

I. Metodología y elección del tema	19
1. La informática como herramienta, sus artificios... y sus resultados.	20
2. Historia y análisis del discurso. Algunas bases metodológicas	34
A. Del análisis de texto en historia al análisis de discurso en lingüística	34
B. La lexicometría y el análisis de un texto	50

I. METODOLOGÍA Y ELECCIÓN DEL TEMA

Toda disciplina secreta su propia cultura; pero las ciencias humanas, no siendo regidas por consignas duras y explícitas están, más íntimamente que las otras ciencias, modeladas por la aceptación de acuerdos informales, por el “saber-vivir” y el tacto intelectual. Decir esto no significa, de manera alguna, excluir de las ciencias humanas el uso de los métodos duros formales, ni atar de manera categórica estas ciencias a las técnicas de la hermenéutica (...) Sin embargo, a diferencia de las matemáticas o de las ciencias experimentales, el empleo de técnicas duras en ciencias humanas será, al fin de cuentas, evaluado en términos de rendimientos significativos, y no únicamente en razón de la coherencia interna del formalismo utilizado.

Thomas PAVEL, *Le Mirage Linguistique. Essai sur la Modernisation intellectuelle*, 1988.⁹

En el transcurso de esta investigación, que se encuentra en los límites de varias disciplinas y ciencias, nos hemos topado con opiniones que no han cesado de alimentar una reflexión profunda sobre los alcances y los límites de la técnicas de investigación que hemos utilizado para su realización.

Efectivamente, este análisis se propuso como objetivo de base, observar si, en la historia de México, existía o no una evolución del lenguaje político, en un momento en que algunas escuelas historiográficas francesas proponían, de antemano, la segunda opción, es decir: la sola continuidad del discurso político.

⁹ Paris, Minuit, pp. 40-41.

Con este fin, estudiamos un tipo de discurso considerado como “típicamente” oficial, discurso que fue rastreado durante un siglo en el que, *a priori*, se podía pensar, que a causa del período revolucionario que lo “divide”, habían existido dos tipos de poder diferente manejando las riendas de la dirección del país.

Por otro lado, para la realización de este análisis, se tuvo que prestar atención a ciertas exigencias de tipo lingüístico. En primer lugar, debido a la ambivalencia del propio objeto de estudio y, en segundo lugar, debido al hecho de que se utilizaron algunos de los métodos y técnicas de análisis de discurso que pertenecen al campo de la lingüística, como es la lexicometría, método que, como su nombre lo indica, consiste en la medición del léxico empleado en un texto determinado.

En tal contexto de análisis, en que métodos y formas de aproximación provenientes de diferentes ciencias se entrecruzan, era necesario esbozar los requisitos que cada una de estas disciplinas nos invitaban a cumplir para una correcta manipulación de sus técnicas. Era pues necesario, plantearse toda una serie de cuestionamientos concernientes a dicha forma de aproximación pluridisciplinaria, tales como: ¿qué requisitos nos solicitaban cada una de las disciplinas puestas en obra?, ¿si las concesiones acordadas a una, no nos hacían traicionar las exigencias de la otra?, ¿llegamos a resolver las contradicciones que en algún momento aparecieron? Todas éstas son interrogantes cuyo planteamiento ha sido desarrollado en esta primera parte de nuestro trabajo.

1. LA INFORMÁTICA COMO HERRAMIENTA, SUS ARTIFICIOS... Y SUS RESULTADOS

Historia, medida... Casar estos dos términos no es algo nuevo. Desde el siglo XIX, algunas mentes ya se habían planteado el problema del recurso a la medida en historia (...) de hecho, la escuela histórica francesa, pronto se forjó una excelente reputación en el dominio de la historia cuantitativa. Sin embargo, la moda ya no radica en ocuparse de este tipo de preocupaciones, otros debates agitan aparentemente, la comunidad científica

ca (...), (así) ante esta nueva situación, existen dos actitudes posibles (...): o rechazar todo en conjunto, ya que no se comprende, o aceptar todo puesto que la sofisticación de los métodos garantiza la cientificidad...¹⁰

Los editorialistas del primer número de la revista *Historia y Medida*, nos muestran, en 1986, cómo la discusión sobre la legitimidad misma de la contabilidad de datos históricos¹¹ pasó al dominio epistemológico, haciendo que los investigadores se cuestionaran sobre la utilización y la generalización de la aplicación de la informática en ciencias humanas. Esto se debió a que, en efecto, la introducción de estas “máquinas” —por no darles otro nombre por el momento— ha ocasionado un cambio radical de la aproximación cuantitativa de la historia, replanteando, a su vez —y otra vez—, la pertinencia de su propia aplicación.

Así, hablando sobre la “naturaleza y la forma de los datos con los que trabaja el historiador”, Jean-Philippe Genêt nos dice:

El “real histórico”, está constituido por el conjunto de *informaciones disponibles* sobre un momento dado de la historia. Ese “real histórico”, a su vez, *no es otra cosa que una parte del real pasado*, puesto que nosotros no podemos conocer *ni aquello que no existe en el estado de información*, entendiéndose, sin embargo, que siempre existen informaciones potenciales más allá de aquellas que ya son conocidas, *ni la extensión de aquello que no existe*. El primer trabajo del historiador consiste (pues) en conocer científicamente lo “real histórico”, lo más (cercanamente) posible, lo mejor posible. *El trabajo del historiador sobre las fuentes (...) transforma el “real histórico” en una colección de*

¹⁰ Editorial de la revista *Histoire et Mesure*, 1986, vol. 1, no. 1, p. 5. Esta afirmación, un tanto irónica debido a la pertinencia real que presentaba la temática de la cuantificación, parece haberse confirmado con el paso del tiempo. Hoy, en 1994, las diferentes escuelas historiográficas francesas realizan todo un balance sobre los métodos cuantitativos, sus usos y abusos a causa de la poca precaución metodológica puesta en obra en varios trabajos de este tipo. El debate parece pues no sólo seguir en pie, sino que ha recobrado su pertinencia, regresándose a esas “otras preocupaciones y debates” de los años 80, a las que el autor hace referencia en estas líneas.

¹¹ Sobre este punto particular y sobre la utilización en general de la informática en historia, ver el artículo de Jean-Philippe Genêt, *Outils et Demarches*, pp. 7-18. Para toda reflexión teórica sobre este tipo de aplicaciones, aconsejamos la lectura de esta revista que presenta tanto reflexiones de tipo epistemológico, como todo un abanico de ejemplos de aplicaciones de estas nuevas metodologías en el campo de la investigación historiográfica.

datos científicamente construida: es esta colección, a la que propongo llamar metafuentes, que será lo que se someterá a la computadora; después de todo, el proceso aquí descrito, no es tan diferente del procedimiento seguido cuando no se utiliza la computadora...¹² (lo cursivo es nuestro).

Así, la utilización de la informática en el dominio de la historia no ha sido tan evidente como muchas veces pudiera parecer y, por el contrario, sí ha provocado numerosas dificultades de tipo epistemológico. En gran parte, el tipo de actitudes descritas en esta cita (aceptar o rechazar todo en conjunto) han sido ocasionadas por varios motivos. Primeramente, por una falta de costumbre del manejo de este tipo de metodologías, enseguida, por la revolución de costumbres que la introducción del “automatismo” causó en el dominio de la investigación historiográfica y, finalmente, por la falta de precauciones con las que, en ocasiones, se han aplicado dichas metodologías.¹³ De ahí que sea necesario replantearse, reflexionar sobre esta aplicación desde su principio —aunque por momentos parezca inútil—, para así esclarecer la manera en que todas estas herramientas fueron repensadas para su aplicación en el marco de la presente investigación.

Cuando los logiciales (paquetes de programas) son concebidos, elaborados y producidos, éstos se hacen partiendo del principio de que hay que facilitar y agilizar ciertas tareas dentro de un sector especializado. En este sentido, dicha elaboración necesita tomar en cuenta los conocimientos de una u otra ciencia, de tal manera que uno de los principales desafíos de los conceptualizadores de dichos logiciales —los ingenieros programadores—, es el de conciliar las formalidades del algoritmo matemático con aquellas formalidades de la ciencia en que un determinado logical se va a aplicar. Esto ya implica, de por sí, una “lucha” entre las leyes de cada uno de los terrenos que entran en juego para la elaboración de dichos programas.

Sin embargo, este tipo de “lucha” siempre ha existido en todo tipo de investigación científica. Un ejemplo similar, sería aquel sufrido por la misma historia que, para devenir Historia, es decir, con ciudadanía dentro del mundo científico, tuvo que distanciarse y aun combatir, formas artesanales de escribir la historia, para frecuente-

¹² *Idem*, p. 7-8.

¹³ Pavel, Thomas, “Le mirage....”, *op. cit.*, pp. 187-189.

mente regresar más tarde, a los mismos tipos de historia que en un momento combatió, pero ya con una visión renovada y contando quizá con otros medios técnicos de aproximación. Estos hechos son tal vez más evidentes en las ciencias llamadas “exactas”, aunque es interesante observar que en este campo, estos altercados no violentan tanto la comunidad que participa ni se llegan a polarizar tanto las posiciones.

Dicho de otra manera: aquello que llamamos las ciencias —¿o la Ciencia?—, no es inmutable ni es un ente aparte a descubrir. Las ciencias son construidas por los hombres, por la época y la historia de esos hombres, lo que, a su vez, construye la propia historia de las ciencias, con sus distanciamientos, sus coaliciones, sus alianzas, sus estrategias y reconciliaciones.

Es por estas razones que siempre nos pareció un poco aventurado discutir sobre la validez o la falta de validez de la utilización de la informática en la investigación histórica (es decir, una utilización diferente al simple tratamiento de texto), antes de intentar llevar a cabo toda una investigación por medios computarizados. En resumen: negarse a esta “aventura” en un momento en que estas prácticas comenzaban apenas a concretarse en Francia, nos pareció negarnos, ante el simple temor de lo desconocido y de lo que aparentemente parecía complicado, a nuevas vías, nuevas lógicas y nuevas formas de reflexionar no sólo las fuentes historiográficas, sino la propia escritura de la historia.

Frente a esta actitud timorata, preferimos tratar de adentrarnos en lo que las nuevas utilidades de la informática en el campo de las ciencias sociales podían ofrecer; observar e intercambiar las diferentes experiencias, las cuales inspiraron líneas y modalidades de sus aplicaciones; y, finalmente, participar con los grupos de personas que reflexionan sobre la manera en que esta herramienta puede llegar a ser más eficaz y adaptable a los diferentes tipos de necesidades.

Durante los años 80, dos factores retrasaron la generalización de una utilización profundizada de metodologías computarizadas: por una parte, lo que podríamos llamar la inercia generacional y las dificultades de “reconversión” de investigadores adultos recientemente iniciados a la computadora; por otro lado, la fulgurante evolución que se sintió en todos los elementos que conforman la herramienta

informática: desde las computadoras mismas, hasta los logicales, pasando por la prodigiosa proliferación que se dio, a partir de 1987, de los *scanners* o “lectores ópticos”.

La carrera que esta evolución técnica impuso, no se dio del todo sin algunos descabros, lo cual reforzó las argumentaciones de los “anti-computacionales”... dentro del sector privado, estas actitudes fueron coronadas con el puro y simple despido de los renuentes a la “reconversión”.

En cuanto a los “aventureros” de las ciencias sociales, ellos tuvieron que hacer evolucionar sus propias investigaciones al ritmo de las evoluciones del material que iba incursionando en los laboratorios de investigación, lo cual implicaba, muchas veces, reandar lo andado para adaptar los datos a los nuevos sistemas. Así, por no dar si no el ejemplo más general, el pasaje, en 4 años, de las hoy dinosaurias “fichas perforadas”, al *main-frame* y, finalmente, a los microsistemas, lo cual no sólo implicaba diferentes formas de “almacenaje” de datos, sino diferentes formas de comunicarse con cada uno de los tipos de máquina y, consecuentemente, todo un aprendizaje y saltos malabarísticos de nuestras lógicas personales.¹⁴

Lo mismo en aquello que concierne a los logicales, para lo cual se tenían que readaptar los datos para su tratamiento, sin olvidar que, en aquella época, las posibilidades de comunicación entre los diferentes logicales eran casi inexistentes, sobre todo cuando se trataba de paquetes especializados (genealogía, población, lexicometría, encuestas, etcétera).

Dicho esto, también habría que señalar que, al principio de los 90, la evolución técnica —al menos de la microcomputación— parece haber llegado a un punto de *stagnation* debido a varios factores. Uno sería de orden comercial (imposibilidad de vender modelos diferentes cada año), y el otro la imposibilidad de continuar sobre las mismas bases la competencia técnica entre los dos grandes de la microcomputación, Apple e IBM, el primero apostando su desarrollo sobre la convivialidad de sus sistemas y el segundo sobre la potencia de sus

¹⁴ Y esto refiriéndonos no sólo a un tipo de lenguaje de programación cualquiera, sino a la manera misma que el usuario tenía para comunicarse con la máquina misma, cualquiera que ésta fuera, pero que cambiaba enormemente de un sistema a otro. Este problema se ha resuelto en gran medida con la expansión del logicial-utilitario.

logicales. De ahí que estos colosos hayan decidido unir sus fuerzas para sacar al mercado un producto común.

