laj casa de los pariente. No debe ser transcrita usando normas ortográficas del idioma.

Esta forma de hablar, característica de personas procedentes del medio rural, sería una constante dentro del discurso; si se destacara en algún momento una verbalización claramente diferente, como ser alguna forma idiomática propia del lugar, ésta debe ser consignada particularmente. Para transcribir instancias en las que el hablante deliberadamente usa formas dialécticas señalando énfasis o humor, ver más adelante.

Niveles y contenidos de la transcripción	Regla	Ejemplo	Observaciones
2) Declaraciones para-verbales	<p>Todos los sonidos o secuencia de sonidos utilizados como muletillas (forma de rellenar el discurso), expresiones o sentimientos de duda, confirmación, inseguridad, meditación, deben ser transcritos en español.</p> <p><i>Afirmación:</i> aha, m-hm. <i>Negación:</i> t-t <i>Evación:</i> ehse, mm. <i>Duda:</i> eee, hm, aa. <i>Interrogación:</i> eh, mmm. <i>Humor:</i> ja, jajá. <i>Exclamación:</i> oh, ah, o-o, m-m.</p>	<p>Algunas de estas expresiones cambian en diferentes países o regiones hispano-parlantes (por ejemplo: la [afirmación], pá [exclamación]); Deberán transcribirse en todo caso con su fonética particular.</p>	
3) Declaraciones no verbales	<p>Cualquier ruido o sonido que provenga del P: (estornudo) (tos) bueno (suspiro), creo que hablante es registrado en el texto, en el momento que ocurre, en la forma de simples comentarios entre paréntesis.</p>	<p>(tos) bueno (suspiro), creo que pillé un resfriado (risa).</p>	
4) Ruidos que ocurren en el contexto situacional	<p>Cualquier sonido producido por el ambiente en que transcurre la sesión es indicado a través de simples comentarios.</p>	<p>P: después, cuando yo (suena el teléfono); necesita atender esa llamada?</p>	
5) Pausas	<p>Se puede utilizar un simple guión, con un espacio a ambos lados para indicar una pausa de aproximadamente un segundo. Si se utiliza más de uno, deben separarse por un espacio. Las pausas de más de cinco segundos no se indicarán con guiones, pero deben ser señaladas, dando cuenta del tiempo transcrito entre paréntesis, según se indica en el ejemplo</p>	<p>P: No puedo pensar en - - - nada, (03:35) en una pausa de cinco segundos y otra de 3 minutos y 35 segundos.</p>	<p>No se usarán los guiones de subrayado (los largos). ERROR: _ _ _ _ ; se usarán los específicos (cortos). CORRECTO: - - - .</p>
6) Tiempo transcrito	<p>Es útil en la transcripción anotar el tiempo que ha transcrito.</p>	<p>P: vimos la película (+03:00) después de la cena.</p>	<p>El + indica el comentario sobre el tiempo transcrito desde el comienzo de la sesión. El 03:00 indica que es el comienzo del tercer minuto a continuación del comienzo de la sesión. Si el minuto cambia en la mitad de una palabra, el tiempo debe ser anotado antes de la palabra. El intervalo entre los códigos de tiempo (si es que va ser utilizado) depende de la naturaleza del estudio. Por ejemplo, estos códigos pueden ser utilizados para relacionar el texto con otra serie de datos temporales (por ejemplo, registros fisiológicos). Esto debe ser colocado al principio o al final de los sucesos específicos o a intervalos regulares, como son, a cada minuto, o cada cinco minutos (hay varias ayudas técnicas para dar entrada a estos códigos temporales).</p>

