

EL ANÁLISIS DE TEXTOS EN LA INVESTIGACIÓN

Fragmento del libro *Análisis del Discurso e Investigación Social* (Padrón, J., Publicaciones del Decanato de Postgrado de la USR, Caracas: 1996, pp. 129-136)

En este tercer módulo intentaremos aplicar principios y técnicas del Análisis del Discurso a la sistematización de la información generada por un texto bajo estudio. Nuestro supuesto elemental será que:

- muchos de los datos que el investigador social se ve obligado a procesar provienen de *textos* (derivados a su vez de actos de habla);
- para obtener datos relevantes de un texto es necesario sistematizar el conjunto de la información contenida en dicho texto;
- para sistematizar el conjunto de la información contenida en un texto hacen falta ciertos principios y técnicas de análisis;
- el Análisis del Discurso puede proveer esos principios y técnicas.

En el mundo de la investigación es ya un prototipo la relación primaria entre el investigador y el “corpus”, entendido éste como el conjunto ya seleccionado de objetos que contienen inmediatamente los datos o de los cuales hay que extraer los datos. Lo que algunos llaman “unidades de análisis”, en la tradición de cierto tipo de estudios es aquí llamado “corpus”, es decir, un inventario de manifestaciones, realizaciones o ejecuciones observables que, al ser analizadas, manipuladas o procesadas, generan información acerca de alguna incógnita. Por ejemplo, si un investigador se plantea alguna pregunta relacionada con la influencia de la publicidad televisiva sobre los niños, es posible que reúna una cierta cantidad de comerciales de TV, que será entonces su *corpus* de investigación. Si, en cambio, se pregunta sobre la reacción de determinados tipos de personas ante conductas agresivas, es posible que filme eventos en el tránsito, en los tribunales, en oficinas de reclamo, etc., filmaciones que constituirán su *corpus* de trabajo¹ (en los enfoques empírico-inductivos los corpus son utilizados para construir explicaciones por vía de generalización progresiva; en cambio, en los enfoques deductivos son utilizados para contrastar explicaciones previamente formuladas o como mecanismos de “control empírico” durante el proceso de derivaciones).

Esta relación primaria entre investigador y corpus se ve crecida o mediatizada por otros componentes:

- i) en primer lugar, está la referencia del problema y el objetivo de estudio;
- ii) en segundo lugar, hay un cuadro de teorías de insumo (“ α -Teorías” o teorías de entrada) que se vincula al problema y al objetivo;
- iii) en tercer lugar, está un sistema de categorías de investigación o de criterios de segmentación, clasificación y agrupación de hechos bajo estudio, sistema que es extraído del anterior cuadro de teorías de insumo;
- iv) en cuarto lugar, hay otro cuadro de *utilidades* teóricas o de teorías de servicio que no se vinculan directamente al problema ni al objetivo de trabajo, sino a grandes clases de problemas e intereses y que proporcionan herramientas de trabajo o tratamientos generales;

¹ Los elementos de un corpus deben ser *representativos* de la relación estudiada, pero no necesariamente en un sentido de *muestra estadística*. El carácter representativo de un elemento puede tener un sentido *lógico*, cuya fundamentación o justificación radica en lo siguiente: si formulamos una clase ‘A’ de hechos mediante una lista cerrada de propiedades “bien-definidas” (que puedan ser reconocidas por una máquina) y dado un hecho cualquiera x, es posible decidir si x pertenece a A. En caso afirmativo, todo lo que encontremos en x y que esté dentro del rango de la lista de propiedades, ipso facto puede ser atribuido a la clase A (o sea, puede ser generalizado sin márgenes de error). En la terminología especializada, la clase a la que pertenece un elemento es llamada “Signo-Tipo” y cualquier elemento representativo de esa clase es llamado “Signo-Ejemplar”. Esa terminología se deriva de Peirce, quien distinguió entre “Type” (la clase) y “Token” (el elemento representativo de esa clase). Desde una óptica racional-deductiva, el sentido *lógico* de la representatividad es mucho más confiable que el sentido *probabilista* (el de las muestras al “azar”, por ejemplo).

v) en quinto lugar, están las técnicas y herramientas proporcionadas por el anterior cuadro de teorías de servicio, las cuales permiten operar sobre el corpus, extraer datos y procesarlos². Todo esta sistema de relaciones investigativas puede ser visualizado en el diagrama 1.