Este período “de calma” ha permitido regresar sobre toda una serie de aspectos de la aplicación computarizada en ciencias sociales, que habían ido quedando rezagados en los años del *boom* informático. De ahí que también se esté de nuevo profundizando sobre las bases epistemológicas de dicha aplicación, lo mismo que repensando la forma de utilización de sistemas de manera que los formatos tengan una mayor compatibilidad, y esto, inclusive, a nivel internacional. En resumen, un tiempo de paz necesario para hacer un balance total, desde el material utilizado hasta las formas de aplicación, lo cual ya se venía resintiendo desde finales de los 80.

Esto nos lleva a hablar de la necesidad de una mayor homogeneización, o al menos, de la pertinencia de aplicaciones cuidadosas según el tema estudiado; homogeneización y pertinencia en la que los centros de cálculo deben seguir jugando un rol mayor para la solución de dichas tareas. Así, los datos de una encuesta sociológica no pueden ser tratados de la misma manera que la misma encuesta en el ámbito histórico. El historiador, remontando hacia el pasado, tenderá a toparse cada vez más con un cambio de género de datos y/o, pura y simplemente, con un vacío de datos, en todo caso, en mayor medida que el sociólogo.

En este sentido, si se plantea una investigación pluridisciplinaria en la que se deseen comparar épocas pasadas con otras más actuales, el microrordenador nos dará resultados. Siempre da. Sin embargo, éstos se prestarán a falsas interpretaciones si no se toman las precauciones estadísticas necesarias para su tratamiento, precauciones que son traducidas en algoritmos por los analistas programadores. Estas precauciones pueden consistir, entre otras cosas, en ponderar los vacíos de los datos historiográficos en relación con los datos sociológicos, en homologar valores y equivalencias y en no perder la perspectiva de la falta de datos más generales, que son los que dan perspectiva a los datos con los que en realidad se cuenta.

Por su parte, estas modificaciones algorítmicas pueden intervenir, como es lógico, en un paquete creado específicamente para un tipo de investigación cuyas funciones son inexistentes en los paquetes vendidos en el comercio. Así, un “paquete”, completamente acabado

y que no encontraremos en el comercio, es por ejemplo el que ha sido desarrollado en la Universidad de Rouen y que permite la observación de la venta de bienes nacionales después de la revolución en Francia; sin embargo, este logicial incluye todo un sistema de conversión que permite aplicar el mismo tipo de función a otros países, adaptando sus cálculos a las medidas del agro de otro país... permitiendo así estudios comparativos sobre esta temática en varios países.¹⁵

Estas modificaciones también pueden introducirse en lo que se denomina un “pro-logicial”, creado para ser aplicado antes o después del tratamiento de logiciales comerciales, con el fin de operar el tipo de modificaciones anteriormente señaladas.

Es en este sentido que pensamos que el rol de centros de cómputo sigue siendo de capital importancia; sin hablar de la capacidad “natural” de estos centros para elaboración de logiciales especializados, como del que hablamos con anterioridad, pero también aquellos otros especializados en genealogías, lexicometría, cartografía, demografía, etcétera, y que se encuentran, y se encontrarán, fuera de una demanda comercial, pero que son esenciales para el desarrollo de la investigación social y el intercambio académico en la era de la globalización.

El reto pues, hoy por hoy, sigue siendo el de la creación, pero acompañada por un esfuerzo de simplificación de utilización y, sobre todo, por un esfuerzo de homologación y compatibilidad. Para esto es necesario una mayor comunicación entre todos aquellos laboratorios, centros y equipos que ya se han lanzado por la vía de la computación. Y aunque en historia sigue existiendo el debate sobre la informatización de fuentes, pensamos que el proceso ya ha sido puesto en marcha —al menos en lo que a Francia se refiere— y que la reflexión que pedíamos hace unos años ya ha sido seriamente abordada por las autoridades burocráticas y académicas¹⁶ correspondientes.

¹⁵ Universidad de Rouen, Instituto de Estudios sobre la Revolución Francesa, Bernard Bodinet.

¹⁶ Así, por ejemplo en el terreno historiográfico, se han hecho varios esfuerzos en este sentido. Primeramente con la reflexión para la construcción y ordenamiento de la “Très Grande Bibliothèque” de París, que recibirá el acervo de la Biblioteca Nacional y se ligará a los Archivos Históricos de Versalles. Más específicamente, dentro del ramo historiográfico, varias instancias ya han realizado grandes esfuerzos en este sentido, así el Central Archive for Empirical Social Science que se encuentra en Colonia, la Asociación internacional “History and Computing”, *cfr. Historical Social Research, Bulletin de l'Association Histoire et Informatique*. Lo mismo sucede con los archivos de Relaciones Exteriores, concentrados de manera completamente informatizada en la ciudad de Nantes, y que ya han recibido en formación a historiadores mexicanos.

Como lo sugería Jean-Philippe Genêt desde 1986, y como es evidente que esta reflexión ha evolucionado, la discusión se sitúa precisamente a ese nivel, es decir: ¿qué criterios se deben utilizar para desarrollar y mejorar los métodos de la aplicación informática?, y ya no saber si es o no un buen criterio aplicar la informática. Es en este sentido y bajo esta perspectiva que radica la importancia de todo trabajo actual que permita alimentar esta reflexión, al mismo tiempo que difundir la investigación computarizada, con todos los atenuantes que ésta pueda presentar, con el objetivo de ir perfeccionando y acoplando las diferentes modalidades que es posible encontrar.

Inclusive, se puede decir que el trabajo realizado en ciencias humanas y sociales en esa época en Francia estuvo alimentado por experimentos y tanteos que, actualmente, están llegando a formar el asiento metodológico para utilidades mejor parametradas. En este caso se encuentra por ejemplo la lingüística, cuyas aplicaciones están siendo ligadas al medio industrial para la creación de sistemas expertos especializados.¹⁷ Pero no es lo mismo tan sólo señalar lo anterior de esta manera y/o utilizarlo hoy en día, cuando los programas ya han pasado por procesos de modificación y perfeccionamiento en su aplicación y cuando las reglas metodológicas ya han sido explicitadas, que en el momento de formación de dichas reglas.

Así pues, es desde este par de perspectivas que nació nuestra voluntad de participar en la experiencia que entonces se desarrollaba en Francia. Y es pensando en las nuevas generaciones, hoy más tempranamente informatizadas, que nos parece necesario hacer un recorrido por algunos de los vericuetos por los que nos hizo pasar la —en aquel entonces— novedosa utilización del automatismo. Las cosas no se dan por sí solas y lo que hoy parece “evidente” y/o fácil de usar, se debe a la reflexión de aquel conjunto de utilizadores, exploradores de terreno.

Tratemos de ver, pues, cuáles fueron las peripecias que, en nuestro caso, se tuvieron que sobrepasar.

La idea historiográfica de partida, era realizar una comparación de los informes presidenciales a largo plazo. Para esto, se tomaron el primer y último discurso de cada mandato presidencial de 1877 a

¹⁷ Dosse, François, *Histoire du Structuralisme*, Paris, La Découverte, col. Textes à l'appui/série histoire contemporaine, Paris, 1991 y 1992; cfr. t. 2: *Le chant du cygne. 1967 à nos jours*, p. 353 y, en forma general, la obra de Thomas Pavel citada anteriormente.

1976, es decir: del primer período de Porfirio Díaz al de Luis Echeverría A., de 1970 a 1976.

Visto así, esto no supone ningún problema particular. Sin embargo, cuando se trata de ligar esta selección a los instrumentos técnicos con los cuales se contaba en 1985 para su tratamiento, las cosas cambian y, aquello que normalmente es un texto común y corriente, se convierte en un rompecabezas en el que entran en juego otra serie de factores como el tipo de impresión, las hojas utilizadas, etcétera.

Estando en París, la única edición más o menos completa que encontramos de la mayoría de estos discursos, fue una edición realizada por la Cámara de Diputados en 1966 de *Los presidentes de México ante la Nación*,¹⁸ recopilación de informes y mensajes presidenciales que la Cámara editaba con una frecuencia poco determinada. Nos faltaban pues, el último discurso de Díaz Ordaz (1970) y el primero y último de Echeverría (1971, 1976), que encontramos en ediciones diferentes de más reciente publicación.

En cuanto a tecnología se refiere, el Laboratorio de Informática para las Ciencias del Hombre (LISH en adelante), laboratorio que apoyó toda la parte técnica de nuestra investigación, adquirió en 1985-86 una máquina de captura automática construida por Helwett-Packard, llamada en la época lector óptico y cuya evolución daría lo que hoy se conoce con el término de *scanner*. Segundo en su género en toda Francia, este *scanner* era, en comparación con los actuales, todo un dinosaurio, tanto por su tamaño descomunal (entre otras cosas, dos grandes unidades de 2 por 1 mts.), como por su funcionamiento (*bandas* de registro, teclado *sólo en inglés* sin posibilidad de cambio, pantalla blanco y negro, etcétera). Sin embargo, no existía otra cosa y nos tuvimos que lanzar en el manejo de una máquina que nadie sabía utilizar y que nunca habíamos visto funcionar.

Pero independientemente de la novedad, y cualquiera que sea el modelo que se utilice, incluyendo los más recientes, todos los *scanners* necesitan de un aprendizaje previo a la captura de textos. Es decir, necesitan que uno mismo les enseñe a reconocer los tipos de caracteres utilizados, de manera que puedan ser reconocidos por el *scanner*.

¹⁸ *Los presidentes de México ante la Nación. Informes, manifiestos, y documentos (1821-1966)*, México, XLVI Legislatura de la Cámara de Diputados, México, 1966, 4 vol.

Sin embargo, como ya lo señalamos, ese lector contaba sólo con el teclado norteamericano, de manera que si el texto que se debía capturar no se encontraba en inglés, la determinación de caracteres dichos “especiales” de cada idioma (*cfr.* “ñ” en español, “ç” en francés, etcétera), se transformaba en la determinación gráfica de todo un “metagrafismo”. De suerte que, para que el lector óptico leyera las vocales acentuadas del español, se le tenía que forzar a interpretar *a' e' i' o' u' n** y *u&*, como *á é í ó ú ñ* y *ü*.¹⁹

Otro tipo de problema técnico, radicó en que el *scanner* utilizado, realizaba el reconocimiento de caracteres en líneas verticales, a diferencia de los *scanners* de “segunda generación”, cuyo reconocimiento es efectuado en puntos. Así, una *c* era visualizada por el lector de la manera siguiente: ⋄, lo cual hacía que el reconocimiento fuese más difícil, sobre todo cuando las ediciones capturadas eran antiguas y de no muy buena calidad, presentaba gran cantidad de imperfecciones tipográficas. Dificultad también para textos que datan por ejemplo, de la Edad Media, como *La Gallia*, que debido a las diferentes ediciones de que ha sido objeto y de los manejos paleográficos correspondientes, se encuentra hoy con una gran variedad de fuentes de caracteres.

El reconocimiento de caracteres por puntos, utilizado por los *scanners* de la “segunda generación” (*c* = ⋄), permitió una mayor fineza de reconocimiento. Lo mismo sucedió con aquellos de más reciente creación y cuyo reconocimiento es efectuado por *imagen*, es decir, casi una fotografía del texto, la cual será corregida en detalle, una vez realizada la captura.

Desafortunadamente, este *boom* de avances tecnológicos de la lectura automatizada, empezó sólo en 1987; los centros de investigación se dotaron de este *hardware* hacia 1988, es decir, una vez que nuestros textos habían sido capturados casi en su totalidad.

La incompatibilidad de terminales implicó también, en la época, un obstáculo para la rapidez del tratamiento, puesto que los paquetes

19 En francés, é era *e/, è e/, î i,* etcétera. Al final de cuentas, debido a una falta de reglas coordinadas, aun si dos investigadores trabajaban sobre textos en francés, sucedía que cada uno determinara su propio metagrafismo, lo cual, de continuar así, hubiera ocasionado problemas de homología para estudios comunes de aquellos textos, o en el caso de elaboraciones de bancos de datos. Este problema ha sido en parte solucionado con la posibilidad de cambio de teclados virtuales. Faltaría una mayor coherencia de formatos. En todo caso, los textos capturados en aquella época, debieron ser “homologados”.

no existían aún para microcomputadora. Tenía pues que enviarse aquella captura realizada en bandas magnéticas por el *scanner*, al Centro de Cálculo Central Universitario (en lo sucesivo CIRCE), aplicar un programa para la reconversión de nuestros caracteres especiales al alfabeto normal y de éste al “lenguaje máquina” utilizado en el CIRCE.²⁰

Todo esto, a intermitencias —por la ida y venida de los textos de microcomputadora a sistema central— de la preparación normal del texto, requerida para el tratamiento por los lógicos de lexicometría.

Para empezar, juntar todos y cada uno de los informes en un solo *corpus*. Enseguida, proceder a la unigrafía de aquellas palabras que aparecían divididas en dos en fin de línea, pues debieron reacomodarse para evitar que fuesen reconocidas por los programas lexicométricos como dos palabras diferentes desprovistas completamente de sentido. Esto es, si la palabra “institucional” aparecía en fin de línea: “... institu-”, para continuar al principio de la siguiente línea “cional...”, tuvo que elaborarse un programa especial para juntar en un solo grafismo todas las palabras que aparecían normalmente divididas de esta forma en los textos.

Por otro lado, si de un informe a otro, en un mismo mandato presidencial, la presentación de los discursos varía, ¿qué será a lo largo de un período de cien años? Así pues, para tratar de homogeneizar lo mejor posible todos los textos, se tuvieron que realizar ciertas modificaciones en la mayoría de los discursos, de manera que se evitara una dispersión de resultados demasiado importante, sin perder de vista los cambios efectuados ya que, en ciertos casos, las formas de presentación no son fortuitas.

