



LAS INDUSTRIAS DEL LENGUAJE: UN CAMPO DE EXPANSIÓN IMPORTANTE EN LA SOCIEDAD DE HOY

M. Teresa Cabré Castellví*

El objetivo básico que se propone este texto es el de presentar el campo impreciso de las llamadas industrias del lenguaje situándolas en un marco definido, el de la lingüística aplicada, y navegar en este campo reflexionando sobre el cómo y el porqué de su aparición y de su creciente importancia, y sobre el objeto que legitima su especificidad y que poco a poco va perfilando sus contornos.

1. LINGÜÍSTICA TEÓRICA Y LINGÜÍSTICA APLICADA

Objetivamente hablando, entre la teoría lingüística y la lingüística aplicada no existe una oposición justificada, a pesar de los múltiples prejuicios y recelos que han existido y todavía existen hoy en día.

Estos recelos y desconfianzas provienen de ambos colectivos. Por una parte, los lingüistas teóricos desconfían de la preparación lingüística fundamental de los aplicados, a los que ni siquiera consideran lingüistas sino profesores de lenguas, terminólogos, traductores, lexicógrafos, correctores, redactores, etc. Por otro lado, los aplicados desconfían de que los avances de la teoría lingüística puedan hacer progresar el tratamiento del lenguaje para resolver necesidades sociales relacionadas con la comunicación y la información. Para los lingüistas aplicados muy a menudo la noción de «aplicado» se reduce a la pura aplicación.

Un gran número de prejuicios han actuado (y continúan actuando hoy aunque cada vez más tenués) sobre las actitudes de ambos colectivos profesionales de lingüistas. Para los teóricos, la teoría lingüística es la lingüística, y el resto pura aplicación de principios alejados de ella. Para los aplicados, la lingüística es pura especulación, socialmente inútil, y científicamente improductiva.

Y pese a esta oposición inútil, en los últimos tiempos el panorama ha empezado a cambiar. Este cambio obedece a causas distintas:

Primeramente, la lingüística aplicada ha ido entrando en los ámbitos de trabajo de algunos lingüistas anteriormente teóricos de manera exclusiva y, con esta entrada, han ido cambiando también los prejuicios de que la teoría era pura especulación inútil para atender a las necesidades lingüísticas que plantea la sociedad actual. Con esta entrada, los prejuicios respecto a la falta de seriedad en la formación fundamental han ido desapareciendo.

* Institut de Lingüística Aplicada Universitat Pompeu Fabra (Barcelona)



En segundo lugar, los aspectos aplicados de las ciencias exactas y experimentales han ido adquiriendo un gran vigor en el mundo actual, y con ello, la lingüística teórica, que había hecho tantos esfuerzos por conseguir un estatuto científico y alistarse así en el conjunto de las ciencias experimentales, ha debido reconocer la lingüística aplicada como una posibilidad paralela a la planteada por otras ciencias.

En tercer lugar, se han ido perfilando las bases y conceptos de la lingüística aplicada y se ha aclarado el tipo de relación que mantenía con la teoría lingüística, de forma que ambas materias se consideran orientaciones metodológicamente distintas, pero complementarias, de un mismo objeto de estudio: el lenguaje humano, situándose ambas en un mismo marco disciplinar: el marco de las ciencias del lenguaje. De esta forma LT y LA constituyen dos polos de estudio complementarios de un mismo objeto, que se encuentran en relación de circularidad, y no de subordinación, formando parte de una única disciplina: la lingüística, entendida como estudio científico del lenguaje humano.

Así, el objeto de la lingüística teórica y el de la lingüística aplicada coinciden, si bien cambian otros elementos como la perspectiva en la que se estudia el lenguaje, la construcción del lenguaje como objeto científico y las finalidades que la investigación lingüística persigue.