Ahora bien, sobre la base del supuesto anterior y de su implícita relación entre investigador y corpus, pasaremos ahora a considerar dos elementos básicos: el Análisis del Discurso, en cuanto especial teoría de servicio, por un lado, y los corpus típicos de investigación social centrada sobre elementos textuales, por otro lado.

En cuanto al primer elemento -el Análisis del Discurso- damos por descontado su dominio teórico, que ya fue expuesto en los dos módulos temáticos anteriores y que, por tanto, no volveremos a tocar por ahora (a menos que se requieran referencias circunstanciales). En su versión técnica, utilizaremos un cierto conjunto muy limitado de herramientas en cuatro instancias sucesivas: abordaremos, primero, una instancia de análisis **pragmático**; segundo, una instancia de análisis **semántico**; tercero, el análisis **sintáctico** y, finalmente, de modo bastante superficial, la instancia del análisis **físico**.

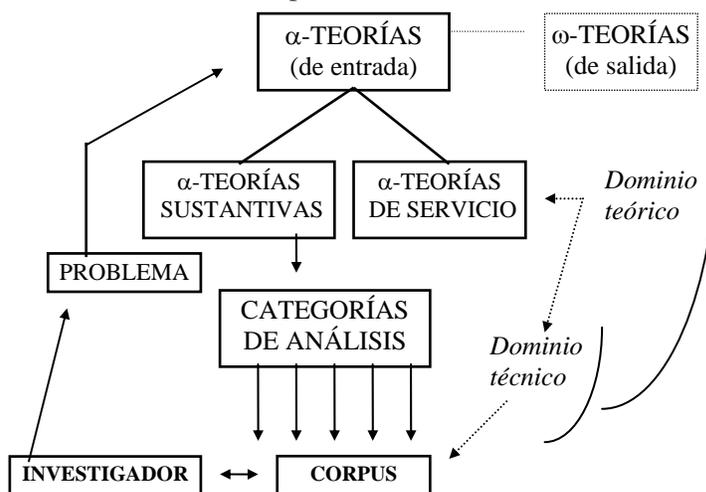


Diagrama 1: relaciones entre investigador y corpus

En cuanto al segundo elemento -los corpus asociados al Análisis del Discurso en una investigación social- consideraremos, dentro de cada una de las instancias de análisis antes mencionadas, cuatro clases de corpus: textos autobiográficos, entrevistas, textos oficiales o institucionales en general y mensajes publicitarios. Precisamente, dividiremos este módulo en secciones correspondientes a esas cuatro instancias.

² Nótese que las teorías de servicio (o los *utilitarios* teóricos) también forman parte de las α -Teorías o teorías de entrada, pero la diferencia está en que, mientras éstas mantienen una relación sustantiva, temática o semántica con el problema (que es el caso del componente ii, arriba), aquéllas mantienen sólo una relación formal, operativa o sintáctica (que es el caso del componente iv, arriba). La teoría del Conflicto, por ejemplo, es de alcance sustantivo o temático con respecto a los problemas en torno a negociación o a relaciones interpersonales. Pero la teoría de probabilidades, por ejemplo, aplicada a esos mismos problemas, tiene un alcance formal u operativo, igual que la teoría de grupos o de caos (y las teorías lógico-matemáticas) e igual que la teoría de sistemas o del Discurso. Evidentemente, dentro de su propia especialidad o disciplina, cada teoría de servicio pasa a ser ya una teoría sustantiva.