En este sentido, la intervención más importante fue la que se efectuó a nivel de los números. Por ejemplo, a veces las fechas aparecían escritas completamente en letras y, en otros casos, en cifras. Así, “22 de Diciembre de 1938”, o “22 de diciembre de 1938”, fue presen-

²⁰ Centre Inter Regional de Calcul Electronique, Universidad de Orsay. Cabe señalar aquí la incompatibilidad que sigue subsistiendo, sobre todo en el pasaje de microcomputadora/ “mainframe”, de los valores hexadecimales de los diferentes caracteres. Estos valores son aquellos con los que un sistema dado va a representar, en su lenguaje-máquina un carácter numérico o alfanumérico, es decir, un carácter del alfabeto o una cifra. Por ejemplo, el valor hexadecimal de á es: 3F en el CIRCE en Francia; en un micro Zenith (hoy Bull) funcionando bajo DOS versión 6 (fabricado en Inglaterra), es 45 y en la micro que me encuentro escribiendo actualmente es 160.

tado “22-diciembre-1938” o “15-septiembre-1956”, etcétera. En cuanto a las cifras —en pesos, dólares o kilómetros, etcétera—, la presentación podía variar de “mil cuatrocientos ochenta y cuatro pesos, con cincuenta centavos” (Miguel Alemán), a “\$ 1484 con cincuenta centavos”. Todas estas variantes fueron transformadas a la presentación general de: “\$1484-50”, “dils-1484-50”, o bien “1458-50-kg”, el punto o la coma y, en general, todo *signo de puntuación* debió quedar sólo con su única función de *separador gramatical*.

Todo este último tipo de correcciones tuvo que efectuarse en detalle. Una vez realizada esta homogenización de forma del corpus, más la corrección de ciertos errores de lectura del scanner (inevitables aun en la escanerización actual) nuestros textos necesitaron recibir una *codificación lineal y referencial* necesaria en todo tratamiento lexicométrico para localizar los diferentes elementos al interior de cada discurso, es decir: Orador, “fecha del discurso” (primero o último), tomo del libro, página, línea. Esto lo veremos de más cerca un poco más adelante.

En cuanto a la preparación de parámetros en los programas del *Lexicloud*, tratamos de dejar la más amplia apertura de manera de poder aplicar y probar todos los programas de dicho paquete.²¹

Sin embargo, la perspectiva cambió un poco, puesto que desde el principio hubiéramos querido probar todos los programas en el *corpus* como un todo, es decir, desde el informe de 1877 de Porfirio Díaz, hasta el de 1976 de Luis Echeverría. Pero como dicho paquete nunca había sido utilizado en un *corpus* de esa magnitud (800,000 palabras aproximadamente), al principio nos topamos con ciertas dificultades de aplicación. De ahí que tuviéramos que dividir el *corpus* en dos *corpus* diferentes: uno que llamamos DIAZ, en el que se encontraban todos los informes del período de Porfirio Díaz, y otro que, por la facilidad de manejo del parametraje necesario para la aplicación de cada programa, llamamos simplemente PRI, y que contenía los informes desde Francisco de la Barra hasta Luis Echeve-

21 Efectivamente, en función de las referencias que uno puede señalar en el parametraje de programas, se pueden aplicar diferentes formas de tratar el texto. Para este trabajo, nosotros lo hicimos por discurso y por presidente, lo cual sin embargo, ya nos dio resultados generales concernientes a los discursos de “entrada” y los discursos de “salida”. Pero pensando en otras formas de explotación de este *corpus*, dejamos el espacio necesario para una codificación temática, o por ministerios, lo que nos permitiría —aún hoy, una vez terminada la investigación por personajes— combinar los resultados con otros criterios de análisis.

ría,²² aplicando todos y cada uno de los programas sobre cada uno de estos *corpus* por separado.

Con todo, esta división de *corpus* no alteró nuestra elección de principio en tanto que análisis comparativo, por un lado. Por el otro, en lo que se refiere a la explotación lexicométrica —es decir estadística—, el hecho de aplicar los programas a dos *corpus* diferentes, tampoco alteró los resultados del conteo de frecuencias absolutas o de ocurrencias de una palabra, de la “repartición” de dichas frecuencias en los diferentes discursos que forman el *corpus* o de sus “concordancias”.²³ Todo lo contrario: permitió que los resultados del *corpus* PRI, numéricamente más importantes que los del *corpus* DIAZ —debido a la diferencia misma de las dimensiones de discursos—, no desviarán nuestra atención de los resultados de este primer *corpus*.

Por la misma razón —y paradójicamente, ya que esta división de *corpus* se había realizado por una “deficiencia” técnica que sería ulteriormente resuelta—, este tratamiento lexicométrico en 2 *corpus* por separado, benefició inclusive la apreciación de resultados de programas que efectuaban aproximaciones más estadísticas, es decir: aproximaciones que, ya de por sí, ponderaban la relación entre los resultados brutos de las partes de un *corpus* —un discurso—, y el tamaño de cada uno de éstos.

Esto lo veremos más en detalle en un capítulo ulterior.²⁴ Por el momento sólo baste decir que: los programas fueron aplicados pri-

22 Hacemos hincapié en el hecho de que este nombre (PRI), no responde, evidente y estrictamente hablando, a un manejo historiográfico del *corpus*, sino única y exclusivamente informático. Este segundo *corpus* pudo haberse llamado “SOL”, “XX” —por el siglo—, etcétera. Sin embargo PRI nos pareció ser más transparente y claro para el manejo de los resultados informatizados, que serían de difícil manipulación debido a su abundancia y diversidad.

23 Ocurrencia = número de veces en los que una palabra aparece en el *corpus* o en una de las partes de ese *corpus* (se puede hablar de un término “genérico”). Formas = número de palabras diferentes que aparecen en un *corpus* o en una de sus partes. Frecuencia total = número de ocurrencias de una misma palabra en la totalidad del *corpus*, por oposición a subfrecuencia = que es el número de veces en que una palabra aparece en cada una de las partes del *corpus*, tomadas por separado. Concordancia = palabras que se encuentran alrededor de una palabra dada —“palabra-clave”—, lo cual sirve para analizar el contexto en el cual dicha palabra está diversamente —¿u homogéneamente?— inmersa. Según los programas, se pueden tener una o varias líneas de contexto, alrededor, antes o después de la palabra-clave y dichos contextos pueden ser acomodados por orden alfabético o de aparición en el texto. Co-ocurrencia = dos (y según los lógicos, hasta tres palabras) que de manera sistemática se presentan una alrededor de la otra, como por ejemplo: “Estado” y “derecho”, que puede ser Estado de derecho, etcétera.

24 Esto también nos permitió comparar resultados, los cuales no presentaron mayores anomalías.

mero a dos *corpus* diferentes y, enseguida al conjunto de estos dos *corpus* como totalidad: al *corpus* “DIAZ”, que contenía 17 discursos de dos personajes diferentes (Manuel González); el *corpus* “PRI”, que contenía 29 discursos de 17 personajes diferentes; y, finalmente, el *corpus* total que llamamos PREMEX, y que consecuentemente está formado por 46 discursos emitidos por 31 oradores.

Las dos primeras aplicaciones resultaron ser “ampliaciones” de la aplicación total, que nos permitieron un análisis más detallado del *corpus* “DIAZ”. La aplicación a la totalidad del *corpus*, nos permitió asegurarnos de aquellos primeros resultados, y nos hizo ver la imposibilidad, lexicométrica, de un estudio detallado cuando los textos presentan una diferencia demasiado importante de sus dimensiones respectivas.

Como lo veremos en el capítulo consagrado al análisis factorial, ambos estudios se complementan. Pero un procedimiento que se originó por un “falla” técnica, nos permitió dar un rodeo que terminó alimentando más nuestro análisis que el habernos restringido a nuestro análisis global.

En términos generales se puede afirmar que con los recientes progresos, tanto a nivel técnico —desde los procedimientos de captura automatizada hasta la tendencia a la simplificación de los paquetes utilizados, pasando por la compatibilidad y rapidez de sistemas—, como a nivel epistemológico —por los balances metodológicos y de aplicación que se están llevando a cabo—, este tipo de “técnicas duras” ha ganado flexibilidad y, consecuentemente, terreno en su adaptación a estudios que tradicionalmente le estaban vedados o cuyas aplicaciones despertaban sospechas. Sin embargo, ante esta perspectiva, las precauciones en dichas aplicaciones nunca deben perderse de vista, ya que, efectivamente, de una herramienta de gran utilidad se podría pasar a una falsa ilusión de rigor.

Es evidente que no se puede evaluar la utilización de la informática en ciencias humanas “a partir únicamente de la coherencia interna del formalismo” con el cual estas técnicas son aplicadas. Pero dicha evaluación, tampoco puede ser juzgada a partir únicamente de los resultados de una técnica cuyos instrumentos, materiales y epistemológicos, se encuentran en rápida evolución.

De forma más general, y más allá del puro campo historiográfico o de las ciencias humanas, esto nos sitúa en la división artificiosa

que actualmente se quisiera operar entre investigación “fundamental” e investigación “aplicada”, como si la primera no alimentara en gran medida los productos de la segunda.²⁵ De ahí que el concepto mismo de “rendimiento significativo”, necesite una mayor definición: rendimiento de qué y ante qué... y a qué plazo.

2. HISTORIA Y ANÁLISIS DEL DISCURSO. ALGUNAS BASES METODOLÓGICAS

Acabamos de ver la manera en que concebimos la utilización, en general, de la herramienta computarizada en historia. Aun así, nos parece necesario redefinir la utilización del análisis del discurso dentro del marco de la investigación histórica y el porqué de la elección, para la realización de ese análisis, del método lexicométrico.

A. *Del análisis de texto en historia al análisis de discurso en lingüística*

Tenía usted razón. Cualquier dato puede llegar a ser importante si se relaciona con cualquier otro. El tipo de conexión cambia la perspectiva. Induce a pensar que cada aspecto del mundo, cada voz, cada palabra escrita o dicha, no tiene el sentido que aparenta, sino que nos habla de un secreto. El criterio es sencillo: sospechar, siempre sospechar. Se puede leer por transparencia, aun un panel de interdicción de sentido de circulación.

U. ECO, *Le Pendule de Foucault*.²⁶

²⁵ El espejismo lingüístico —refiriéndonos al título de la obra de Thomas Pavel, con cuyos análisis, por otro lado, coincidimos en gran medida—, y su largo trayecto de investigación teórica y fundamental, dio paso precisamente a una lingüística de aplicación industrial —como la llama el mismo autor—, como los sistemas expertos de traducción, entre otras muchas cosas. Sin embargo, el hecho de que la investigación lingüística haya entrado en una fase de “aplicación”, se debe a todos aquellos “experimentos” que muchas veces parecían fantásticos. No obstante, el que por el momento la lingüística se encuentre en esta fase de aplicación, no quiere decir que dicha aplicación no está engendrando cuestionamientos que no podrán ser resueltos más que por la vía de la llamada investigación “fundamental”. Esta separación estricta, pues, no responde a aquello que pura y simplemente es la investigación.

²⁶ Ed. Enaudi, 1989, pp. 386-387. Traducido del francés. Preferimos, para la presente publicación, hacer nuestra propia traducción de la versión francesa de esta obra, ya que fue la primera versión que leímos, y la que inspiró su utilización en el transcurso de la redacción de

En términos generales, el análisis de discurso no parece presentar grandes dificultades ya que, de una forma u otra, el historiador siempre lo ha practicado.

En efecto, a pesar de que el investigador actual regrese a las más antiguas fuentes de información —como es el caso de la historia oral—, o que recurra a métodos de otras ciencias de más reciente desarrollo —como por ejemplo la arqueología o la sociología—, la fuente por excelencia del historiador continúa siendo el texto escrito e inclusive, más recientemente, grabado o filmado. Son los textos los que proporcionan lo esencial de la información al investigador, y es a través del análisis de textos y de su mutuo cotejo, que tradicionalmente el historiador ha tratado de abordar los tiempos pasados.

Para esto, el historiador reflexiona sobre las condiciones de producción del texto, al mismo tiempo que sobre aquello que el texto dice, lo que suministra al investigador informaciones sobre el periodo estudiado. Y aunque en un momento dado los historiadores resintieron la necesidad de establecer qué textos podían ser considerados como falsos y cuáles como auténticos, la experiencia les ha mostrado que aun el estudio de documentos falsos, proporciona informaciones preciosas sobre un momento o un evento analizado.

Por otra parte, la aparición de técnicas que permiten el archivado sistemático, no sólo de textos —cualquiera que sea su género: literarios, panfletos, actas notariales o parroquiales o elocuciones—,²⁷ sino también de imágenes y de voces —y, por tanto, de entonaciones— que acompañan ciertos textos, están produciendo de nuevo grandes mutaciones en el oficio de historiador.

Efectivamente, en la actualidad, observar mediante video-grabaciones cómo una mirada, un gesto, una entonación cambian el contenido de una palabra expresada oralmente —en relación a la misma palabra expresada por escrito—, ha modificado también las modalidades de aproximación de textos en historiografía.

nuestro trabajo. Advertimos que nuestra traducción busca más una fidelidad al sentido de la versión que nos inspiró, que una armonía estética; sin embargo, esta versión ha sido cotejada con la versión en español de Bompiani-Lumen-Orbis-Patria, 6a. ed., sept. 1990, sin haberse encontrado mayores contrasentidos.

27 De ahí la creciente importancia de la elaboración —y la correspondiente homogeneización de formatos— de bases de datos, como fue señalado en el precedente apartado.