Protocolo de Transcripción de Sesiones de Psicoterapia / Mergenthaler, Gril, Ávila y Vidal			
Niveles y contenidos de la transcripción	Regla	Ejemplo	Observaciones
ASUNTOS QUE REQUIEREN UNA TRANSCRIPCIÓN ESPECIAL.			
1) Palabras incompletas	Aquellas partículas de la palabra, generadas por interrupciones en la verbalización de la misma, incluyendo tartamudeos (repetición de sílabas, palabras incompletas) se indican colocando un guión (-) y un espacio tras ese fragmento de palabra. Una palabra partida es definida como una palabra incompleta que no es repetida de nuevo por el hablante.	P: seguram- yo no pueda ir solo. P: e- e- ella t- t- te pidió que no la llamara más	El tartamudeo se define como: (1) una o más partículas de la palabra, que comparten la letra inicial de la siguiente palabra completa, o (2) una secuencia de más de una partícula de palabra; cada una de estas partículas compartiría la letra inicial, pero no estaría seguida de la palabra completa.
2) Declaraciones indescifrables	Se introduce una barra sencilla (/) en la transcripción delante de cada palabra que no puede ser claramente comprendida pero que puede distinguirse como una palabra separada. Hasta cinco palabras se señalarán con una barra por palabra separadas entre sí de un espacio. En caso de superar las cinco palabras se señalará con (?).	P: yo estaba solo (?) allí, toda la noche (?) / hasta que él///.	Si no es posible determinar el número de palabras en una declaración o ninguna de las posibles palabras, eso debe ser indicado simplemente con el siguiente comentario: P: (incomprensible)
3) Citas.	Si el hablante cita directamente algún trozo de discurso, el texto de cada hablante se coloca entre comillas, utilizando una sola comilla a cada lado. (').	P: yo le pregunté '¿tu lo harás?', y él empezó a gritar 'deja de hablarme de esa manera' y dio un portazo.	No usar los dos puntos del estilo directo.
4) Cambios en la forma de hablar.	Si el hablante cambia su manera usual de hablar y utiliza una voz diferente de la habitual, por ejemplo en un tono irónico o agresivo, esas palabras deben ser entrecorilladas (situadas entre dobles comillas " "). También pueden entrecorillarse de esta manera expresiones en argot o jerga así como transcripciones literales.	P: ella me dice que no diga aquí "vete al diablo" y "toma ya!". Qué se cree que es esto, terapia gramatical?	
5) Comentarios del terapeuta.	Sean al principio o al final de la sesión se transcriben entre paréntesis.		

Niveles y contenidos de la transcripción	Regla	Ejemplo	Observaciones
6) Puntuación	Los indicadores de puntuación son utilizados para ayudar al lector a reconstruir el flujo original del discurso. No se utilizan de acuerdo a las reglas gramaticales tradicionales, porque el discurso habitual está rara vez tan bien ordenado. Quien realiza la transcripción debe usar los indicadores de puntuación para indicar cambios en la forma de hablar, énfasis, entonación y cadencias. Debemos diferenciar las siguientes situaciones:		Si hay dudas, los indicadores de puntuación no deben utilizarse. Deben colocarse al final de la palabra y no dividir la palabra.
6.1) Finalización de un pensamiento.	El punto (.) indica el fin de un pensamiento completo y está usualmente acompañado de una caída en la entonación.		Cuando se de un cambio de la intervención del terapeuta a la intervención del paciente, o de otras personas dentro de la sesión, sólo se utilizará puntuación al final de cada intervención de acuerdo con las reglas comentadas. De tal forma que si se expresa un pensamiento interrumpido esa frase terminará con un punto y coma, si es un pensamiento completo se terminará la frase con punto, etc.
6.2) Pensamiento interrumpido.	El punto y coma (;) indica un pensamiento interrumpido, seguido de otro pensamiento.	P: yo odio la manera en que usted; le comé la boda?	
6.3) Vacilación.	La coma (,) indica una duda, seguida por la continuación del mismo pensamiento y está frecuentemente acompañada por una ligera caída en la entonación.	P: usted, nunca parece, mirarme cuando estoy hablando.	
6.4) Pregunta.	El signo de interrogación (?) indica una pregunta, usualmente se acompaña de una subida en la entonación, o de una clara subida del tono. Debe utilizarse al final de las posibles preguntas si es que hay un aumento de la entonación aún si la frase no tiene gramaticalmente forma de pregunta.	T: a usted no le agrada cuando él hace eso? P: me debería gustar! cuándo él hace eso?	El signo de interrogación se utiliza solamente al final , teniendo en cuenta particularmente, que la entonación de pregunta no se anuncia desde el comienzo de la verbalización.
6.5) Énfasis	El signo de exclamación (!) va a continuación de palabras claramente enfatizadas por el hablante.	P: esto a él no le importa! pero para mí! es distinto.	Es importante tener en cuenta que el signo de exclamación, en las transcripciones, es utilizado solamente con la función de enfatizar y no indica la finalización de una frase gramatical. De igual forma que el signo de interrogación se utilizará sólo al final
6.6) Pronunciación prolongada	Los dos puntos no son utilizados en su forma gramatical tradicional pero se utilizan para indicar cuando se extiende la pronunciación de una palabra, sílaba o letra.	P: bueno: a mí realmente s-: siempre me gustó eso. (ERROR: buenooo a mí realmente sssiempre me gustó eso).	