Desde el punto de vista de la perspectiva, la lingüística teórica se ocupa de la formulación de hipótesis sobre cómo funciona el lenguaje, y no hace descripción de las lenguas particulares, cuáles son sus bases universales, comunes a toda la especie humana, y sobre qué significa ser un individuo competente lingüísticamente con independencia de la lengua que se hable. A la LT le preocupa la base común de las lenguas, por el hecho de ser lenguaje, y no los fenómenos específicos de tal o tal lengua concreta. Por ello, la LT parte del principio de que el lenguaje es una capacidad cognitiva (como lo son la facultad de ver, de oír, de moverse, de conocer, de orientarse, etc.) de base biológica, propia de toda la especie humana. El lenguaje se concibe así como una capacidad innata que se desarrolla en una u otra dirección (es decir, que se materializa en una u otra lengua particular) en virtud de los estímulos a los que los individuos se encuentran sometidos. Las diferencias entre las lenguas se consideran simples variaciones paraméricas de una base común, que configura la Gramática universal (GU). El objetivo fundamental de la LT es describir esta capacidad y los mecanismos que se ponen en marcha para ejercerla, es decir, para hablar y entender cualquier lengua concreta.

Complementariamente a esta concepción, y al mismo tiempo fuertemente diferenciada de ella, la lingüística aplicada, partiendo de los logros de la teoría, se centra en el estudio del funcionamiento lingüístico de los individuos en sociedad y en la descripción de las lenguas particulares para utilizarlos en la resolución de necesidades sociales tales como el aprendizaje de lenguas, la construcción de programas de tratamiento del lenguaje natural, la traducción general y especializada, la confección de diccionarios y terminologías, la clasificación documental o la preparación de programas de recuperación de trastornos lingüísticos, por citar sólo unos pocos ejemplos.



El objeto científico que sobre el lenguaje construye la teoría lingüística tampoco coincide con el que construye la LA. En efecto, para la LT, preocupada fundamentalmente por actuar como las ciencias experimentales, las lenguas son modelos idealizados de la realidad. Su objeto de estudio es una construcción parcial hipotetizada como un modelo global. Este modelo es analizado solamente en tanto que sistema de principios y condiciones comunes a todas las lenguas; es decir, sólo desde el punto de vista de la competencia (o conocimiento innato que los hablantes tenemos del lenguaje). La LA, en cambio, parte de la idea del lenguaje como sistema de comunicación social, para cuya descripción necesita tener en cuenta tanto la competencia (qué sabemos sobre el lenguaje) como la actuación (cómo usamos el lenguaje).

En esta misma lógica, el sujeto lingüístico de la teoría es el hablante-oyente ideal, individuo hipotético considerado fuera de las contingencias de espacio y tiempo que lo sabe todo sobre el lenguaje. Sólo así, la LT avanza en el establecimiento de modelos hipotéticos. El sujeto de la LA, en cambio, es un sujeto real que funciona en sociedad y que necesita el lenguaje como sistema de expresión e intercambio y que, por el hecho de estar sometido a la situación real, tiene del lenguaje un conocimiento parcial e interferido por los conflictos que se producen inevitablemente entre las lenguas en contacto.

Finalmente, para la teoría lingüística, el lenguaje como sistema debe ser inmanente y, por tanto, el lingüista debe poder explicarlo por las reglas propias de su sistema. El conocimiento lingüístico se concibe como un módulo, uno más entre las capacidades cognitivas del ser humano, que tiene sus propias leyes. De la interacción del conjunto de los distintos módulos cognitivos se explica el conocimiento general. Para la lingüística aplicada, en cambio, el lenguaje concebido como sistema de expresión y comunicación solo puede explicarse a partir de la interdisciplinariedad. Sin contar con la participación de todas las materias de base psicológico-social, difícilmente puede darse una idea fiable de cómo funciona el lenguaje.

