

LA LEXICOMETRÍA COMO MÉTODO DE LOCALIZACIÓN DE RASGOS IDEOLÓGICOS *

Hablamos de texto o enunciado ideológico cuanto éste es expresivo de las relaciones sociales reales o deseadas. El término ideológico es definido como «conjunto coherente de representaciones, valores y creencias que reconstruyen en una dimensión imaginaria dichas relaciones sociales»¹. Estas relaciones sociales se presentan siempre en el marco de una ideología dominante, que intenta presentarse como si cubriera la totalidad del campo ideológico, y una o varias ideologías dominadas que aparecen, en el campo del discurso político, como subconjuntos relativamente autónomos en relación a la ideología dominante.

«En toda forma de sociedad es una producción determinada y las relaciones engendradas por ella las que asignan a las demás producciones y relaciones su rasgo y su importancia»². La ideología se presenta, dentro de una formación social, constituida por diversos modos de producción o por diversas formas procedentes de ellos y reestructuradas en función de la dominación de uno de estos modos, y, por tanto, parece evidente que cuando hablamos de ideología no nos referimos a un ente sin sentido sino, en términos de Althusser, a «la manera según la cual los hombres viven sus relaciones con sus propias condiciones de existencia» y, en consecuencia, las ideologías

* Esta comunicación fue presentada al Simposio de Lingüística organizado por la S. E. L. y celebrado en Sevilla en diciembre de 1977.

¹ Garcés, J., *Chile: El camino político hacia el socialismo*, Barcelona, 1972, págs. 80-81.

² Marx, K., *Contribución a la crítica de la economía política*, Madrid, 1970.

serán «formations complexes de montages de notions, de représentations, d'images d'une part, et de montages de comportements, attitudes-geste d'autre part, l'ensemble fonctionnant comme des normes pratiques qui gouvernent l'attitude et la prise de position concrète des hommes à l'égard des objets réels de leur existence sociale et individuelle et de leur histoire»^{3 4}. Con esta definición Althusser nos lleva ya a la consideración de dos tipos de ideología, una ideología teórica y una ideología práctica, o lo que sería igual, en términos de Hérbert, a dos formas de ideología, una forma empírica, referida a las relaciones sociales de producción, «reorganizadora de los elementos del proceso de trabajo» y reflejada en los hechos; y una forma especulativa, referida a las relaciones sociales de producción, cuya función es «dar a conocer a los agentes de la producción su lugar en el seno de la producción». Estas dos formas de ideología llevarían a dos funciones distintas: una función «semántico-metafórica», dominante en la forma empírica, y una función «sintáctica», propia de la forma especulativa, que permitiría a los sujetos «identificarse con las estructuras políticas e ideológicas»⁵.

Los cuatro puntos caracterizadores de una ideología serían, siguiendo a Hérbert, los siguientes:

Las ideologías son históricamente necesarias y constituyen el terreno en que los hombres se mueven y se conciencian de su posición.

Las ideologías tienen una función específica en una formación social.

Las ideologías son inconscientes de su sistematicidad.

Las ideologías tienen una existencia material porque informan prácticas e instituciones que constituyen los aparatos ideológicos del estado.

Introduciendo el término «formación ideológica» pretendemos caracterizar un elemento capaz de intervenir como fuerza confrontada a

³ Althusser, L., *Pour Marx*, París, 1965.

⁴ Althusser, L., «Cours pour scientifiques». Cit. por Robin, F., *Histoire et linguistique*, París, 1973.

⁵ Hérbert, Th., «Remarques pour une théorie générale des ideologies», *Cahiers pour l'analyse*, núm. 9, 1968.

otras en una formación social y en una coyuntura determinada. Las formaciones ideológicas son conjuntos complejos de actitudes y de representaciones relacionadas más o menos directamente con posiciones de clase. Dentro de este conjunto situaremos las formaciones discursivas, entendidas como conjunto de manifestaciones lingüísticas gobernadas por una formación ideológica. En ningún momento afirmamos que las manifestaciones lingüísticas se reducen a ideologías, ni que las ideologías se superponen a las manifestaciones lingüísticas; lo único que pretendemos decir es que cada formación discursiva no es sino uno de los conjuntos que componen las formaciones ideológicas; es, en suma, un subconjunto de un conjunto ideológico determinado y que, por tanto, se manifiesta en función de las condiciones de producción, de las instituciones que las sostienen y de las reglas que constituyen el propio discurso. En consecuencia, un análisis de un texto ideológico deberá incluir, por un lado, la relación entre las formaciones discursivas y las posiciones que sostienen en el campo social e ideológico, y, por otro, la selección y combinación de las unidades lingüísticas en relación a su propio enunciado.