Niveles y contenidos de la transcripción NORMAS ADICIONALES	Regla	Ejemplo	Observaciones
1) - Línea en blanco al inicio del texto.	El texto debe comenzar con una línea en blanco (sólo una línea).	Línea en blanco al inicio del texto P: (risas) puedo fumar, no?, como te veo a ti (el T. tiene un cigarrillo encendido; gesto de asentimiento de T.). es que no sé muy bien por dónde empezar en cuestión de tiempo atrás, yo creo que, toda la inseguridad que yo tengo con respecto a mi pareja, no con respecto al resto de la gente, es porque desde pequeña digamos, siempre he salido con chicos, no salir de una forma seria, en un pueblo donde digamos, no sé si es una impresión mía pero siempre parecía que yo destacaba más que las demás. T: decías que había pasado antes con *cristina. Espacio en blanco tras T: y decías	
2) Espacio en blanco tras los identificativos de paciente y terapeuta.	Los identificativos de paciente y terapeuta han de ser correctos (P; y T:) y que exista un espacio en blanco tras los dos puntos.	T: decías que había pasado antes con *cristina. Línea en blanco entre párrafo y párrafo. P: sí, hace muchos años - - - o sea, *cristina - yo la conozco de toda la vida porque sus padres y los míos son familia, y nos hemos criado en un pueblo pequeño, bueno, en *valencia; pero todos los veranos íbamos allí.	
3) Espacio en blanco entre párrafo y párrafo.	Entre párrafo y párrafo ha de haber sólo una línea en blanco.	T: Son unos sentimientos muy fuertes los que te despierta. (fin de la sesión).	Línea en blanco con el cursor parpadeante bajo la misma al final de cada documento de texto. T: qué pueblo?. Espacio en blanco tras la última palabra del párrafo.
4) Línea en blanco al final de cada documento.	Al final de cada documento de texto debe haber sólo una línea en blanco con el cursor parpadeante bajo la misma.	Al final de cada párrafo se tiene que dar a la barra espaciadora una vez con lo cual quedaría un espacio en blanco después de la última letra o punto del párrafo)	
5) Todo final de párrafo ha de terminar con un espacio en blanco tras la última palabra	Al final de cada párrafo se tiene que dar a la barra espaciadora una vez con lo cual quedaría un espacio en blanco después de la última letra o punto del párrafo)		

- ¹ Este estudio tiene lugar con ayuda del Research Advisory Board de la International Psychoanalytical Association (1999). Han colaborado además los becarios de investigación de la Universidad de Salamanca Silvia Rodríguez Oliva y Luis Angel Saúl Gutiérrez.
 - ² Director de la *Sección de Informática en la Psicoterapia*, Universidad de Ulm - Clínica Universitaria, Alemania.
 - ³ Investigadora y Profesora Adjunta del Departamento de Psicología Médica en la Universidad de la República-Uruguay.
 - ⁴ Catedrático y Director de la *Unidad de Investigación en Psicología Clínica y Psicoterapia* de la Universidad de Salamanca, España. Dirección postal: Avenida de la Merced, 109-131 37005-Salamanca (España) Fax +34-923294607; E-Mail: avil@gugu.usal.es
 - ⁵ Becaria de investigación de la *Unidad de Investigación en Psicología Clínica y Psicoterapia* de la Universidad de Salamanca, España
-

Es cada vez más clara la importancia de realizar transcripciones de sesiones y de otros materiales clínicos de acuerdo a estándares para poder utilizar estos materiales en estudios en colaboración y trans-nacionales en el ámbito de la investigación de procesos en psicoterapia. Desarrollando aportaciones precedentes (Mergenthaler y Stinson, 1992; Mergenthaler y Gril, 1996) se presentan aquí nuevas orientaciones más completas y detalladas para realizar la transcripción del discurso en el idioma español. Estas orientaciones pueden ser utilizadas no sólo con estos fines, sino también para investigaciones en otras áreas (educación, lingüística, antropología, entre otras). El protocolo de transcripción que se propone es utilizable para varias tareas: a) transcripción de materiales y su valoración por jueces; b) Análisis de textos por computadora, facilitando la obtención de indicadores cuantitativos en el análisis de contenido y la adecuada preparación de los textos para el posterior análisis del contenido cualitativo. Se efectúa una breve presentación de las principales herramientas de análisis computacional de textos.