2. LA LINGÜÍSTICA INFORMÁTICA EN EL MARCO DE LA LA

Dentro de la concepción de la LA que acabamos de describir y en el marco de las ciencias del lenguaje, situamos por un lado las vertientes aplicadas de la lingüística propiamente dicha: la fonología aplicada, la lexicología aplicada, la sintaxis aplicada, la pragmática aplicada; y, por otro lado, las llamadas disciplinas de intersección, esto es, las materias que tienen como objeto el lenguaje, pero que, a diferencia de la lingüística teórica, lo tratan como un fenómeno social y comunicativo, actúan desde una perspectiva de base social. Estas materias son fruto de la conjunción explícita de por lo menos dos disciplinas: la lingüística y la sociología (la sociolingüística), la lingüística y la historia (el análisis del discurso histórico o ideológico), la lingüística y la patología (la patología del lenguaje y la logopedia), la lingüística y la psicología (la psicolingüística), la lingüística y la informática (la lingüística computacional), etc. En los ámbitos de las interdisciplinas de base lingüística, siempre será legítimo que el lingüista considere estas materias interdisciplinarias desde la perspectiva del lenguaje, como también lo será que el sociólogo, el neurólogo, el psicólogo o el informático, las traten desde el punto de vista específico complementario.



La llamada por algunos lingüística informática y por otros informática lingüística (y aun para algunos —con una concepción precisa— lingüística computacional) presenta los ingredientes elementales que hemos descrito en el apartado anterior cuando hemos presentado las interdisciplinas de lingüística aplicada: tiene como objeto de análisis, estudio y tratamiento, el lenguaje. Para su análisis se sirve de la informática y, al mismo tiempo, produce sus análisis para ella. No podemos pues hablar de lingüística informática sin contar con ambas materias que son al mismo tiempo las dos origen y finalidad del campo de estudio:

- la lingüística proporciona el objeto de base: el lenguaje
- la informática permite tratar el lenguaje
- el lenguaje tratado informáticamente es objeto a su vez de la informática
- la informática aplicada al lenguaje permite utilizar los datos lingüísticos para la investigación sobre el propio lenguaje

En esta relación interdisciplinaria, la LI se ocupa de aspectos diferentes del lenguaje y, en cada uno de ellos, desde perspectivas distintas:

a) desde el punto de vista del nivel lingüístico objeto de estudio:

- * nivel fonológico/fonético
- * nivel prosódico
- * nivel morfológico
- * nivel léxico
- * nivel sintáctico
- * nivel textual

b) desde el punto de vista en que los trata:

- * transmisión/traspaso
- * análisis
- * generación
- * manipulación para otros procesos

c) desde el punto de vista del grado de tratamiento a que somete los datos y el nivel de actuación inteligente del ordenador:

- * primer nivel (elemental): trata los datos como señales gráficas. El ordenador es un simple transmisor.
- * segundo nivel: convierte los datos en señales codificadas y los trata como datos formales de carácter lingüístico. El ordenador es un analizador.
- * tercer nivel: comprende los datos y los procesa teniendo en cuenta cada situación. El ordenador es un mecanismo «inteligente» que actúa como un sistema experto dotado de un cierto grado de conocimiento.

Solo en este tercer nivel podemos hablar propiamente de inteligencia artificial.



3. LAS INDUSTRIAS DEL LENGUAJE

La expresión *industrias del lenguaje*, relativamente reciente (primera cumbre de jefes de estado de países de lengua francesa celebrado en París del 17 al 19 de febrero de 1986), sirve para designar un vasto campo de actividad industrial que trata informáticamente los datos lingüísticos como elemento constitutivo de un producto cuya finalidad no es lingüística.

Para muchos, las IL constituyen un campo de actividad de límites todavía imprecisos en el que se conciben, fabrican y comercializan ingenios y programas informáticos capaces de manipular, interpretar, generar, comprender y tratar el lenguaje humano, tanto en su forma oral como escrita, fundamentándose en los trabajos de investigación de las ciencias relacionadas con la información y las lenguas (la informática, la lingüística, las ciencias cognitivas, la documentación).

En un sentido amplio, consideramos que forman parte de las industrias del lenguaje todos los productos, técnicas, servicios o actividades que requieren un tratamiento automático del lenguaje natural. Las IL se integran en sistemas capaces de llevar a cabo alguna de las funciones siguientes:

- recibir, almacenar y tratar información de cara a facilitar su selección y consulta;
- proporcionar una descripción de la lengua escrita o hablada en forma de señales digitales, inteligibles para un ordenador, de manera que éste pueda reconocer e interpretar enunciados en lenguaje natural o generar mensajes de síntesis;
- perfeccionar, estandarizar y automatizar los procesos cognitivos y lingüísticos del tratamiento de la información, desde el punto de vista del sentido y de la significación.

