

## INTELIGENCIA ARTIFICIAL EN FILOLOGÍA

Agradezco a los estudiantes organizadores del ciclo *Humanidades: metodología para los '90'* el haberme invitado a dar mi opinión sobre este asunto, de un modo informal, que sirva para el diálogo, el intercambio de pareceres y una toma de postura intelectual ante un evento futuro, que se desconoce cómo se desarrollará como es el de la inteligencia artificial.

El curso responde a una inquietud de los estudiantes de Humanidades por el porvenir de esta rama del saber. Además, en un momento en el que prevalece en la política del Estado el fomentar las denominadas ciencias duras y eliminar o aparcas las ciencias del espíritu. El avance de la tecnología parece querer ocultar la importancia del pensamiento libre del hombre para sacarle jugo inmediato al lado pragmático de la ciencia. Es curioso que sean los hombres de laboratorio —entre ellos Severo Ochoa— los que aportan más luz, en esta hora, para dar sentido a una vida colectiva deshumanizada por la tecnología y piden una formación general del individuo, basada en el estudio de las lenguas muertas y en materias filológicas e históricas. Ramón y Cajal, sobre las reformas de los planes de estudio de su tiempo, opinaba que era tarea inútil y proponía que se debía:

«Transformar la enseñanza científica, literaria o industrial, no aumentando, como está ahora de moda, el número de asignaturas, sino enseñando de verdad y prácticamente las que tenemos. Bajo este aspecto habría que decir de nosotros cosas

---

1. Cáceres, del 5 al 7 de febrero de 1990.

atrocés. La media ciencia es, sin disputa, una de las causas más poderosas de nuestra ruina»<sup>2</sup>.

Los cultivadores de las ciencias humanísticas deben —debemos— de dar un toque de alarma que frene el exceso de tecnicismo con el que nacerá el hombre del siglo XXI.

Los pensadores del siglo XIX estuvieron atentos para descubrir y analizar nuevas facetas de la vida humana. El investigador de disciplinas humanísticas deberá atender, en la década de los noventa, a los avances de la técnica para progresar en sus estudios.

En Francia e Italia ya hace años que, con el apoyo del Estado, se aplican las máquinas para adelantar en la investigación de la filología. En España, como siempre, y a pesar de pagar el coste europeo, vamos a la zaga.

Por ello creo que el asunto de esta conferencia es de suma importancia para intuir la evolución de la ciencia en la última década del siglo XX. Mi postura ante el tema de la inteligencia artificial es de meditación y reflexión desde la perspectiva del filólogo, del hombre de letras, que movido por los avances tecnológicos se ve obligado a abandonar el acarreo tradicional de fichas y papeletas para ordenar los datos de su investigación.

El primer problema que se plantea es el de definir el concepto de inteligencia artificial.

### 1. *Concepto de inteligencia artificial*

La inteligencia artificial ha surgido como consecuencia y en oposición al concepto de inteligencia humana. Dicho con palabras sencillas, la inteligencia artificial se ocupa de la manera de «pensar» y de «razonar» de una máquina. En este sentido el desarrollo de la Cibernética ha hecho progresar a las restantes ciencias. Se tendría que admitir la capacidad de las computadoras para realizar las mismas operaciones que realiza la mente humana. Si el pensamiento del hombre no es perfecto para entender el mundo que le rodea —véanse las distintas concepciones filosóficas de la historia de la humanidad—, el lenguaje de las máquinas estará, evidentemente, lleno de imperfecciones.

En la última década<sup>3</sup>, significativos progresos en inteligencia artificial han conducido a los llamados sistemas expertos que se definen como:

«...un sistema informático que simula el razonamiento de un experto humano que, de este modo, ayuda a otros expertos humanos a resolver problemas o tomar decisiones en dominio limitado pero difícil del mundo real»<sup>4</sup>.

Una idea que está en la mente de todos los investigadores es la de que un sistema experto no puede ni debe reemplazar a los expertos humanos.

2. García Durán Muñoz y Francisco Alonso Burón, *Cajal. Vida y obras*, 2.<sup>a</sup> edición, Barcelona, Dossat, 1983, t. II, 258-259.

3. Ramón López de Mántaras, «Inteligencia artificial y sistemas expertos: ¿la alquimia del siglo XX?», *Ciencias Sociales*, 1989, 4, 242-245.

4. *Op. cit.*, p. 242. En esta exposición seguiré la terminología de este autor.

Otra faceta evidente es que para usar un sistema experto es necesario ser bastante experto.

No todo son inconvenientes en un sistema experto. Este constituye una herramienta muy útil para ayudar al usuario a evitar que pase por alto aspectos importantes de la tarea mecánica que tiene que desarrollar. Por ejemplo en la ordenación alfabética de una lista de alumnos, el experto humano puede cometer fácilmente errores, en cambio, el sistema experto salva enseguida dichos errores.