Palabras clave: Investigación en Psicoterapia, Reglas de Transcripción, Análisis de Textos

Referencias bibliográficas:

- BAKER, J. C. (1988). Pace: A test of authorship based on the rate at which new words enter an author's text. *Literary and Linguistic Computing*, 3, 36-39.
- BERELSON, B. (1952). *Content analysis in communication research*. Glencoe: Free Press.
- CARTWRIGHT, A. (1996). An Introduction to Code-A-Text. A software system to code and analyse psychotherapy texts and recordings. Canterbury: Centre for Study of Psychotherapy-University of Kent
- CHOTLOS, J. W. (1944). *Studies in language behavior*. Psychology Monography (pp. 56)
- DAHL, H. (1979). *Word frequencies of spoken American English*. Essex, CT: Verbatim.
- EDELSKY, E. (1981). Who's got the floor? *Language in Society*, 10, 383-421).
- EDWARDS, J. A. & LAMPERT, M. D. (Eds.). (1991). *Transcription and coding methods for language research*. Hillsdale, NJ: Erlbaum.

- HERDAN, G (1960). *Type-token mathematics*. The Hague, Mouton & Co.
- HOLSTI, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading MA Addison-Wesley
- HOWE, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die-hard. *Educational Researcher*, 17, 10-16.
- KÄCHELE, H. (1983). Verbal activity level of therapists in initial interviews and long-term psychoanalysis. In W. Minsel & W. Herff (Eds.). *Methodology in psychotherapy research* (pp. 125-129) Frankfurt, Germany: Lang.
- KRACAUER, S. (1952). The challenge of qualitative content analysis. *Public Opinion Quarterly*, 16, 631-642.
- KRACAUER, S. (1972). Für eine qualitative Inhaltsanalyse. *Ästhetik und Kommunikation*, 3
- LORENZ, M. & COBB, S. (1954). Language patterns in psychotic and psychoneurotic subjects. A. M. A. *Archives of Neurology and Psychiatry*, 72, 665-673).
- LUBORSKY, L. & SPENCE, D. P. (1978). Quantitative research on psychoanalytic therapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy research and behavior change: An empirical analysis* (pp. 331-427). New York: Wiley.
- MAHL, G. F. (1956). Disturbances and silence in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology*, 53, 1-15.
- MARTINDALE, C. (1975). *Romantic progression: The psychology of literary history*. Washington, Hemisphere.
- MERGENTHALER, E. (1985). Textbank systems: *Computer science applied in the field of psychoanalysis*. New York: Springer-Verlag.
- MERGENTHALER, E. (1990, June). *Parts of speech: A measure of therapeutic alliance?* Trabajo presentado en la Sociedad de Investigación en Psicoterapia, Wintergreen, VA.
- MERGENTHALER, E. (1991) *Zur Reliabilität von Transkriptionsregeln*. Manuscrito sin publicar.
- MERGENTHALER, E. & STINSON C. H. (1992) *Psychotherapy Transcription Standards*. *Psychotherapy Research*, 2 (2), 125-142.
- MERGENTHALER, E. (1993). *TAS/C User Manual*. Ulm: Ulmer Textbank.
- MERGENTHALER, E. & GRIL, S. (1996). *Descripción de las reglas para la transcripción de sesiones de psicoterapia* (5)2, 163-176.
- OXMAN, T. E., ROSENBERG, S. D., SCHNURR, P. P., & TUCKER, G. J. (1985). *Linguistic dimensions of affect and thought in somatization disorder*. *American Journal of Psychiatry*, 142, 1150-1155.
- OXMAN, T. E., ROSENBERG, S. D., SCHNURR, P. P., TUCKER, G. J., & GALA, G. (1988). The language of altered states. *Journal of Nervous and Mental Disease*, 176, 401-408.
- ROSENBERG, S. D., SCHNURR, P. P., & OXMAN, T. E. (1990). Content analysis: *A comparison of manual and computerized systems*. *Journal of Personality Assessment*, 54, 298-310.
- SHANNON, C. E. (1948). *A mathematical theory of communication*. Bell Syst Techn J 27: 379-423.
- SPENCE, D. P. (1968). *The processing of meaning in psychotherapy: Some links with psycholinguistics and information theory*. *Behavioral Science* 13:349-361
- SPENCE, D. P., OWENS K. C. (1990). *Lexical Co-Occurrence and Association Strength*. *Journal of Psycholinguistic Research* 19(5):317-330
- SOTELO, J. A. & ÁVILA, A. (2000). *Utilización de Software para el Análisis del lenguaje y la Cognición – USAI* ver 1. 1. Madrid. CD-ROM demo y completo©.
- STILES, W. B. & SULTAN, F. E. (1979) Verbal Response modes and psychotherapeutic technique. *Journal of Consulting and Clinical Psychology*, 47, 611-613.
- STINSON, C. H. (1991). *Transcript verification analysis*. Manuscrito sin publicar.
- THOMPSON, J. (1990). CONC. Dallas: International Academic Bookstore- Summer Institute of Linguistics