Es evidente que el historiador no dispone de la misma gama de datos, como lo hace el sociólogo o el politólogo, quienes pueden ofrecer interpretaciones más “precisas” sobre la producción y las consecuencias de un discurso dado, ya que tienen la posibilidad de contar, por ejemplo, con los hoy tan solicitados “sondeos de opinión”, una “mayor” información sobre las condiciones de producción e, inclusive, con la opinión misma del autor de aquel discurso o texto.

Sin embargo, el simple hecho de descubrir la influencia de estos factores de creación e impacto en la elocución o difusión de un discurso en el presente, ha hecho que el historiador interroge, desde una perspectiva renovada, los textos que le corresponden —aun los anodinos libros de contabilidad—, tratando en la medida de lo posible, de paliar la falta de fuentes interpretativas —existentes para los tiempos presentes— recurriendo a otro tipo de fuentes de información: artículos periodísticos, grabados o caricaturas y, aun revistas humorísticas.

En el fondo, el procedimiento sigue reposando sobre la misma base, es decir, la confrontación de fuentes. Simplemente se ha ensanchado la gama de fuentes consultadas, pues se ha incluido todo aquello que podía ser imagen, chascarrillo, modas y, de forma general, representaciones de personajes, ambientes y momentos, materiales que, hasta hace relativamente poco tiempo, habían sido considerados como fuentes poco nobles.²⁸ Diversificación, pues, de fuentes y afinación de las formas interpretativas, es decir: la manera en que pensamos aquel material, la manera en que lo organizamos y lo relacionamos uno con el otro, para así extraer una interpretación o hipótesis.

Esto nos lleva a pensar en el debate que existe actualmente sobre la “historia narrativa” en su aspecto “desconstruccionista”. Evidentemente, el propio discurso historiográfico es una construcción, y esto tanto por los elementos que se toman en cuenta para la reali-

²⁸ Como ejemplo, podemos citar los excelentes trabajos sobre la caricatura en la época de la Revolución francesa, realizados por Pascal Dupuy —Universidad de Rouen—, quien estudia a través de estas imágenes, la percepción de las personalidades de la época, de los cuerpos, etcétera. Otro ejemplo, serían los trabajos de Robert Darnton, quien analizando, por ejemplo, libros de contabilidad trata de aprehender no sus aspectos económicos —cuya importancia no niega de manera alguna—, sino las costumbres que estos escritos pueden reflejar. Varios ejemplos de este tipo de aproximaciones en *L'Image de la Révolution Française*, Paris, Pergamon Press, 1989, 4 vol.; publicado bajo la dirección de Michel Vovelle. Actas del Coloquio, 6-12 de julio de 1989.

zación de tal construcción —tal o cual documento, tal o cual referencia, y la relación que entre ellos se establece—, como por el lenguaje con que estos elementos son descritos y puestos en escena en nuestra narración histórica.

Sin embargo, como nos advierte Roger Chartier, no porque la escritura de la historia sea una narración organizada —un discurso construido—, se puede hablar de una no-diferenciación entre historiografía y ficción, ya que en la primera entraría en juego toda una serie de precauciones metodológicas que conciernen a la científicidad misma del oficio historiográfico, a la que es justificable aspirar en tanto que investigador.²⁹ El que no se respeten dichas reglas es, como en cualquier otra profesión, un asunto de principios deontológicos. Esto no impide que, al no dejar de lado dichos principios, el historiador —a diferencia del aficionado o del escritor de ficción—, tiene una mayor dependencia del archivo que está tratando, y una mayor obligación de cumplir —ante una comunidad de historiadores— con los requisitos metodológicos que poco a poco han hecho de su oficio una profesión.³⁰

Pero regresando a las nuevas influencias metodológicas, ante las cuales la historiografía se ha ido sensibilizando en los últimos 20 años, podemos citar dos que nos interesan de manera más cercana: por una parte, la lingüística y, por otra, las ciencias de la información y de la comunicación. En este proceso de “integración”, las problemáticas de la última han sido más recientemente abordadas que las de la lingüística; pero ha sido esta disciplina la que, con frecuencia, ha alimentado a las técnicas de análisis de las ciencias de la comunicación.

²⁹ Ese “regreso” a la historia narrativa estaría relacionado a una respuesta a modos muy estrictos de hacer y escribir historia correspondientes, *grasso modo*, a las décadas de los 60 y de los 70, por un lado; y, por el otro, a una reacción relativista que esas mismas rigideces acarrearon hacia fines de los años 70 y durante todos los 80. Detallar este debate nos desviaría demasiado de los propósitos que aquí se desean señalar. Sin embargo, para mayor información sobre este debate, *cfr.* Dosse, François, *Histoire du Structuralisme*, cit., t. II, los varios capítulos relacionados a la Historia; ver también Pavel, Thomas, *Le Mirage Linguistique* y, Chartier, Roger, “Narración y verdad”, en *Reforma*, supl. El Ángel, 15 de mayo de 1994, pp. 9 y 10.

³⁰ Roger, Chartier, “Narración y verdad”, *Reforma*, supl. El Ángel, 15 de mayo de 1994, p. 10. Dejamos voluntariamente de lado las cuestiones referentes a comunidad científica, a la cual consideramos junto con Chartier, pertenecen —o aspiran a pertenecer— los historiadores, debido al reacomodo que se efectúa actualmente en los parámetros de dichas comunidades, obediendo a los cambios de internacionalización científica.

Así, algunas de las temáticas de la lingüística que han sido retomadas en historia, son aquellas que conciernen a la producción del discurso, la relación entre las “formas discursivas” y las “formas no discursivas”,³¹ y las diferentes teorías sobre “significantes y significados”, sus formas de interpretación, etcétera.

En cuanto a la influencia de las ciencias de la información y de la comunicación en el ramo historiográfico —sin hablar de los temas epistemológicos que tratan sobre las relaciones entre las ciencias y sobre la construcción de modelos de comunicación entre éstas—, se percibe en todas aquellas temáticas que analizan más específicamente las formas de difusión y de recepción del discurso, ya sea oral o escrito,³² además del interés que ha suscitado recientemente, el estudio de la evolución de los propios medios de comunicación tanto a nivel de la historia en general, como en el de la historia cultural.³³

Como ya se señaló, es evidente que estos dos campos se han encontrado íntimamente ligados en ciertos momentos; así, muchos de los análisis periodísticos (o televisivos) de los comentaristas de la vida política, se han inspirado directamente en metodologías desarrolladas por los lingüistas.

Si para la realización de esta investigación, en 1985, nos hubiéramos encontrado en otro contexto intelectual, esa “invasión” de terminologías dentro del campo historiográfico, nos hubiera parecido un tanto extraño. Imagen (“imaginario”), discurso, “lo no-dicho” (*le ‘non-dit’*), “construcción del discurso sobre el discurso”, no hubieran tenido —en ese momento— más eco que un ejercicio novelístico utilizado por Italo Calvino en su obra *Si par une nuit d’hiver un voyageur*.³⁴ Lo que nos hizo reaccionar, fue ver la importancia que tales proposiciones lingüísticas tomaban dentro del conjunto del instrumental analítico de todas las ciencias sociales en Francia,³⁵ y el interés de algunos resultados de las investigaciones que habían segui-

31 Dosse, François, *Histoire du Structuralisme, op. cit.*, t. 2, pp. 307 a 311, 321 a 326, y, en general todos los capítulos de esta obra concernientes a Foucault y la historia.

32 Como el interesante trabajo de Chartier, Roger, *Pratiques de la Lecture*, Paris, Rivages, 1985, por no citar más que el más representativo.

33 Le-Petit, Bernard, *Chemins de terre et voies d’eau. Réseaux de transports. Organisation de l’espace en France: 1740-1840*, Paris, Ed. EHESS, 1984.

34 Calvino, Italo, Paris, Seuil, 1981.

35 Lo cual ha venido a confirmarse con las obras ya citadas de Thomas Pavel y François Dosse, quienes detallan la manera en que se llevó a cabo dicha incursión lingüística.

do esta vía en otros dominios de la historiografía, particularmente en lo que se refiere a la historia sobre Francia. Fue por esto que nos pareció interesante probar este tipo de metodologías en el ámbito latinoamericanista.

Sin embargo, al adentrarnos más concretamente dentro del mundo de la lingüística, nos dimos cuenta que esta disciplina tenía sus propias problemáticas no resueltas, y que muchas de éstas estaban siendo simplemente transpuestas —sin resolver— al dominio de la historia. Ante la imposibilidad real —desde entonces reconocida por una gran parte de los investigadores—³⁶ de profundizar en los cuestionamientos de aquella disciplina sin abandonar completamente el campo historiográfico, nuestra utilización de sus metodologías trató de ser lo más neutra y moderada posible, tomando para ello como guía al grupo de Textos políticos y análisis de discurso, de la Escuela Normal Superior de Saint-Cloud (paquete *Lexicloud*), cuya proposición —con las pertinentes advertencias metodológicas que marca la lingüística en general—, expone sus métodos *en tanto que técnicas* y cuyo principio es basarse en el cuerpo analítico a partir del cual se piensa estudiar el discurso: en nuestro caso, la historia.

Es por esto que —y aunque no podamos extendernos aquí sobre los pormenores de las discusiones de las diferentes escuelas lingüísticas—, es necesario hacer un pequeño rodeo sobre la utilización de la lexicometría, tal como el grupo aquí mencionado la concibe y sobre la manera en que nosotros mismos la pusimos en práctica.

Maurice Tournier, uno de los principales animadores del *Lexicloud*, citando a Montaigne afirma que “*‘la palabra es mitad del que habla y mitad de aquel que escucha’*. *El lenguaje habla a través de nosotros y para el otro; un discurso cuyo dominio no lo tenemos más que a la mitad, pero cuya responsabilidad nos pertenece completamente*”. Así, las consecuencias de la palabra pueden ser medidas, tanto por aquello que *el locutor ha deseado o intentado expresar a través de sus palabras*, como por *la interpretación que hace aquel que lee o escucha dichas palabras*.

Como es bien sabido, todo intercambio de palabras entre dos interlocutores se realiza en un espacio delimitado social y culturalmente, y es dentro de una *gama lexical* más o menos amplia, más o me-

36 Dosse, François, *op. cit.*

nos especializada, más o menos “internacionalizada” o “regionalizada”, que dicho intercambio se llevará a cabo *con un mayor o menor margen de posibilidades de interpretación*. Aunque esto no implica que tales delimitaciones sociales y culturales no varíen en el tiempo y en el espacio.

Ahora bien, ¿cuál sería la diferencia existente entre aquello que se llama léxico y aquello que, más comúnmente llamamos vocabulario? Los lexicólogos contemporáneos parecen haber llegado a una cierta unanimidad al definir el *léxico* como “el conjunto de palabras por medio del cual los miembros de una comunidad lingüística se comunican entre ellos”, en tanto que el *vocabulario* sería “un subconjunto situado (o) especializado del léxico”.³⁷

Así, para J. Dubois existirían:

dos definiciones del sistema lexical: una paradigmática que concierne al léxico propiamente dicho, esto es, el conjunto de unidades lexicales que se basan en las costumbres de la comunidad y en las relaciones sociales que se expresan (a través de dichas unidades); la otra, sintágmica, por el vocabulario, es decir, por la distribución de esas unidades dentro de la frase.³⁸

Para aprehender las implicaciones del concepto de “delimitación cultural” en lingüística, pongamos un ejemplo que puede parecer anodino, pero que nos ofrece la dimensión de lo que esto significa. Una frase tan sencilla como “vivo en el último piso del edificio B”, no tiene el mismo valor si esto se pronuncia en la ciudad de México o en la de Nueva York. En el primer caso, generalmente no sucede que un “último piso” sea el piso 31 de un edificio, y aunque esto pudiera suceder, no es un caso usual. En tanto que en Nueva York, los trigésimos pisos representan una realidad más cotidiana, el saber si un elevador está o no descompuesto, pudiendo llegar a ser una legítima y “natural” causa de preocupación, ya que forma parte de una cotidianidad.

³⁷ Todas estas definiciones las encontramos en *Jornadas de lexicometría*, 13-14 abril 1988, en el artículo de Maurice Tournier, “*La Réduction Lexicométrique*”, p. 15 à 27.

³⁸ *Idem.* extrait de Dubois, J., *Le vocabulaire politique et social en France de 1869 à 1872*, Paris, Larousse, 1962, p. 188.

En cuanto a la diferencia entre un léxico y un vocabulario, pongamos el ejemplo de la frase “tengo que subir al último piso a buscar mis *floppys* para hacer una transferencia de ficheros sobre tu *Mc-Intosh*”. En esta frase, encontramos 19 palabras que, hoy por hoy, pertenecen al léxico de la comunidad hispano-hablante, aunque las palabras *floppy* y *Mc-Intosh* hayan sido últimamente adoptadas del inglés. Como quiera que sea, es evidente que estos dos términos hacen más específicamente referencia a un vocabulario de los dominios de la computación, sin presentar en sí ninguna ambigüedad. Tal no es el caso de la palabra “transferencia”, la cual puede pertenecer a otros tipos de vocabulario: psicológico —transferencia de personalidad—, financiera —transferencia de divisas—, etcétera. En este caso, no podemos dar un *sentido preciso* a la palabra “transferencia” si no recurrimos a su *contexto*, es decir, a las palabras que se encuentran alrededor de este término en un momento dado de un texto o discurso dado. Lo mismo sucede con las otras palabras de la frase expuesta (hacer, subir, ficheros).³⁹

Umberto Eco, en su obra *Lector in Fabula*⁴⁰ resume esta problemática de la siguiente manera: “...un locutor normal puede inferir, de una expresión aislada, su contexto lingüístico posible y sus circunstancias de enunciación posibles. Contexto y circunstancias son indispensables para poder conferir a la expresión su significación plena y completa, pero la expresión posee una significación virtual que permite al locutor adivinar su contexto” (las cursivas son nuestras).