Los sistemas expertos, al poseer una limitada forma de autoconocimiento, pueden explicar y justificar parcialmente sus decisiones y sus preguntas a los usuarios.

El circuito de un sistema experto es aparentemente muy simple. En el lenguaje de un sistema experto intervienen dos factores: el emisor y el receptor. El emisor es el especialista en una determinada disciplina (médico, matemático, filólogo, historiador, etc.) y el receptor es el ingeniero del conocimiento. Estos dos elementos crean un sistema experto que puede ser utilizado por posibles compradores o usuarios. Se reconoce que en los últimos años ha habido una cierta trivialización por parte de informáticos, que, con fines lucrativos, vendían pseudosistemas expertos a un público escasamente conocedor del asunto.

El ingeniero del conocimiento suele ser un informático que en un tiempo muy corto debe familiarizarse con el lenguaje experto. Este hecho aumenta la posibilidad de malas interpretaciones y, por tanto, de errores.

La mayoría de los sistemas expertos actuales representan el conocimiento superficial, no el conocimiento profundo o causal. Por tanto, solamente las frases denotativas son representables. Los valores de la ironía y el lenguaje poético y metafórico no hay ningún sistema experto capaz de representarlos por ahora. La connotación de un enunciado escapa a la inteligencia artificial.

En la elaboración de los sistemas expertos, un emisor (experto en una materia concreta) proporciona una información que el receptor (ingeniero del conocimiento) debe interpretar y modelizar mediante un esquema ejecutable por el ordenador, evitando los errores y las ambigüedades que se pueden generar.

En este aspecto sigue vigente la idea de Ramón y Cajal que sostiene que:

«Hay que crear ciencia original en todos los órdenes del pensamiento: filosofía, matemática, química, física, biología, sociología, etc. Tras la ciencia original vendrá la aplicación industrial de los principios científicos, pues siempre brota al lado del hecho nuevo la explotación del mismo, es decir, la aplicación al aumento y la comodidad de la vida»<sup>5</sup>.

Es evidente que en el ámbito de las disciplinas humanísticas se deberá crear primero ciencia original para poderla transformar posteriormente en ciencia artificial que facilite su aplicación para mejorar el bienestar social. Ciertamente, la inteligencia artificial será una ayuda imprescindible para el progreso de la ciencia en el futuro.

Otro elemento que complica la creación de sistemas expertos es la comunicación no del todo perfecta entre el emisor, experto especialista, y el usuario. Ello genera incertidumbre en la transmisión del conocimiento. Uno de los aspectos

---

5. *Op. cit.*, en nuestra nota número 2, p. 259.

tos fundamentales de este problema es identificar los criterios que permitan emitir algún tipo de juicio acerca de si un sistema experto lleva o no a cabo un buen razonamiento, un buena explicación o una buena interacción con el usuario.

## 2. *Aplicaciones de la inteligencia artificial en filología*

Me voy a detener, brevemente, en repasar la historia de la aplicación de la inteligencia artificial a los estudios de filología durante el siglo XX.

Roberto Busa, en la mitad de la década de los años cuarenta, utilizó las computadoras en el análisis lingüístico y filológico de las obras de Santo Tomás de Aquino. Resultado de su investigación son los cuarenta y cinco tomos del *Index Tomisticus*<sup>6</sup>.

J. W. Ellison publica en 1957 una concordancia de la Biblia, generada gracias a una computadora electrónica.

Más de medio millar de referencias bibliográficas se pueden agrupar sobre el uso de las computadoras en la investigación literaria y lingüística en textos españoles. Los primeros trabajos sobre esta lengua aparecieron en la década de los años sesenta. A. Juilland y E. Chang-Rodríguez dieron a conocer, en 1964, un léxico lematizado con indicación de frecuencias, basado en una muestra de 550.000 palabras-textuales.

En la Universidad de Madison se desarrolla el proyecto DOSL (*Dictionary of Old Spanish Language*) y en México el programa DEM (*Diccionario del Español de México*), patrocinado por el Colegio de México.

En la ciudad de Pisa funciona el *Instituto de Lingüística Computacional*, que se ha convertido en promotor de la actividad de este campo.

Investigaciones individuales o pequeños grupos trabajan en las más variadas áreas en Lovaina, París, Bonn, Sevilla, Barcelona, Götteburg, México, Nueva York, Quèbec, Pisa.

Esta primera etapa se ha caracterizado por ser necesario el uso de las viejas y voluminosas computadoras mecánicas y electrónicas, que obligaban al filólogo a estar relacionado con un centro de cálculo universitario o con una empresa privada importante.