Por otra parte, estas teorías de servicio pueden ser incorporadas a una investigación de dos modos: a modo de dominio *técnico* (es decir, a nivel de simples herramientas ciegas) y a modo de dominio *teórico-técnico* (o sea, a nivel de herramientas justificadas teórica y conceptualmente). El caso de la estadística es un excelente ejemplo: al analizar hechos en términos muestrales, podemos limitarnos sólo a calcular medidas de tendencia central y de dispersión, lo cual representa el dominio *técnico* de una teoría de servicio. Pero debajo de esos cálculos hay teorías de distribución muestral, cuya incorporación a la investigación sería el caso de un dominio *teórico-técnico*. El componente iv representa el dominio *teórico* de un utilitario, mientras que el componente v representa su dominio *técnico*.

La manera en que se interrelacionarán estos dos elementos en las cuatro secciones de este módulo queda representada en la tabla 1, donde además aparece, en cada casilla, el rasgo definitorio básico de cada clase de corpus según cada instancia de análisis:

	Sección 1: Textos Autobiográficos	Sección 2: Textos institucionales	Sección 3: Entrevistas	Sección 4: Mensajes Publicitarios
Análisis pragmático	Macroacto: informar sobre las propias experiencias	Macroacto: Definir lineamientos y concepciones institucionales	Macroacto: atender solicitud de información	Macroacto: persuadir
Análisis semántico	Macrosigni-ficado: experiencias familiares, laborales, profesiona-les, educativas, artísticas, etc.	Macrosignifi-cado: acciones socializadas (macro/micro-grupales, organizacionales...)	Macrosigni-ficado: el definido por la solicitud del entrevistador (lo que sabe el entrevistado)	Macrosig-nificado: las virtudes o conveniencias de un bien o servicio
Análisis sintáctico	Lenguaje: verbal escrito	Lenguaje: verbal escrito	Lenguaje: verbal oral	Lenguaje: múltiple-integrado (icónico, gráfico-dia-gramático, verbal escrito, verbal oral, musical)
Análisis físico	Medio original: impreso	Medio original: impreso	Medio original: cara-a-cara	Medio original: impreso, video, foto, etc.

Tabla 1: Tipos de corpus analizados en cada sección

Antes de cerrar esta introducción, consideremos algunas generalidades técnicas sobre cada una de las cuatro instancias de análisis.

El Análisis Pragmático

Para cualquier texto en general, el análisis pragmático busca datos luego de definir, como referencia de examen, la situación socio-espacio-temporal en la que se produjo el texto, contextualizada dentro del respectivo marco situacional (el comercio, el entretenimiento, la recreación, la educación, la academia, la política, etc.) e identificada bajo un Macroacto. Estas definiciones previas permiten fijar un patrón interpretativo y evaluativo para la obtención de los datos.

Partiendo de eso, es conveniente diseñar la estructura de macro/meso/micro-actos del texto, con el objeto de clarificar lo que ‘va haciendo’ el autor del texto a medida que desarrolla su discurso y de evaluar la eficiencia del texto. Pero, más que eso, ese diseño de la estructura pragmática permite visualizar las intenciones del autor, especialmente si sus actos son *bivalentes* (i. e., si existen acciones encubiertas).

Evidentemente, todo depende de las categorías de análisis de la investigación o de lo que se esté buscando. Pero, en general, el análisis pragmático provee información sobre lo siguiente:

- las relaciones del autor con respecto a la situación socio-espacio-temporal y al marco situacional correspondiente: su grado de adaptación, su dominio de las convenciones del caso, su equilibrio socio-contextual, etc.
- su eficiencia como actor-hablante: sus mecanismos de enlace, la cohesión y coherencia de sus meso/micro-actos, el esfuerzo invertido en la comunicación, etc.
- su sistema de valores, motivaciones, creencias, etc.: aunque buena parte de los valores se identifican en la instancia semántica, en esta instancia pragmática suele haber algunos ‘actos’ especiales que reflejan los valores y los sistemas normativos del autor. Tal es el caso de actos como prohibir algo, persuadir de algo, ensalzar algo, manifestar temor de algo, defender algo, etc., donde el nombre del acto (ubicado en el verbo) indica valores y normas, mientras que el ‘algo’ remite a un posterior análisis semántico que dirá, concretamente, en qué consisten tales valores y normas.