Eco nos dice que, precisamente, es esta intuición la que:

dio origen a las *teorías textuales de segunda generación*.⁴¹ Éstas reconocen efectivamente que, para comprender un texto, se necesitan re-

³⁹ En general, cuando se habla de un estudio lingüístico, *todas* las palabras deben ser analizadas puesto que todas tienen un sentido por sus combinaciones dentro del discurso, inclusive la utilización de artículos, preposiciones, etcétera. Estas palabras, se denominan “palabras-útiles”. Es evidente que un discurso que presenta una mayoría de verbos conjugados en su forma imperativa, no tendrá la misma significación —aun histórica— que un discurso cuyos verbos están mayoritariamente conjugados en su forma condicional (cfr. introducción de Régine Robin a la obra *Langage et Idéologies*, J. Ghilhaumou, D. Maladier, A. Prost, R. Robin, ed. Ouvrières, 1974, p. 9). Sin embargo nuestro estudio no desarrolló este tipo de estudio, no por su falta de interés, sino por el tipo de aproximación histórica que nos propusimos: lanzarnos en un estudio más fino del lenguaje, nos hubiera alejado demasiado de los objetivos de este trabajo.

⁴⁰ Paris, Bernard-Grasset, 1985, p. 18.

⁴¹ *Idem*, p. 13: “En el desarrollo de las semióticas textuales, desde el principio se perfilaron dos tendencias. Las definiremos como las teorías de primera y de segunda generación, sin que

glas que no se limiten a una gramática del enunciado, pero, al mismo tiempo (estas teorías) no pretenden dejar de lado los resultados de un análisis semántico de los términos aislados. Al contrario: las teorías de la segunda generación pretenden construir —o postular— un análisis semántico que estudie los *términos aislados en tanto que sistemas de instrucción orientados hacia el texto*. Para ello (estas teorías) deben pasar... de un análisis en forma de diccionario a un análisis en forma de *tesaurus*.⁴²

Y es a este nivel que pensamos que las investigaciones iniciadas por el equipo de lexicometría política de la Escuela Normal Superior de Saint-Cloud se sitúan, ya que sus postulados toman en cuenta la interactividad que U. Eco nos describe:

1. Por la aceptación de un cierto “formalismo estructural”, es decir: el reconocimiento de una gramática y de un léxico pre-existentes —formados a partir de convenciones establecidas por una comunidad lingüística—, formalismo que es tomado en cuenta en el momento de elaboración de los programas informáticos más sofisticados del equipo de Saint-Cloud.

2. Puesto que el objetivo de estas investigaciones no es la creación de programas en sí y/o para la determinación de un diccionario fijo

esta distinción sea de tipo cronológico. La primera generación es, según nosotros, aquella que se mostró extremista y fuertemente polémica ante la lingüística de la frase (y más aún, del código); la segunda, es aquella que trató, al contrario, de llevar a cabo una hábil fusión entre las dos visiones, al construir *puentes entre un estudio de la lengua en tanto que sistema estructurado que precede las actualizaciones discursivas y un estudio de los discursos o de los textos como productos de una lengua que ya se habla* o, en todo caso, ‘para hablar’...” Es muy posible que la línea que representa Maurice Tournier dentro del equipo de Saint-Cloud, sea cercano de esta segunda generación cuya tendencia U. Eco nos describe. Así, M. Tournier, hablando de la *lexicología política estructural* nos dice que, al principio, este tipo de lexicología tendía a distinguir “discurso didáctico” de “discurso polémico” con base en su función. Sin embargo, argumenta este autor, tal distinción reduciría la problemática del discurso a una dicotomía entre un discurso didáctico que sería *estabilizador de empleos* (terminológicos, el ejemplo de esto serían los diccionarios), y un discurso polémico que se concebiría a sí mismo como *desestabilizador en su acción sobre el lenguaje*. M. Tournier se pregunta si en caso de que esta distinción fuese operativa, el discurso “didáctico” implicaría entonces “una neutralidad ideológica” que se encontraría por encima de todo embate, en tanto que el discurso “polémico” implicaría más bien “una toma de palabra partisana y militante”. “Ilusión óptica”, responde el mismo autor, ya que tanto uno como otro de estos discursos hacen política: el primero estructurando y “como leyendo las representaciones sociales dominantes” y, el segundo, al someter a rompimientos el *consensus* así formado por el diccionario. Tournier, *op. cit.*, p. 19-20.

42 Eco, U., *op. cit.*, p. 18.

y cerrado. Al contrario, con base en los datos de los resultados de estos programas, aplicados a un cierto tipo de textos —políticos, pero analizados desde un punto de vista sociológico, historiográfico, etcétera—, el objetivo es *observar los usos, las formas de enunciación, la evolución y las interacciones que han sufrido dichos usos y formas*, tanto de una formación a otra (política, por ejemplo), como en el transcurso del tiempo. Lo que se pretende pues obtener con estos programas informatizados, no es una “fijación”, una palabra “congelada”, sino un método que nos permita aprehender de manera sistemática, el movimiento de un tipo de lenguaje o de un vocabulario determinado; el cómo y el porqué del funcionamiento y de la evolución de éstos. En una palabra: las interacciones entre las estratificaciones del lenguaje y sus modificaciones.

En este sentido, Maurice Tournier toma también en cuenta el punto de vista de P. Guiraud y de C. Muller, quienes hablando de la diferencia que existiría entre “lenguaje virtual” y “discurso enunciado”, afirman que es importante hacer la distinción entre el léxico y un vocabulario. Según estos lingüistas, debe reservarse “el primero de estos términos al idioma y el segundo al discurso”, ya que un vocabulario puede ser calificado con cierto rigor, en tanto que con el léxico obtendremos, a lo sumo, una evaluación.⁴³ De la misma manera, para el equipo de la Escuela Normal Superior de Saint-Cloud, “*el vocabulario de un texto no constituye más que una ilustración del léxico implicado en una situación particular (dada), lo cual le prohíbe ser ipso facto ‘una muestra’ globalmente representativa*”.⁴⁴

¿Y qué otra cosa podría ser ese estudio de “términos ‘aislados’, en tanto que sistemas de instrucción orientados hacia el texto”?⁴⁵

En última instancia, los términos “aislados” serían los grafismos a partir de los cuales los programas informatizados van a efectuar el reconocimiento de cada una de las palabras. Estos grafismos representan, pues, una palabra —o un conjunto de palabras— de un texto cualquiera. Y son estas palabras que, tomadas también de forma aislada en el análisis de reconstrucción del texto, darán un sentido a

⁴³ Citado por Tournier, M., *op. cit.*, p. 19, nota 5; Muller, C., *Initiation à la statistique linguistique*, Paris, Larousse, 1968, p. 136, 140.

⁴⁴ *Ibidem*.

⁴⁵ *Cfr.* Umberto Eco, *supra*, nota 26.

éste último, tomando, a su vez, ellas mismas un sentido determinado por el propio texto y su historial.

Estas palabras, tomadas “aisladamente” y que pueden así ser analizadas a partir de diferentes ángulos del texto —en sus diferentes contextos dentro de un mismo texto—, son las que nos darán aquello que U. Eco llama “la instrucción orientada hacia el texto”, es decir, “una ilustración del léxico implicado en una situación particular...”⁴⁶ Una instrucción sobre aquello que esas palabras quieren decir *en ese texto* y lo que ese texto quiere decir en un contexto extralingüístico dado (es decir, fuera de toda consideración lingüística), pero, precisamente, debido a los sentidos que las palabras de aquel texto tomaron en él.

Y para esta interpretación, como el propio Umberto Eco lo dice, “se debe pasar... de un análisis en forma de diccionario a un análisis en forma de enciclopedia o *thesaurus*”, *thesaurus* cuyos datos de base son extraídos, en su versión lexicométrica, de los resultados de los programas de concordancias, co-ocurrencias, segmentos repetidos, etcétera, programas cuya elaboración respeta, en el establecimiento de sus parámetros, la gramática del idioma analizado.

Ahora bien, no exageremos nada. Por el momento esta gramática no ha podido ser completamente traducida en “algoritmo matemático”. Aun así, es necesario señalar que los términos matemáticos no deben atemorizar. De hecho, esta gramática siempre está presente, tanto en la mente de los investigadores que elaboran los programas, como en la de aquellos que operan una interpretación a partir de sus resultados. En este sentido, nos parece normal que, en tanto que lingüistas, el equipo de Saint-Cloud trate de llegar a un reconocimiento cada vez más fino de los segmentos repetidos o de las co-ocurrencias que tienen funciones gramaticales específicas.⁴⁷ En todo caso, tal como los programas funcionan hasta ahora y viendo la manera en que el equipo continúa ocupándose de dicho reconocimiento fino, podemos decir que los programas del *Lexicloud* no alteran en forma al-

⁴⁶ Cfr. Tournier, *supra*, nota 37.

⁴⁷ Lo cual sería importante, aun si no somos lingüistas, ya que el perfeccionamiento de ese tipo de reconocimiento permitiría, de una parte, el reagrupar términos o frases cuya función es estrictamente gramatical, permitiendo ocuparse de manera más directa de los términos que interesan nuestro tema específico. Por otro lado, este tipo de reconocimiento del texto, permitiría obtener una como “radiografía” más fina y menos dispersa del texto analizado.

guna el seguimiento sintáctico del texto, aunque su presentación parcial puede hacer pensar lo contrario.

Esto que acabamos de decir viene de nuevo a ser complementado por Umberto Eco,⁴⁸ quien hablando siempre de las teorías lingüísticas de análisis de texto, nos dice:

...En realidad, las dos teorías han demostrado que un texto posee propiedades que no pueden ser las mismas que las de una frase; las dos teorías admiten que la interpretación de un texto es también —y quizá esencialmente— debida a factores pragmáticos, y que, en consecuencia, un texto no puede ser abordado a partir de una gramática de la frase que funcionaría sobre bases puramente sintácticas y semánticas. En general, las teorías de primera generación consideran que los proto pseudos de una gramática de la frase es su límite lexical, de manera que ninguna teoría de orientación lexical puede explicar la significación de una frase dada como una pura agregación o amalgama de los significados lexicales previamente y definitivamente codificados.

Y U. Eco define dichos factores “pragmáticos” como la “dependencia esencial de la comunicación dentro del lenguaje natural del locutor y del auditor, del contexto lingüístico y del contexto extra-lingüístico...”⁴⁹

De la misma manera, cuando Maurice Tournier se pregunta sobre la pertinencia de medir un léxico a través de útiles informáticos, este autor considera dos soluciones: “... elevar las posibilidades del autómatas hasta la simulación de una competencia idiomática o disminuir el discurso al nivel de sus capacidades iniciales...”. Prácticamente, la primera de estas posibilidades necesitaría una transcripción previa de cada enunciado; además, se necesitaría recurrir:

⁴⁸ Habría que señalar que el equipo de lexicometría de la Escuela Normal Superior de Saint-Cloud y U. Eco, no trabajan juntos ni pertenecen a la misma “escuela”. Lo que aquí exponemos son las coincidencias de varios de los puntos de vista entre las dos formas de aproximación del texto: lexicométrica y semiológica, siendo este más bien un “diálogo” entre ambas maneras de proceder.

⁴⁹ Eco, *op. cit.*, p. 31. Para la definición de factores “pragmáticos”, Eco cita, Yehoshua Bar-Hillel, “Communication and argumentation in pragmatic languages”, 1968; in AA.VV., *Linguaggi nella società e nelle tecnica*, Milán, Comunità, 1970.

para cada desambiguación (de palabras, de expresiones), a una *multiplicidad de modelos lingüísticos y situacionales*, algunos de los cuales están lejos de estar establecidos (...), de manera que al cabo de dicho trabajo, ya no (sería) el texto original el que (sería) puesto en la máquina, sino un conjunto complejo de códigos sofisticados.

Coincidencia de visiones de nuevo. El *Lexicloud* opta por ir de lo simple a lo complejo, del texto real tal cual a las palabras que lo componen, y a partir de las cuales se “reconstruirá” el texto mediante el análisis de sus partes. Son estas partes que nos ayudarán a saber el rol que ellas mismas tienen dentro del texto, y no un formato previo (traducido en códigos) que establecerá dicho rol. Si estos resultados nos llevaran a establecer una “regla” sobre el texto, ésta sería dictada por el propio texto y sólo aplicable a dicho texto —después de su estudio a través de varias modalidades— y, sobre todo, dicha regla no sería propuesta de manera fija y definitiva. Pero con esto nos adentramos ya en el dominio interpretativo, el cual desarrollaremos en los siguientes capítulos.⁵⁰

Para resumir, digamos que si todos estos puntos, que ya han sido y siguen siendo discutidos por los especialistas del lenguaje, se plantean para el análisis de toda palabra (*floppy*, transferencia, hacer, “31o. piso”, etcétera), ¿qué sucederá cuando se trate de términos que son analizados desde una perspectiva histórica? Palabras cuyos sentidos representaron ciertas prácticas; prácticas que sin embargo, se han modificado, modificando a su vez el propio sentido de las palabras que nos las representan. Sentidos y cambios de sentidos, que con frecuencia no son ni legislados, ni codificados, ni oficializados —en los diccionarios, por ejemplo—, sino años después en que dichas modificaciones se han concretado en la práctica. Y, ¿qué sucede con aquellos otros términos que, sin haber encontrado un consenso de utilización o representación en un momento dado de su historia, caen en el olvido, para ser reutilizados años más tarde, cuando ciertos aspectos de la práctica que una vez trataron de representar —sin consenso— han cambiado en la realidad? ¿Cómo se les recupera? ¿Qué cara se les

⁵⁰ Hacemos referencia a la problemática: análisis-de-texto/discurso-analizado/discurso(s)-sobre-el-discurso-analizado, que algunos perciben en todo análisis, aunque no sea de discurso... con mayor razón cuando se trata de un análisis de discurso.

da? ¿Hablamos entonces de la misma palabra o del mismo término veinte años atrás que veinte años después?