Si aceptamos los principios expuestos hasta este momento, es lógico pensar que es posible localizar en un discurso tanto los rasgos ideológicos como los sistemas estereotipados de representación de las formaciones ideológicas. Hemos de tener en cuenta, no obstante, que una ideología nunca se manifiesta en un texto del tipo que tratamos de un modo patente, sino que hay que descubrirla «leyendo entre líneas».

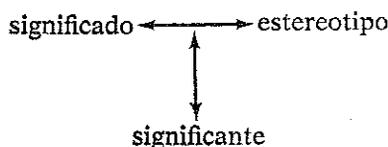
E. Veron considera que una ideología a nivel lingüístico no es más que uno de los muchos niveles de organización de los mensajes desde la base de sus propiedades semánticas⁶. El nivel de significación ideológico, presente en cualquier tipo de mensaje, únicamente puede ser localizado mediante el análisis de la organización de un discurso, aclarando previamente las características de selección y combinación de las unidades que lo integran. La ideología por tanto no es comunicable, sino metacomunicable, porque no es otra cosa que «un

⁶ Veron, E., «Ideología y comunicación de masas: La semantización de la violencia política», *Lenguaje y comunicación social*, Buenos Aires, 1971.

sistema de reglas semánticas expresivas de un determinado nivel de organización de un mensaje»⁷.

La ocultación de la ideología en un discurso nos hace patente que un discurso ideológico se caracteriza por asumir un tipo especial de función conativa, llamado por Veron «normativa», función determinada por su intento de convencer a un receptor, no siguiendo una estructura manifiesta, sino reforzando de una manera oculta un determinado universo de pautas económico-sociales. La primera función, la que se manifiesta, es aparentemente referencial y descriptiva de la realidad; la segunda, implícita y oculta, es normativa, es decir, controladora de la conducta humana a través de la propia organización lingüística de un discurso⁸.

Para definir y caracterizar el objeto discursivo-ideológico, a fin de no confundirlo con la significación, tomamos como punto de partida la distinción establecida por A. Schaff dentro del plano del contenido⁹. Si un signo lingüístico era para Saussure la relación significado-significante, para Schaff es la relación significante/significado-estereotipo



El concepto es el reflejo generalizado de la realidad en el conocimiento humano; el estereotipo, el reflejo particular y subjetivo de la realidad que incluye, en primerísimo lugar, los elementos emocionales, evaluativos y volitivos de dicha realidad. Ambos términos designan un modo de ver la realidad, pero el proceso de conocimiento del concepto marca una tendencia «objetivamente» descriptiva, mientras que el proceso que lleva al estereotipo se fundaría en el conoci-

⁷ Veron, E., *op. cit.*, pág. 141.

⁸ En un sentido parecido se manifiesta Morris cuando combinando las variables «uso» y «modo dominante de significar» establece los distintos tipos de signo y caracteriza el signo político como «valorador y prescriptor» porque «prescribe acciones para la organización institucional de la sociedad como un todo, a fin de encontrar aprobación por haber procedido de tal manera». Morris, Ch., *Signs, language and behavior*, Nueva York, 1946 (traducción castellana en Losada, 1962).

⁹ Schaff, A., *Lenguaje y acción humana*, Barcelona, 1971.

miento evaluativo; en palabras de Schaff, «el concepto es una categoría de pensamiento lógico; el estereotipo, de conocimiento pragmático»¹⁰. Concepto y estereotipo configuran la ideología que, a su vez, condicionada por concepto y estereotipo, desencadena un comportamiento social, que, exidentemente, cambiará en función del estereotipo que lo origine. Podemos afirmar en consecuencia que grupo social y estereotipo se requieren y caracterizan mutuamente. Todo grupo ideológico posee un estereotipo, y a éste le es absolutamente necesario un grupo social para marcar opiniones preconcebidas de clases de individuos, grupos u objetos, opiniones que no son, más que el resultado de nuevas valoraciones de la misma realidad.

Para llevar a cabo estudios lingüísticos seguidores de pautas como las que acabamos de presentar, fue fundamental la distinción de L. Guespin entre «enunciado» y «discurso»¹¹. El enunciado sería la serie de frases emitidas entre dos blancos semánticos; el discurso, el enunciado considerado desde el punto de vista del mecanismo discursivo que lo condiciona. Esta distinción lleva a establecer las diferencias entre lingüística del enunciado y lingüística del discurso. La lingüística del enunciado tendrá en cuenta el plano lingüístico; la del discurso deberá estudiar las condiciones de producción de un enunciado y, en consecuencia, el proceso de enunciación; y, además, tendrá que poner en relación el plano vertical de las condiciones de producción de un discurso y el plano horizontal característicos de las unidades lingüísticas seleccionadas y combinadas por un sujeto emisor.