En este período la lingüística automática ha iniciado el empleo de las máquinas para adelantar paulatinamente en los dominios y áreas tradicionales en los que suele dividirse la filología. En principio, los lingüistas se han mostrado recelosos a la intromisión de la informática en sus estudios y han tenido que pasar más de veinte años para que sus frutos fueran reconocidos.

Los filólogos se han planteado el problema teórico del uso de las máquinas y cómo realizar los programas informáticos adecuados. Ha sido una tarea costosa y de pioneros. En Madrid, el Centro de Cálculo de la Universidad Complutense ha apoyado iniciativas atrayentes. En el terreno de los problemas generales de la adaptación de los ordenadores a la lingüística se deben tener en cuenta aquellos

6. Leopoldo Sáez-Godoy, *BICC*, XXXVIII, 1983, 340-375, la cita en la p. 340.

trabajos que establecen relación entre la informática y la filología y también aquellos que se preocupan por generar un lenguaje artificial, a imitación del humano.

El empleo de la máquina no debe asustar al filólogo, porque con ella se consigue eficacia, rapidez y economía<sup>7</sup>. La eficacia permite que en la labor mecánica no se pase por alto ningún elemento, la rapidez es nota sobresaliente y la economía resalta en proyectos de gran envergadura.

Una de las facetas que más asustan al filólogo es la lentitud en la preparación de un programa y la dependencia de un informático para ejecutar la investigación.

Los trabajos de lingüística computacional que se han realizado hasta ahora se pueden clasificar en análisis lingüístico por niveles del lenguaje, ediciones de textos, aprendizaje del idioma y geografía lingüística.

En el análisis lingüístico propiamente dicho se ha avanzado en estudios de ortografía y grafemas, fonética y fonología, morfología y sintaxis, y lexicografía y lexicología.

De capital importancia ha sido el uso de las computadoras en el examen de textos literarios, en temas relacionados con la métrica (verso, rimas y concordancias), ediciones críticas, estilística y en la atribución de la autoría de un documento o paternidad literaria. En este último apartado se debe citar a los investigadores franceses, reunidos en torno a la revista *Cahiers de Linguistique Hispanique Médiévale*. René Pellen, profesor de la Universidad de Poitiers, ha trabajado diversos aspectos lingüísticos del *Poema de Mio Cid* y de las obras de Gonzalo de Berceo. Jean Roudil, profesor de la Universidad de la Sorbona, se ha ocupado de la familia de los fueros de Cuenca y de diversos libros medievales de carácter jurídico.

En cambio, son escasos los títulos dedicados al aspecto práctico de la enseñanza del español y a la traducción de esta lengua a otros dominios.

Digno de ser anotado es el aprovechamiento de la informática en la elaboración y almacenaje de los datos recogidos por la geografía y cartografía lingüísticas.

### 3. *Mirada hacia el año dos mil*

La inteligencia artificial en la década de los noventa puede dar muchas sorpresas. El avance en la investigación de programas nuevos y masivos respecto al número de usuarios puede poner al alcance del filólogo una sofisticada herramienta que deberá saber manejar para poder acceder al progreso en sus disciplinas.

Las viejas y antiestéticas computadoras se han transformado ya en modelo del más avanzado diseño italiano. Los establecimientos que hasta ahora vendían lavadoras y televisores se han convertido en promotores de una variada serie de material informático. El estudiante medio en esta década va a disponer de un compatible en su casa. Si este tipo electrodoméstico invade los hogares, la filología tampoco se verá privada de su presencia. El abaratamiento del producto socializará su posesión, pero como acontece con otros electrodomésticos, no todo el mundo

---

7. Manuel Alvar, «Informática y Lingüística», *RSEL*, VII, 1977, 192-212.

racionalizará su ayuda. El estudiante de ciencias humanísticas se verá irremediablemente inmerso en la órbita de la máquina.

Voy a poner orden a una serie de ideas sobre el rendimiento de la informática en el estudio de la filología.

Confieso que soy un neoconverso al posible disfrute activo de la máquina en la investigación y que, por consiguiente, me encuentro en las primeras lecciones del manual.

Mi contacto con la máquina se produce en 1972 cuando visité y practiqué, como alumno, en el Laboratorio de Fonética del Consejo Superior de Investigaciones Científicas de Madrid. En 1987 se consiguió montar el Laboratorio de Fonética de la Universidad de Extremadura. Su utilización ha permitido que las últimas promociones de alumnos hayan visto generarse un espectrograma a partir de su propia voz.

En estos años he repetido que la máquina deshumaniza al hombre de letras y he notado la falta de un equipo interdisciplinar, sueño dorado inalcanzable en la universidad española. Pero esa desazón intelectual, en este último año, se ha convertido en agrado, al comprobar que los fabricantes de aparatos han llenado el mercado de productos que se consumen, simultáneamente, en Estados Unidos, Europa y Japón, y responden, en grandes líneas, a las necesidades de la investigación que se llevará a cabo en la década de los noventa.