Un ejemplo de este tipo de mutación es el término de “revolución”, que hace su aparición en el dominio de la física, para luego ser adoptado por la geografía y la astrofísica, pero cuya extensión semántica no llega al campo de lo social sino hasta fines del siglo XVII.⁵¹ Y aunque el término de “revolución” ha significado, de manera general, “cambio brusco” en el campo de lo social, también ahí tomó varios sentidos a través de su propia historia. Así, la “revolución” inglesa de 1643, es un término empleado *a posteriori* para identificar este evento, ya que este tipo de “actitud psicológica y de acción social” es más bien señalada por sus contemporáneos como “revuelta”.⁵² Lo mismo sucede con la “revolución” de 1789 en Francia, la cual, a pesar de ser señalada ya con este término, no tiene exactamente las mismas connotaciones que la revolución de 1848 o la de 1917 en Rusia. De la misma manera que 1989 en los países del Este, no es 1789 en Francia, ni 1917 en México... y aunque se pueden establecer ciertas analogías entre los contenidos de todas estas “revoluciones”, también se debe hacer hincapié en las particularidades que las diferencian.

De esta manera, a pesar de que este término continúa siendo utilizado en otros campos con sus propias connotaciones, y que la definición general en el campo de lo social significa “cambio brusco”, desde el punto de vista historiográfico, todas las modificaciones del término arriba mencionadas, deben ser analizadas de una forma más próxima, dilucidadas y diferenciadas de un periodo a otro, de una formación a otra, de la misma manera que se analizan aquellos términos que le son yuxtapuestos en los diferentes periodos de su existencia, como “tradición”, “comunidad” o “comunitario”, individuo, etcétera.

Así, se puede establecer un paralelo entre la historia de los términos y su simbólica, y la historia de las imágenes y su simbólica. A este respecto, las reflexiones de Maurice Agulhon sobre la historia de la representación republicana de la Mariana⁵³ es ilustrativa,

51 Alain Rey, “*Révolution*” *Histoire d'un mot*, ed. Gallimard, 1989. Cfr. pp. 54 a 64.

52 *Idem*, pp. 30-31.

53 “Marianne, réflexions sur une histoire”, *Annales Historiques de la Révolution Française*, 1992, núm. 3. Sobre los avatares detallados de la historia de Marianne, ver: t. 1: *Marianne au*

cuando este autor, después de un breve resumen de los avatares de dicha representación en Francia, sobre todo entre 1920 y 1930, nos dice: “...De ahí la profusión de imágenes contradictorias, de manera que el desciframiento del sentido de la *Marianne* en caricatura no depende ya tanto de los atributos que se le dibujan, como del contexto del episodio, de la leyenda y del mensaje global”.⁵⁴

De la misma manera, en el asunto que nos concierne, el discurso político y sus términos, el desciframiento de éstos, el análisis de sus sentidos en tanto que “sistema de instrucción orientado hacia el texto” —que nos da así el propio sentido del texto—, no depende únicamente de la definición que se les atribuye explícita o implícitamente en la retórica discursiva, sino también del contexto del episodio, del contenido de la leyenda que han acarreado consigo en los diferentes contextos en que se enfrentan (oficiales, oficiosos, oposicionistas o legalistas), lo que daría su mensaje más global, que es seguramente el que perdura por más tiempo en la memoria de la sociedad.

Las palabras acompañan la historia, por lo cual tienen su propia historia. Pero también pueden “hacer historia”. De ahí la atención que se debe dar tanto a su *enunciación* como al *análisis de su enunciación...* pero *sin olvidar las propias palabras que utilizaremos para comunicar los resultados de dichos análisis*. Así, por ejemplo y aunque no parezca más que un simple título, no es lo mismo decir en Francia “América Latina y la Revolución francesa”, “Repercusiones (o) Interacciones (etcétera) entre América Latina y la Revolución francesa”, que “L’Amérique Latine face à la Revolution Française” (América Latina frente (o ante) la Revolución francesa),⁵⁵ ya que la expresión “face à” en francés, puede fácilmente incurrir en una connotación de contraposición u oposicional.⁵⁶

combat (l'imprimerie et la symbolique républicaine de 1789 à 1880); t. 2: *Marianne au pouvoir (l'imagerie et la symbolique republicaines de 1880 à 1914)*; t. 3: en preparación. Paris, Flammarion, 1979 y 1989.

⁵⁴ *Idem*, p. 319 (lo cursivo es nuestro).

⁵⁵ Título de la exposición y del coloquio organizado en París I-Sorbona del 28 al 30 de junio de 1989 (AFSSAL), y esto a pesar de que, en la propia Sorbona, no era ésta la única concepción que prevalecía sobre la relación de ese evento histórico con América Latina, ni de hecho, la más consensual, *cfr.* las Actas del Coloquio *L'image de la Révolution Française*, París, Sorbona, 6-12 de julio 1989. Pergamon Presse, 4 vol.

⁵⁶ *Petit Robert*, t. 2, Dictionnaire de la Langue Française, París, 1985; p. 746: *Face à*: en faisant face à (- faire face: front, parer à, répondre à); vis-à-vis de (p. 2100), face à face, tête

En historiografía, a pesar de que ya se ha hecho evidente que muchos de los cuestionamientos que el historiador plantea a los documentos del pasado, se inspiran en su presente, nunca debe perder de vista que —sobre todo cuando se trata de análisis de discurso—, *es la historia de las palabras la que está tratando de abordar y no su propia historia presente (es decir, la del investigador), plasmada en aquellas palabras del pasado.*

Este señalamiento adquiere quizá una mayor pertinencia cuando, como en nuestro caso de análisis de discursos *políticos*, no se trata de palabras “neutras” o “frías”... Aunque no debemos olvidar que desde el momento en que un investigador se apasiona por un dominio cualquiera —no importa qué tan matemático éste sea—, las palabras que se eligen para representar dicha práctica, no son del todo inocentes ni frías.

No obstante, si convenimos en que el discurso político toca un dominio particularmente acalorado —“caliente”, como lo califica Maurice Tournier, es decir, efervescente—; un discurso que, como nos dice este mismo autor, “más particularmente que los otros trata de convencer” ocultando y disfrazando más cosas que otro tipo de discurso, para analizarlo necesitamos, pues, “desbaratar la trampa política. Mientras más ‘acalorado’ (*brûlant* - hirviente) sea un texto, más necesario es enfriarlo; mientras más apasionado sea un intercambio, su aproximación debe ser más objetiva, desimplicándose en relación a él...”⁵⁷

Es en este sentido que pensamos que junto con la tradición historiográfica, el método lexicométrico tal como es propuesto actualmente por el equipo de Saint-Cloud, puede ser una técnica eficaz para lograr dicho “enfriamiento”. Podemos aplicar cada una de estas técnicas o podemos utilizar sólo algunas; lo importante es que esta metodología quede, como hasta ahora, abierta a otras disciplinas, para que así el “formalismo” que nos señala Thomas Pavel, no llegue a ser un obstáculo para la investigación pluridisciplinaria, de la que tanto se habla actualmente a favor o en contra.⁵⁸

à tête). *Diccionario Larousse*, Espagnol-Français, 1986: vis-à-vis: enfrente de, frente a frente, persona colocada enfrente de o frente a otra, cara a cara, etcétera.

57 M. Tournier, *op. cit.*, p. 16.

58 “Propuestas para un ejercicio limitado de la interdisciplina”, por Bernard Lepetit, en revista *Iztapalapa*, núm. 26, México, jul-dic. 1992, artículo traducido por Carlos Aguirre-Rojas.

Así, a partir de nuestra propia experiencia, pensamos que ya que los puentes han sido tendidos, este intercambio debe continuarse, *respetando los límites y las exigencias de cada disciplina*. Al menos tomar en cuenta de manera prioritaria, los límites y exigencias de la disciplina a partir de la cual se piensa realizar la aproximación.⁵⁹ Pero esto no implica que se hagan a un lado otras posibilidades de investigación, ya que este enclaustramiento no llevaría más que a una estratificación de nuestra ciencia, y de aquellas posibilidades que, precisamente a fuerza de experimentos y tentativas, se han construido un camino de honor dentro de la historiografía francesa.⁶⁰

B. *La lexicometría y el análisis de un texto*

Mezclar las letras del Libro significa mezclar el mundo. Es un laberinto sin salida (*on ne s'en sort pas*). No importa qué libro, aun el abecedario... (p. 575, *Le Pendule de Foucault*).⁶¹

La lexicometría, tal como es practicada por el equipo de la Escuela Normal Superior de Saint-Cloud, va un poco más lejos de lo que su simple nombre indica: medida del léxico. Pero si pensamos un poco en las diferencias establecidas por los lingüistas entre lo que es un

de la *Revue de Synthèse*, 4a. serie, jul-sept. 1990, a quien agradecemos el habernos proporcionado estos textos.

⁵⁹ Si Michel Foucault dice no interesarle algunas de las exigencias que para un historiador son consideradas como prioritarias, tampoco es obligatorio que se le considere un historiador (*Cfr.* F. Dosse, *op. cit.*, pp. 322-323 a 326). Pero, inversamente: aceptando lo saludable que su *inspiración* (o ayuda) pueden resultar en ciertos momentos, que su referencia no se divinice en el ramo del historiador, permitiéndose puntualizar las deficiencias que su trabajo presenta desde el punto de vista historiográfico, sin que ello implique una ofensa al creador (*cfr.* Bernard Lepetit, "Propuestas para un...", es sobre la poca importancia acordada a el "grado" de exigencias prioritarias de una disciplina que diferiríamos de las proposiciones de este último autor, pp. 29-30).

⁶⁰ Ver, el artículo de François Dosse, traducido al español en la revista *Secuencia*, núm. 24, México, 1992, "La historia contemporánea en Francia". Esta traducción desafortunadamente omite la fecha de su primera aparición en Francia, en la revista *Espace Temps*.

⁶¹ Insistimos de que se trata de nuestra traducción de la versión francesa, cotejada con la española. Ver razones en nota 26.

léxico y lo que es un vocabulario, nos daremos cuenta que tal definición no se basta por sí misma.

Así, André Salem, en su introducción al artículo “Análisis factorial y lexicometría”⁶² nos dice que, ante el “discurso sobre los aspectos cuantitativos del vocabulario”, existen dos actitudes:

La primera actitud, resultado del análisis textual tradicional, consiste en recurrir a argumentos cuantitativos sólo para confortar opiniones *a priori* (...) puede suceder que las predicciones avanzadas de esta manera, recubran una realidad que es medible por medios cuantitativos. Pero en este caso, y en general, el lector dispone sólo de algunos ejemplos para poder forjarse una opinión... ejemplos que están destinados a probarle... que el texto que se estudia posee un vocabulario “rico” y “variado”, etcétera, sin que uno pueda saber a ciencia cierta si dichos adjetivos recubren un sentido estrictamente lírico, o si el autor está rindiendo cuentas sobre resultados de conteos que realmente se realizaron (...) La otra forma de proceder (la lexicometría) pretende ser científica. Es decir, que se propone *crear y sistematizar un conjunto de conocimientos*, de estudios con valor universal (concerniendo la totalidad del objeto que se considera), *caracterizados por un objeto y un método determinados, y fundados sobre relaciones objetivas, verificables*. Nacida de una profunda necesidad, por parte de los especialistas, de superar las aproximaciones tradicionales, frecuentemente juzgadas como subjetivas (la lexicometría) se propone aportar sobre los textos un esclarecimiento novedoso del texto, fundado sobre el conteo (y las deducciones que éste aporta) y sobre la localización de las formas que contiene...

Por su parte, otro integrante del equipo de Saint-Cloud, Pierre Lafon, afirma que:

la operación de indexación automática de un *corpus*, atribuye una frecuencia a cada una de las formas presentes en dicho *corpus*. Este resultado bruto, considerado aisladamente, esclarece (más) una lectura minuciosa del texto. Así, la atención del lector es atraída por los términos cuyas frecuencias resultan elevadas, y cuya repetición no había forzosamente llamado la atención durante una previa lectura. Para otros tér-

62 *Mots*, núm. 4, 1982, p. 147.

minos que, al contrario, se esperaban (con una frecuencia más elevada), son su escasez o ausencia lo que llama la atención. La jerarquía de usos y la comparación de frecuencias de empleo (de ciertos términos), dan ya lugar a ciertos comentarios. Hasta aquí, esta práctica esencialmente descriptiva, no es en absoluto del dominio estadístico...⁶³

Sin embargo, existen otros investigadores para quienes el simple conteo está también sujeto a sospechas:

Fui formada en un paradigma lingüístico, en el que la teoría de Chomsky y su concepción de la competencia innata de los hablantes nativos, juegan un rol predominante (...) Algunas de las proposiciones metodológicas del grupo francés...⁶⁴ no han alcanzado el nivel de desarrollo equivalente al cuerpo teórico de sus proposiciones. (...) En lo que concierne a la competencia lingüística: si se asume que todos los hablantes poseen una competencia innata (y precisa) para manejar las enunciaciones de sus propias lenguas, ¿por qué se deberá negar a los lingüistas la posibilidad de emplear esta habilidad en un sentido más amplio, como instrumento metodológico, controlado por otros dispositivos analíticos complementarios? (...) El laborioso nacimiento del análisis de discurso como una nueva disciplina de la lingüística, así como los esfuerzos del grupo inicial para adquirir un carácter científico reconocido, han contribuido a la adopción de un modelo rígido y tendenciosamente dogmático del análisis de discurso, en el que la eliminación, por parte del analista de varios, si no es que de todos los procesos analíticos, y su sustitución por procedimientos altamente formalizados, o preferentemente automatizados, parecen ser la principal garantía metodológica de objetividad (...) Me resolví, pues, por una aproximación manual lingüística del análisis de discurso. Esto significó, desde el punto de vista metodológico, *un uso sistemático de las habilidades descriptivas e interpretativas del propio analista*, de una forma artesanal...⁶⁵

63 Lafon, Pierre, "Sur la variabilité de la fréquence des formes dans un corpus", *Mots*, núm. 1, oct., 1980. pp. 127 a 165.