En ningún momento queremos dar a entender que las condiciones de producción de un texto «son el contexto extralingüístico», sino que este contexto está contenido en un discurso, lo constituye y lo caracteriza, y, en consecuencia, es localizable mediante un análisis lingüístico del mismo. En otras palabras, llevar a cabo un estudio del discurso significa reintroducir en los estudios lingüísticos las variables extralingüísticas, la relación emisor-receptor, el marco institucional en que ambos se mueven y el contexto en que se produce la emisión; y ello no puede conseguirse partiendo únicamente de la dico-

¹⁰ Schaff, A., *op. cit.*, pág. 22.

¹¹ Guespin, L., «Problématique des travaux sur le discours politique», *Langages*, núm. 23, 1971. Véase en un sentido parecido, Dubois, J., y Sumpf, J., «Problèmes de l'analyse du discours», *Langages*, núm. 13, 1969; Guespin, L., «Types de discours, ou fonctionnements discursifs?», *Langages*, núm. 41, 1976.

tomía saussuriana «langue» «parole», ni de la dicotomía inicial de Chomsky «competence» «performance», porque un estudio del discurso excede a la simple lógica denotativa¹².

No se trata tampoco de reducir la lengua a una superestructura, ni de negar el análisis intralingüístico, ni de reducir la lingüística del discurso a una sociología, sino de constituir el nivel discursivo como nuevo objeto de estudio en el campo de la lingüística; y ello comporta una interpretación del discurso que necesariamente tendrá de incluir su función, su eficacia y sus procesos de inserción en un formación social.

El método lexicométrico¹³ —estudio del uso léxico partiendo de su cuantificación— puede ser una introducción en el análisis del discurso, a fin de llegar a localizar rasgos ideológicos del mismo sin perder de vista las condiciones de producción del propio discurso.

Cuantificar un texto significa hacer un recuento de las unidades que lo integran a partir de los siguientes pasos:

1. Especificación detallada de la norma lexicológica adoptada.
2. Preparación del corpus que pretendemos estudiar.
3. Elección de un método de tratamiento del corpus.
4. Aplicación de tests estadísticos.
5. Presentación de resultados cuantificados.
6. Interpretación de los resultados.

1. La norma lexicológica no es más que el conjunto de criterios que garantizan la constante de tratamiento de un texto. Escoger una norma óptima significa elegir la más simple, la más fácilmente interpretable, la más autónoma con relación a recurrir al texto y la menos arbitraria, en el sentido de no depender del

¹² En una línea paralela podríamos citar la obra de Austin, J. L., *How to do things with words*, Oxford, 1962; Ducrot, O., «Présupposés et sous-entendus», *Langue Française*, núm. 4, 1969 y Searle, J. R., *Speech Acts*, Cambridge, 1969.

¹³ Preferimos, siguiendo a ERA-56 (Equipe de Recherche Associée), el término lexicometría —y no el de Estadística Léxica— porque, como muy bien justifican, es más preciso en cuanto al objeto y más amplio en cuanto al método. Véase para ello ERA-56 «Analyse lexicométrique et formalisation», *Les applications de l'informatique aux textes philosophiques*, París, 1970; Gefroy, A., «L'analyse lexicométrique des textes», texto policopiado del Centro de Saint Cloud, 1971.

criterio particular y subjetivo de cada investigador. El punto primordial para el establecimiento de la norma lexicológica es elegir la unidad de base sobre la que habrá de llevarse a cabo el recuento.

2. Preparar un *corpus* en vistas a su tratamiento significa elegir entre diversas alternativas: homogeneidad y / o heterogeneidad en la fragmentación de longitud y tema.

3. Si previamente hemos decidido hacer un estudio cuantitativo de un texto tenemos que proceder nuevamente a otra elección para su vaciado: método manual o método automático. El progreso de la informática ha mostrado que la utilización del ordenador garantiza un tratamiento más rápido y seguro por una parte, y por otra, más rentable, si el corpus que deseamos estudiar es de longitud considerable¹⁴.

Vaciar automáticamente un texto significa:

1. Pasar el texto al soporte automático.
2. Hacer un algoritmo que resuelva el problema con un número finito de pasos.
3. Adaptar el algoritmo al lenguaje-máquina más adecuado. Para cuestiones lingüísticas suele utilizarse el lenguaje PL/I.

Antes de proceder a la introducción del texto en la máquina deberemos tener resueltos dos problemas:

1. La unidad que vamos a considerar como base del recuento¹⁵.
2. La ambigüedad de las formas lingüísticas.

¹⁴ Ver Geffroy, A., Lafon, P.; Muc y Tournier, M., «Traitement automatique des textes. La perforation», texto policopiado por ENS de Saint Cloud, 1967; Pecheux, M., *Analyse automatique du discours*, París, 1969; Tournier, M., *Vocabulaire politique et inventaire sur machines*, texto policopiado de la ENS de Saint Cloud, 1967.