El estudiante, el profesor y el investigador se podrán comportar ante la inteligencia artificial con dos tipos de actitudes complementarias. En la primera como sujetos pasivos, serán consumidores de productos y programas. Pronto figurarán en los catálogos —ya figuran algunos— programas que desarrollen el contenido de variadas asignaturas. Desde esta perspectiva el usuario se mostrará como un consumidor del producto informático.

En el segundo tipo de actitud el usuario será elemento activo de esa inteligencia artificial. Inicialmente tendrá que aprender el funcionamiento del sistema experto para saber manejar, con el máximo rendimiento, cada producto. Y posteriormente acomodar los programas en beneficio de la investigación.

Se plantea un problema ético. La máquina dará más posibilidades de llegar a la información sobre una materia, pero ¿quién podrá tener acceso a ella? Creo que el más poderoso y el más inteligente y experto. Por ello, la universidad, el Estado, debería de poner medios para que los ciudadanos tuvieran acceso a la información como servicio público. Y esta vertiente nos lleva a la cuestión de la formación de los usuarios —estudiantes, profesores e investigadores— en las nuevas tecnologías.

En el año dos mil deseamos que, en los centros universitarios y de investigación del país, las bibliotecas estén informatizadas y unidas a los principales puntos de información del mundo. ¿Cómo se utilizará la informática en la docencia y en la investigación? Como este país es diferente, profesores y alumnos habrán aprendido, autodidactamente, los programas imprescindibles para estar al día. Veo, entre la niebla, el fantasma del dinero, del coste económico, que será un impedimento para destinar exhaustivamente estos medios a la enseñanza. Y, además la masificación del alumnado universitario será otro factor negativo. El sector privado obtendrá, sin duda, los mejores resultados, a no ser que el Estado decida, de

una vez, alejado de la demagogia habitual, convertir la universidad pública en verdadera universidad y no en aparcamiento de las jóvenes generaciones, y dotarla de los medios necesarios para el aprendizaje y la investigación que la década de los noventa se merece.

El alumno constante y motivado —egresado de la masa amorfa— podrá formarse de una manera más cómoda gracias, por lo mínimo, al uso de un programa de tratamiento de textos —que sustituirá las funciones de la anticuada máquina de escribir—, de una base de datos —donde archivará sus fichas— y de un paquete de gráficos, con el que ilustrará sus trabajos.

En filología se adelantará muchísimo en aspectos de recuento de datos y en el análisis de la voz humana. Las ediciones críticas se harán con mayor facilidad, merced a la reproducción por láser de los textos y a la rapidez para generar índices y concordancias. La inteligencia artificial permitirá que la traducción de un idioma a otro sea más llevadera, gracias a los copiosos bancos de datos de léxico de todas las lenguas.

Al hombre de letras le quedará una importante misión que cumplir cual es la de humanizar los aspectos mecanicistas de esta nueva inteligencia artificial de acero, sin sentimiento. No creo que la máquina pueda llegar a pensar; en todo caso, será una mera repetidora del pensamiento humano, pero no será capaz de expresar libremente unas ideas y unos conocimientos. Lo último de que vienen dotados los programas de proceso de textos es de la ortografía y del Thesaurus de sinónimos y antónimos. Sin el hombre, la máquina no es apta para distinguir si una palabra debe ir o no acentuada. Con la ayuda del diccionario de sinónimos y antónimos es el usuario quien decide qué voz debe ser sustituida por otra.

En el aspecto de generar textos la máquina ayudará, de hecho ya ayuda, al periodista y al escritor. Es como si la máquina tuviera metido en su interior un excelente corrector de pruebas, el manual de ortografía de Gili Gaya y una parte considerable del Diccionario ideológico de Julio Casares. El corpus que tiene es pequeño, pero ampliable por el usuario, la creatividad del individuo humano no se destruye, sino todo lo contrario, se potencia, con lo cual, con el uso de la informática, el inteligente será más poderoso, porque sabrá sacarle mejores rendimientos, y el inexperto no progresará como todos deseáramos.

La informática es un elemento más de poder. Un reciente anuncio publicitario afirma que, con una determinada marca de microordenadores, «Las secretarías ya tienen secretaria», lo que quiere decir que el jefe de esas secretarías suma a la inteligencia humana de sus empleadas la inteligencia artificial de la máquina.

Se puede llegar a un acentuamiento del individualismo y prepotencia de unos pocos que dominarán a los más indigentes y necesitados. Un instrumento que podría servir para que todos los hombres fuéramos más libres, más iguales y más cultos, se puede convertir en alejamiento entre sectores sociales y en dominio de los poderosos sobre la inmensa mayoría.

ANTONIO VIUDAS CAMARASA