64 ¿A qué grupo se hace aquí referencia? Existieron y existen varios que se dedican al análisis de discurso, no sólo en París, sino en Francia.

65 Carbo, Teresa, Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS): "The silent pleasures of discourse analyses: Some remarks on a methodological alternative", *Sociolinguistics*, vol. XVII, num. 1, june 1987. pp. 3-4. (Traducido del inglés por nosotros.) Esta autora no es la única que deplora de forma tan vehemente la utilización de métodos "automatizados", si concebimos éstos como la organización sistematizada de las palabras de un texto, ejecutada por una computadora. La norteamericana, Estelle Manette Raben, del Departamento de Inglés del Queen College de Nueva York, critica —también de forma muy generali-

Ante tal encrucijada de opiniones especializadas concernientes a las metodologías cuantitativas para el análisis del discurso, nuestros cuestionamientos fueron los siguientes: ¿Dónde nos colocamos pues, nosotros los historiadores, simples utilizadores de dichas metodologías, ante un debate que en la época concernía sobre todo a los lingüistas, y de qué forma afectaba esto nuestro oficio de historiadores (*cfr.* capítulo precedente)? Al utilizar sólo la parte “descriptiva” (conteo y localización de formas), ¿utilizábamos de manera conveniente la lexicometría? Y, finalmente, si lanzándonos, de una forma u otra, por esta vía, ¿nos lanzábamos, al mismo tiempo, en la “adopción de un modelo rígido y tendenciosamente dogmático” del análisis del discurso?

Tratando de responder a estas preguntas, y ligándolo con lo ya establecido en el precedente subcapítulo, comenzaremos por explicitar la forma en la cual hemos concebido la utilización de la lexicometría en el cuadro de nuestra investigación; de qué manera pensamos que esta forma de utilización no conlleva una modelización rígida, dogmática o tendenciosa, para luego regresar sobre los aspectos prácticos—casi prosaicos— de la manera en que utilizamos este método.

Dentro del abanico de opiniones anteriormente expuesto, la definición que para nuestro caso y estudio nos parece más pertinente, es la ofrecida por André Salem. Es decir, que la lexicometría nace de una necesidad de dar nueva luz a textos que en ocasiones, habiendo sido estudiados ya por otros medios, necesitaban de formas de sistematización del conjunto de conocimientos que sobre ellos se tenían, sin que esto haya implicado una negación de los aportes que las aptitudes personales de otros analistas hayan podido aportar por otros medios.

Un texto literario es una cosa, su análisis le pertenece, lo mismo que su crítica. Un texto o un discurso político no son lo mismo, ya

zada— todo aquello que puede ser llamado “automatismo”, focalizando su atención en un programa denominado “*wordcrunch*”. Esta autora traza un paralelo entre este tipo de métodos y una suerte de “investigación cabalística”, agregando que ninguno de los resultados que estos programas nos proporcionan, “se relacionan con la comunicación en doble sentido (existente) entre el lector y el autor, esto es, el interés de la crítica literaria moderna...” (“*relates to the two-way communication between reader and author that is the concern of modern literary criticism...*” *cfr.* “*Can wordcruncher enable the user to ask significant questions about literary texts?*”, Cologne Computer Conference, 7 sept., 1988, p. 6. Ver también Dosse, F., t. II, *op. cit.*, pp. 16-29, para mayores detalles sobre el automatismo del análisis textual.

que los objetivos, y aun el público son diferentes, y para abordarlos pueden existir varios tipos de aproximaciones: político, historiográfico, estadístico y, ¿por qué no?, literario. Nosotros nos colocamos en el punto de vista de la historia política; en este sentido, un conocimiento sistematizado (es decir, que nos permita regresar sobre nuestros resultados, independientemente de nuestros estados de ánimo), nos parece bastante pertinente. Luego veremos de qué forma se puede regresar sobre dichos resultados, aunque dicha sistematización también pueda ser posible mediante otros métodos, *siempre y cuando éstos sean explícitos*.

En nuestro caso, habría que señalar que, al principio de la investigación, estábamos interesados sobre todo por la evolución de ciertos términos de orden político, tratando de saber a qué universo “mental” hacían referencia y, de forma general, saber si existió o no una evolución lexical del discurso político mexicano entre el siglo XIX y el XX.

Los otros estudios que podían efectuarse sobre un discurso, es decir: la dinámica de los enunciados, *la forma según la cual los universos políticos a los cuales se hacía referencia en los discursos* —y a los cuales nos interesábamos al principio de la investigación—, concretados en palabras y frases, *trazaban su camino dentro de aquel conjunto de discursos*, son temas que nos saltaron a la vista durante el estudio mismo de la metodología. Hablamos de los años 1984 a 1987.

Con base en este interés —evolución lexical o no—, todo el bagaje teórico del cálculo computarizado de los “segmentos repetidos”, del “análisis secuencial”, etcétera, lo dejamos en manos de los especialistas, quienes bastante hacen con prestar oídos a las exigencias de cada especialidad y a las particularidades de cada tipo de estudio, para tratar de introducir las, de manera pragmática, en la elaboración de sus programas: hablamos de debate, traducción algorítmica, y no de modelos.⁶⁶

66 Salem, André, en su obra, *Pratique des Segments Répétés. Essai de statistique textuelle* (pp. 9 a 21), nos explica muy claramente los avatares por los cuales han pasado, desde su inicio, los métodos de estadística textual, lo que quizá explica en parte la visión que Teresa Carbo tiene sobre el “grupo francés” (aunque, como lo señalamos, habría que identificar a qué grupo se refiere esta autora), en un momento dado de la evolución de estas metodologías. Efectivamente, André Salem, coincidiendo con lo dicho por Thomas Pavel (*cf. supra*, subcap. I.2.A,), sobre el hecho de que la estadística textual se ha inspirado de algunos modelos de la lingüística estructural, nos describe también los últimos debates que se han llevado a cabo al interior mismo de su equipo, a propósito de la manera en que debe efectuarse el reconocimiento por la computadora de los componentes de un texto, cada una de estas formas presenta sus propias

Es por esto que decimos que, finalmente, dimos nuestra confianza a los especialistas, sin que ésta haya sido una confianza “ciega”. Nos interesamos en sus debates, en sus problemáticas, tratando de ver cuáles eran los factores en juego y de qué manera podían afectar nuestro quehacer historiográfico. Lo cual es ya en sí, un enriquecimiento puesto que se trata de una elección sin *a priori*, en el terreno de las prácticas de investigación.

De esta manera pudimos constatar que los denominados “modelos”, presentan un gran margen de fiabilidad, la cual se debe no sólo a “una coherencia interna del formalismo” de los “modelos”, sino más bien a la coherencia de los resultados obtenidos en diversas y reiteradas aplicaciones en los diferentes campos que forman la investigación en ciencias sociales.

Dicho esto, nos parece que la desconfianza ante la “informatización del lenguaje”, toma magnitudes desproporcionadas, quizá un tanto similar a la generalización primera de la propia informática. Es verdad que la falta de costumbre de ciertas prácticas, pueden darnos una visión “abracadabrante” o demasiado difícil de algunos procedimientos. Sin embargo, una vez que la dificultad de la adquisición de la técnica es rebasada (sin olvidar las evoluciones que ésta ha sufrido, *cfr.* subcapítulo I.1), nuestra percepción se abstrae de ese sentimiento mágico ante lo desconocido.

La computadora realiza la identificación de “*formas simples*”, de grafismos simples para *el* reconocimiento de una palabra cualquiera. Con base en este primer reconocimiento, la computadora va a detectar, a lo largo de nuestro texto, las diferentes apariciones de dicha forma y va a darnos como resultado *el* número de veces que dicha forma aparece en nuestro texto. Así, en *el* presente párrafo la forma “*el*”

ventajas y desventajas, dependiendo del tipo de análisis que se desee realizar. Así, André Salem nos dice sobre esto: “Por nuestra parte, nos parece... indispensable que la cadena de tratamiento en su conjunto, sea el objeto de una *evaluación colectiva*, tomando en cuenta los objetivos particulares que se fijan los investigadores. Esta posición es igualmente válida para la toma de decisiones concierntes a la captura y segmentación del texto, los cuales deben ser también evaluados más allá de toda consideración *a priori* de orden lingüístico o filológico, en función (pues) de los resultados obtenidos y en relación a los objetivos que se persiguen.” Salem, A., *op. cit.*, p. 14. Cualquiera que sea el desenlace de esto, lo que sí es evidente es que dicho debate se desarrolla actualmente más bien sobre bases pragmáticas —y a partir de numerosas y diversas experiencias ya realizadas—, que sobre un formalismo estructural cualquiera; es decir, lo que U. Eco ha calificado como teorías de análisis de texto de la segunda generación (*cfr.*, subcap. I.2.A).

ha aparecido hasta ahora 3 veces (en cursiva). Se dirá pues que *la forma “el” tiene 3 ocurrencias en esta parte del texto.*

El *número de ocurrencias* es pues la *manera general de denominar* el número de veces en que un grafismo o forma cualquiera aparece dentro del texto, la *frecuencia de una palabra*, siendo pues, el *número total de ocurrencias de una sola forma o grafismo* dentro de la totalidad (*corpus* entero) o subtotalidad (una parte definida dentro del *corpus*) del *corpus* analizado. Por el contrario, *el número de formas* de un texto, es *el número de formas diferentes* detectadas en un mismo texto.

Así, en el párrafo anterior, entre los dos últimos signos de puntuación fuerte (los puntos seguido y aparte), tendríamos una frase de *21 ocurrencias*, pero con sólo *13 formas gráficas diferentes*: Por / el / contrario/ número / de / formas / un / texto / es / diferentes / detectadas / en / mismo /. De estas 13 formas diferentes: *Por*, tiene una ocurrencia; *el* tiene 3 ocurrencias, *contrario* = 1 ocurrencia; *número* = 2 ocurrencias; *de* = 3 ocurrencias; *formas* = 2 ocurrencias; *un* = 2 ocurrencias; *texto* = 2 ocurrencias; *es* = 2 ocurrencias; *diferentes* = 1 ocurrencia; *detectadas* = 1 ocurrencia; *en* = 1 ocurrencia; *mis-*
mo = 1 ocurrencia.⁶⁷

Es sobre la base de estos principios simples que los programas se desarrollarán, y la “complicación” del asunto —cálculo estadístico— comenzará por la manera en que los resultados serán organizados, y por las operaciones estadísticas que se efectuarán sobre la base de estos conteos simples, para obtener, por ejemplo, el análisis factorial de las palabras contenidas en nuestro texto.⁶⁸

⁶⁷ Habría que explicitar por qué se prefiere el término de “palabra” o simplemente “forma”, a lo que aquí utilizamos simplemente como explicación figurativa del asunto (forma gráfica). Por ejemplo, la palabra *de* en francés, aparece con esta representación gráfica antes de toda palabra que comience con consonante; cuando la palabra que este término precede comienza con vocal, la forma gráfica de esta preposición sera *d'*: *d'accord?/de plein gré*. Aquí tenemos pues una sola forma representada con dos grafismos diferentes. La computadora los reconocerá como formas gráficas diferentes, y es al analista, con sus aptitudes personales a quien corresponde, determinar y contabilizar las ocurrencias de ambos como una sola forma. El mismo problema subsiste en español con los *homógrafos* (u homónimos), es decir, aquellas palabra que escribiéndose de manera similar, tienen significados diferentes, para lo cual es necesario proceder a una *desambigüización* —manual— de los términos, ya sea previamente al tratamiento informático o, como en nuestro caso, después de realizado dicho tratamiento.

⁶⁸ Para el cálculo estadístico que implica el análisis factorial, ver, Salem, A. “Analyse factorielle et lexicométrie”, *Mots*, núm. 4, 1982, *op. cit.*, p. 150.