¹⁵ Muller, Ch., *Initiation à la Statistique linguistique*, París, 1968 (traducción castellana en Madrid, Gredos, 1973) toma la palabra, como unidad de análisis, entendida como «grupo de letras separado de los grupos vecinos por un blanco o un signo de puntuación».

El sistema utilizado normalmente para la desambiguización es la «indexation» o codificación previa de cada forma. Nosotros pensamos, no obstante, que para un texto de notable longitud no es rentable la codificación previa y por ello sugerimos, previa consulta a estadísticos que avalan nuestro proceder, recurrir al procedimiento manual lo menos posible mediante la muestra representativa establecida por la misma máquina para cada una de las formas que podrían resultar ambiguas.

Aplicando programas adecuados podemos obtener un listado total del texto y un listado sin repetición de todas las unidades del texto —previamente establecidas por la norma lexicológica— con expresión de la ocurrencia (frecuencia absoluta) de cada una de ellas. Así conoceremos, en términos lingüísticos, la longitud global de un texto, la longitud gramatical y lexical del mismo —si hemos diferenciado previamente entre unidades léxicas, pertenecientes a inventarios abiertos, y unidades gramaticales, pertenecientes a inventarios cerrados (el establecimiento de estos dos tipos de unidades deberá ser rigurosamente constante a lo largo del tratamiento).

Si el corpus lo tratamos mediante un programa de frecuencias, conoceremos la frecuencia de cada una de las unidades y, a partir de la frecuencia, será posible la aplicación de tests estadísticos de distribución —función χ^2 , test de Kendall, «R» de Spearman, etc.—, así como la obtención de curvas de uso de una o diversas unidades léxicas.

Un estudio completo de un texto no puede acabar en el análisis de las palabras aisladas sino que debe continuar analizando unidades más amplias, unidades de vocabulario, «vocablos» en palabras de Muller, mediante la lematización de las unidades (agrupación de todas aquellas formas que pertenezcan a un mismo lema). Conoceremos, a partir de la lematización, la extensión del vocabulario de un texto y podremos establecer el coeficiente de gramaticalidad o funcionalidad, el de lexicalidad y el grado de redundancia lexical del mismo.

Prosiguiendo el estudio descontextualizado de las unidades que integran un texto, podemos analizar la correlación de frecuencias de distintos fragmentos —habiendo cuarteado el texto en la preparación del corpus— y, a partir de ello, la distribución de cada una de las unidades, que nos informará de qué vocablos o palabras son diferen-

ciadores ideológicos de un fragmento en base a la irregularidad de su distribución. La aplicación del test de Pearson o función χ^2 nos permite apreciar la regularidad o irregularidad de una unidad:

$$\chi^2 = \sum \frac{(FO - FT)^2}{FT}$$

donde,

FO es la frecuencia observada de cada unidad,

FT es la frecuencia teórica de la misma, y

Σ de la fórmula es el sumatorio de cada unidad en los distintos fragmentos del texto.

Los resultados obtenidos de la aplicación de la función χ^2 en un texto nos permite establecer también aquellos vocablos que son representativos de los temas del texto —distribución regular, χ^2 de valor bajo— y que no actúan como vocablos ideológicos, aquellos de alto valor de χ^2 que son los diferenciadores ideológicos del texto o de los fragmentos.

La aplicación de otras pruebas, todavía descontextualizadas, nos llevará a conocer el grado de correlación entre grupos o entre vocablos:

$$R = 1 - \frac{6 \cdot \Sigma d_i^2}{(n-1)n(n-2)}$$

donde, n es el número de fragmentos que consideramos

Σd_i^2 la suma de los cuadrados de las diferencias de ordenación de los distintos vocablos de dos fragmentos distintos.

También estadísticamente es posible estudiar el campo significativo de una unidad o conjunto de unidades mediante la aplicación de coeficientes tales como la coocurrencia o la distancia a fin de determinar el hilo significativo de un período histórico determinado.

Estos son algunos de los datos que puede proporcionarnos la estadística aplicada al léxico, datos objetivos, aunque sin significación.

Los datos no tienen valor sin una interpretación de los mismos, sin embargo el método lexicométrico debe ser tenido en cuenta por parte del estudioso para una interpretación ideológica más completa de los acontecimientos sociales.

Como conclusión, afirmamos la necesidad de las investigaciones interdisciplinarias, de las investigaciones en colaboración de lingüistas, historiadores, informáticos, estadísticos, sociólogos, economistas, etc., para llegar a poder explicar el sentido global de la historia. El centre de Lexicologie Politique de L'École Normale Supérieure de Saint Cloud, centro que depende del Centre National de la Recherche Scientifique francés, puede servirnos como modelo de una investigación racional, compartida, interdisciplinaria y comprometida, tanto por los textos que estudia como por su propia estructura y composición.

M.^a TERESA CABRÉ CASTELLVÍ