El Análisis Semántico

Básicamente, el objetivo del análisis semántico consiste en sistematizar los mapas representacionales que el hablante intenta transmitir. La clave aquí consiste en diseñar el sistema de macro/meso/micro-significados sobre la base de la estructura pragmática, bajo el supuesto teórico de que a cada x-acto corresponde un x-significado. Con el apareamiento del esquema semántico al esquema

pragmático, podemos ya especificar y comprender mejor los actos textuales, ya que, además de la identificación de una acción, disponemos ahora de los distintos mapas representacionales asociados.

Otra estrategia de análisis, que puede ser refinada con técnicas estadísticas, consiste en, primero, aislar los nombres, los adjetivos, los verbos y los adverbios presentes en el texto (que son los elementos textuales semánticamente cargados o semánticamente eficaces), junto a la cantidad de veces que se repite. Segundo, agruparlos en categorías o ejes de significación formulados por el propio investigador mediante ensayo y error. El tipo de ejes formados por las palabras y sus repeticiones indicará entonces los significados principales y secundarios del texto (los psicoanalistas poseen estrategias específicas que profundizan mucho más en estos ejes semánticos). Más adelante mostraremos algunos ejemplos de aplicación que ilustrarán mejor esta técnica.

Al margen de datos específicos que interesen al investigador, el análisis semántico permite, en general:

- determinar los valores, normas y creencias del autor,
- establecer la adecuación entre lo que ‘hace’ y lo que ‘informa’,
- precisar las relaciones entre los mapas representacionales del autor y los que corresponden al marco situacional,
- determinar la coherencia, consistencia y relevancia de su información (trivialidades, redundancias, omisiones, etc.).

El Análisis Sintáctico

Para la investigación social (no lingüística) el objetivo del análisis en esta instancia es profundizar en los hallazgos de los dos componentes anteriores a través de las estructuras formales del lenguaje, tales como el orden de las palabras, las formas lógicas (pregunta, admiración, afirmación, etc.), las relaciones y conexiones entre frase y frase, entre párrafo y párrafo, entre sección y sección, entre encabezados y textos, entre sublenguajes de uno y otro tipo, etc.

Las estructuras con “tú” y “Ud.”, por ejemplo, remiten a actitudes frente al interlocutor, lo mismo que los imperativos escuetos con respecto a los imperativos disfrazados de pregunta. Los desórdenes sintácticos (mal uso de los signos de puntuación, discordancias o inconexiones verbales, redacciones forzadas...) remiten a correspondientes estructuras de pensamiento y a orden lógico, etc.

El Análisis Físico

El objetivo de esta instancia es, igual que en el caso anterior, ampliar la información obtenida en las dos primeras instancias mediante la observación de las características específicas del medio de transmisión/recepción. En investigación social debe tomarse en cuenta que unas veces se trabaja con medios diferentes al medio original en que se produjo el texto (como ocurre en las grabaciones y en las transcripciones), lo cual hace variar completamente el análisis³. Las principales estructuras del medio, a las que debe atender el analista, son la calidad y legibilidad del material, las tecnologías empleadas, la distancia espacio-temporal que el medio fija entre los interlocutores, etc.

³ Otras veces ocurre que el entrevistado, por ejemplo, al advertir que su voz está siendo grabada o que está siendo filmado, cambia notablemente su producción textual. Es competencia del analista discriminar esta variable.