No es el momento de describir los programas que forman el lote del “Lexicloud”, ya que lo haremos en un subcapítulo aparte. Sólo quisiéramos aclarar que, es a partir del reconocimiento de estas formas “aisladas” (o simples) que tendremos nuestras listas de palabras: por orden decreciente o ascendente de sus frecuencias; con o sin una referencia que nos permita encontrar una palabra cualquiera en el o los lugares donde se encuentra en el texto; que existe otro programa —CONCOR—, que nos proporciona las palabras que rodean los términos que escogemos con el objeto de analizarlos de cerca —lo que se llama *palabras-clave*—, con sus referencias respectivas, de manera de que siempre podamos encontrar aquel grupo de palabras dentro del texto mismo. Con cantidades de información tan importantes, se entiende fácilmente el interés que representa el tener una especie de fichero que nos indicará inmediatamente, hacia donde debemos dirigirnos para encontrar la definición de tal o cual concepto en un texto de 800,000 palabras. Por ejemplo, se podrá localizar fácilmente y en cualquier momento, el término “democracia” y verificar cómo ha sido definido en el transcurso del tiempo por nuestros presidentes; en qué momentos y refiriéndose a qué lugares del país, se ha evocado el tema de la municipalidad; con relación a qué temáticas se ha hecho referencia a los Estados-Unidos de América, cuando en un discurso la frecuencia de esta forma aumenta de manera vertiginosa, etcétera.

Vistas ya sólo desde esta perspectiva, este tipo de programas presentan un interés considerable, al menos refiriéndonos a textos analizados desde un punto de vista historiográfico o politológico. Sin embargo, pensemos también en las atenuaciones, y/o confirmaciones, que pueden aportar a estas “simples observaciones”, la aplicación al *corpus* de los programas de cálculo más sofisticados.

En nuestro caso, el análisis factorial vino a comprobar todas las observaciones realizadas mediante el análisis de los resultados, de los más sencillos a los más complicados. Ahora sabemos, y tenemos la seguridad —gracias a este procedimiento, que por elección lo quisimos inductivo—, que podríamos perfectamente haber utilizado el análisis factorial como punto de partida de nuestra investigación, lo que puede ser la forma de proceder en una futura investigación de este tipo.

Y a pesar de que este último tipo de programas implique cálculos estadísticos más elaborados, *la base de reconocimiento sigue siendo*

la misma. La diferencia es que estos programas van a detectar no sólo un conjunto de caracteres entre dos espacios —el grafismo aislado—, sino un conjunto de grafismos.

Así, si los índices simples reconocían en la expresión “el Benermérito de las Américas” 5 formas simples diferentes, el programa de co-ocurrencias, reconocerá en este conjunto de 5 términos *una poliforma*,⁶⁹ al detectar que dichos términos son “recurrentes” en el texto, es decir, que aparecen sistemáticamente juntos a lo largo de un *corpus*, e informándonos del número de veces en que dicho conjunto aparece en ese *corpus*. La co-ocurrencias también pueden consistir en 2 o 3 palabras que se encuentren de manera sistemática en relativa proximidad, separadas sólo por algunas palabras.

Otro tipo de poliforma sería la detectada por el índice de “segmentos repetidos” (ISR), en el que se toman en cuenta las co-ocurrencias, pero que son identificadas a partir de caracteres “delimitadores”, es decir, que se encuentran entre signos de puntuación fuertes.⁷⁰

Por otro lado, habría que subrayar que no es porque el orden “natural” de las palabras de nuestro texto ha desaparecido en nuestros resultados lexicométricos, que se ha operado un divorcio entre el sentido del texto mismo y las palabras que lo componen, pero que han sido organizadas de otra manera (alfabética, por frecuencias, etcétera). No. Las palabras siguen ahí, en el lugar que las habíamos dejado cuando capturamos nuestra base de datos textual; con sus sintaxis, sus funciones gramaticales y sus sentidos, ambiguos o bien definidos.

Es como si leyéramos un texto y, que en lugar de escribir manualmente interminables fichas de lectura, estos programas nos dieran la posibilidad de elaborar un fichero de varias entradas, el cual nos permite volver en cualquier momento sobre una de las temáticas que nos interesa. Esta posibilidad, una vez capturados los textos, nos ofrece la facilidad de regresar a las diferentes temáticas del *corpus*, lo que

⁶⁹ La especialización de esta terminología ha resultado necesaria para facilitar, en ciertos casos, el desarrollo (y comunicación) del propio análisis.

⁷⁰ Signos que, al menos en este programa, pueden ser determinados por el propio investigador. Generalmente se aconseja utilizar como *Signos fuertes* . ¡ ! ¿ ? , siendo la “frase” el conjunto de ocurrencias que se encuentra entre dos delimitadores de secuencia (Salem, *op. cit.*, pp. 46-47). Estos delimitadores de secuencia comprenden los *signos de puntuación débiles* (; : - / ') . Pero, por ejemplo, en nuestro estudio los paréntesis fueron definidos como signos fuertes, ya que son utilizados como delimitadores de notas.

permite estimular y enriquecer las formas de abordar y de analizar una fuente cualquiera de información.⁷¹

Definida de esta manera, no vemos de qué manera la lexicometría puede ocultar una misteriosa “nomenclatura”, respaldo de un aparato dogmático, rígido o tendencioso. Que algunas veces —y en alguna época— este tipo de métodos hayan querido ser utilizados de esta manera, no significa que ese ha sido su desarrollo, ni que una utilización... ¿tramposa? haya sido la que ha predominado en la práctica. Actualmente se trabaja bajo el signo de la apertura, al mismo tiempo que del reconocimiento explícito de las reglas de cada disciplina, de tal manera que la lexicometría no pretende ser hoy más que un método y una herramienta de análisis entre muchos otros. Hoy por hoy es más bien la hora de los balances y no de las estigmatizaciones.

Que las “aptitudes” tengan mucho que ver con la importancia de los resultados e interpretaciones que se deducen sobre una temática cualquiera de estudio, es algo que nunca se ha puesto en duda, y esto en cualquier campo de investigación.

Ahora bien, si tal es el caso, ¿por qué oponer a las aptitudes personales, toda una serie de instrumentos que, en ciertos casos —e inclusive sólo a título experimental—, pueden ser de gran utilidad? Hay mil maneras de ir de Roma a Londres, simplemente ahora es posible realizar este viaje por barco, por avión o por tren (a través del Eurotúnel). Todo depende de los objetivos del viaje. Difícilmente tendríamos la tentación de efectuar nuestras lecturas de esparcimiento a través de programas lexicométricos.

Desde el punto de vista historiográfico, podemos decir que si la historia ha hecho de su propio análisis de discurso “tradicional” toda una metodología que la resguarda y fortifica, el experimentar nuevos métodos le ha aportado indudablemente una renovación de sus cuestionamientos ante el texto y su pasado (*cf.* subcapítulo I.2.A), lo que, al mismo tiempo, le está permitiendo abandonar —estudiándolo— el eterno movimiento pendular de argumentaciones que, con regularidad casi cíclica, vienen a retroalimentar ciertas discusiones desde hace 200 o 500 años.⁷²

71 Lo que puede ir desde el estudio de un concepto cualquiera, hasta la observación de formas “estilísticas” de aquel texto.

72 *Cfr.* Revolución francesa o Descubrimiento de América.

Sin embargo, a este respecto, es decir, al problema de las “interpretaciones sobre las interpretaciones”, cabría hacer ciertos señalamientos que, pertinentemente, Umberto Eco nos recuerda en su obra *Lector in fabula*, obra que, junto con *Los límites de la interpretación*, representa una profunda reflexión sobre su enorme experiencia en el mundo de la semiología:

... el Interpretante final no es final en el sentido cronológico. La semiosis muere a cada instante y renace de sus cenizas en su muerte. Las acciones individuales carecen de generalidad, *pero una serie de acciones, uniformemente repetidas, puede ser descrita en términos generales* (y citando a Pierce, Eco agrega): “¿Pero cómo se podría describir una costumbre, si no es a través de la descripción del tipo de acción a la cual (dicha costumbre) da nacimiento, con la especificación de las condiciones y de la motivación?” Así, una acción repetida que responde a un signo dado, deviene a su vez un nuevo signo, el *representamen* de una ley que interpreta el primer signo y da nacimiento a un nuevo proceso infinito de interpretación...⁷³

¿Por qué la historiografía escaparía de este movimiento?, sobre todo cuando se sabe que lo esencial de su actividad ha reposado, casi desde sus inicios, en la interpretación de textos y en el cotejo de los mismos. Integrar nuevos métodos, como la lexicometría, le permite pues, observar de otra manera no sólo sus fuentes primarias, sino también su propia producción en el transcurso del tiempo, lo que, como ya se dijo, le ha permitido también examinar constantemente sus propias problemáticas interpretativas.⁷⁴

Por otra parte, ¿qué otra cosa podría ser una metodología si no es esa “serie de acciones, uniformemente repetidas” y que pueden ser descritas en términos generales? Dicho de otra manera: *acciones y operaciones que pueden ser repetidas y descritas* y que, por lo mismo, *pueden ser transmitidas*.

Es en este sentido que entra en escena el concepto de *práctica metodológica*; es decir, un procedimiento que nos induce, de forma general, a ciertas conclusiones, pero que puede ser recorrido no sólo por

⁷³ Eco, U., *op. cit.*, p. 56.

⁷⁴ Este tipo de estudios ha comenzado a desarrollarse en el cuadro del equipo de “Historia de las Ciencias y de las Doctrinas” del CNRS, *cf.* también: Salem, A., *Pratique des...*, *op. cit.*, nota 1, p. 55, y Dosse, F., “Histoire du Structuralisme”, *cit.*

la persona que lo ha creado, sino por otra persona que, poseyendo no sólo una aptitud para ciertos tipos de análisis, posee también instrumentos de trabajo que le permiten transmitir su propio saber.⁷⁵

Ahora bien, sabiendo precisamente que las metodologías también están sujetas al movimiento semiótico que Umberto Eco nos describe, no es ilegítimo buscar vías que, siendo transmisibles, nos permitan al mismo tiempo, establecer cierta distancia entre el investigador que somos y nuestra propia producción; entre el investigador y aquello que el propio medio en el que evolucionamos nos enseña como el método. El tiempo dirá si los métodos de hoy serán olvidados o criticados en el futuro...

En todo caso, es un hecho que la historiografía, contando ya con técnicas de trabajo y métodos de investigación suficientemente afianzados en el gremio de historiadores, ha podido integrar nuevas técnicas a sus “antiguas” costumbres (las cuales, por su parte, no han sido hechas a un lado), integración que, a su vez, ha creado nuevas prácticas y costumbres en el oficio de historiador, prácticas que, por su sola experimentación crítica, están siendo afianzadas a su vez.

Un tanto independientemente de lo anterior, pero refiriéndonos siempre al aspecto metodológico de las interpretaciones textuales, pensamos que la advertencia que Umberto Eco hace a propósito de la dificultad semiótica de interpretación de una palabra fuera de su contexto y sin ciertas “reglas”, puede aplicarse también a las formas de interpretación de la historiografía en general, de un texto, de un término, ¡vaya!, de todo aquello que tiene algo que ver con el discurso:

... sólo quedaría el recurso, más bien decepcionante, a la intuición del locutor —categoría desagradable a la cual ninguna teoría semiótica sería debiera recurrir, ya que, si una teoría semiótica tiene un objetivo, éste es precisamente el *explicar cómo funciona la intuición del locutor y de explicarlo en términos no-intuitivos*.⁷⁶

⁷⁵ Cfr. Carbo, Teresa, *Discurso político. Lectura y análisis*, Cuadernos de la Casa Chata, núm. 105, Centro de Investigaciones y Estudios Superiores en Antropología Social. En este trabajo, la autora, hace referencia, para la interpretación de textos, a las “condiciones sociales de producción del texto y de la utilización verbal” empleada en el texto, lo cual parece apoyarse en las propias ciencias históricas de interpretación textual, y con lo que estamos completamente de acuerdo.

⁷⁶ Eco, U., *op. cit.*, p. 24.

Y es desde este punto de vista que la lexicometría puede constituir un excelente instrumento de análisis de textos, puesto que su utilización ayuda, no sólo a crear esa distancia necesaria entre el analista y su objeto de estudio —el cual, siendo político, presenta mayores ambigüedades—, sino que también nos ayuda con procedimientos objetivos.

Efectivamente, los programas de lexicometría nos permiten “desmontar” de manera transitoria el texto, con el objeto de efectuar una reorganización sistemática de sus componentes —palabras, formas gramaticales, adjetivos, etcétera—. Esta reorganización, nos proporciona una visión suplementaria y, frecuentemente, diferente a la lectura de un texto leído en su estado original. Sin embargo, después de haber observado el texto desde otras perspectivas, se tiene que regresar al texto original para restituir las modalidades que saltaron a la vista durante el análisis de los diversas formas de reorganización de sus componentes. Del cotejo de estos resultados —con las conclusiones de nuestra primera lectura, con nuestras intuiciones o con comentarios de otros analistas—, se puede llegar a aportar nuevas conclusiones no observadas por otros medios; a aportar ciertos matices o, simplemente, a corroborar, y hasta refutar algunas de nuestras primeras hipótesis.⁷⁷

Después de esto, se pasa al análisis tradicional del texto, es decir, a su confrontación global —como unidades específicas— con el contexto que las pudo haber producido, pero armados de una mayor información sobre los componentes de aquella unidad que es el texto. Método tradicional de la historiografía, pero que nos ayuda a darnos cuenta —sobre todo si uno se lanza en el estudio a profundidad de un pequeño conjunto de textos—, si ciertos eventos de dicho contexto “de producción”, son causa o consecuencia del propio texto. Las palabras tienen su historia, pero recordemos que también pueden llegar a hacerla.

Tratemos, pues, más bien de avanzar en un campo cuyos frutos comienzan apenas a ser repensados, después de años de arduos y serios experimentos, sin escondernos detrás del temor a las palabras.

77 Lo que fue el caso de nuestra investigación sobre la Convención Revolucionaria de Aguascalientes.