Sobre esta base, los especialistas en informática aplicada al lenguaje señalan nueve polos de orientación de las actividades de las llamadas IL:

- comunicación hablada: síntesis y reconocimiento de la voz
- traducción asistida por ordenador o traducción automática
- interfaces en lenguaje natural
- tratamiento del lenguaje escrito, impreso y manuscrito
- asistencia al análisis, la comprensión y la producción de textos
- tratamiento avanzado de textos
- medios avanzados de edición de textos y de programas informáticos
- informática documental
- edición electrónica y publicación asistida.



En estos ámbitos, las IL fabrican y comercializan autómatas capaces de manipular, interpretar, generar lenguaje, tanto en forma oral como escrita, como si fueran humanos. Por lo tanto, las IL fabrican el conjunto de productos derivados de los resultados de la investigación en informática lingüística.

Pero poner en pie desde la base de la situación existente un campo nuevo de las características de las IL no ha sido, ni es, una tarea fácil. De entrada, las IL han debido sortear por lo menos tres tipos de problemas: 1) la formación de especialistas (en este punto se ha dado prioridad a un perfil de ingeniero especializado en lingüística, más que de al de lingüista con conocimientos de informática); 2) la puesta a punto de las lenguas a través de una reorientación de las investigaciones lingüísticas (los modelos formales de análisis lingüístico no eran adecuados para el procesamiento automático); y 3) la relación entre los investigadores y la empresa (ha sido necesario interesar a la industria no sólo en la financiación de la investigación en este campo, sino también en el desarrollo preindustrial de los productos elaborados).

Los resultados de la investigación lingüística a partir de los modelos que las propias ciencias del lenguaje han proporcionado no podían ser directamente utilizables por parte de los ordenadores. Y por ello, la lingüística informática ha tenido que cambiar de modelos y producir formalismos más adecuados al ordenador, con lo que se ha creado una separación entre los modelos formales que sigue la teoría lingüística actual (por una parte los paradigmas chomskianos; por otro los modelos de base sociofuncional) y los que sigue la lingüística computacional (fundamentalmente modelos secuenciales de un único nivel de representación).

Con estos supuestos resueltos, se han abierto diversas líneas de trabajo en el campo de las IL:

- la creación de bancos de recursos lingüísticos utilizables por los autómatas:

- *datos lingüísticos: bancos de textos, de palabras, de términos, de formantes, de fonemas, etc.),
- *reglas gramaticales
- *sonidos;

- analizadores lingüísticos y lematizadores;

- correctores y programas de redacción y edición asistidas;

- programas de síntesis y reconocimiento de voz;

- programas de traducción; etc.

Desde el punto de vista de la industria informática, las IL se suelen clasificar a partir de dos parámetros: a) por el tipo de producto que se genera, que es siempre (o casi siempre) un



programa de ordenador (o software); y b) por las herramientas previas que son necesarias para desarrollar estos productos. Estas herramientas previas son de tres tipos:

- a) herramientas de realización de productos lingüísticos: programas de tratamiento de texto, de corrección ortográfica, de edición electrónica, de indización automática, de producción automática de textos;
- b) herramientas de utilización de los datos lingüísticos: interpretación del lenguaje natural, sistemas de traducción automática, bancos de datos textuales, control vocal de aparatos, síntesis de la voz, sistemas expertos, programas de autoaprendizaje;
- b) herramientas de desarrollo del lenguaje: instrumentos de búsqueda y reconocimiento de unidades lingüísticas, redes de comunicación, bancos de datos lingüísticos.

EL desarrollo adecuado de todos estos productos presupone que se lleven a cabo:

1. Investigaciones en informática lingüística sobre el análisis de la señal (oral y escrita), sobre la estructura de la lengua (fonológica, morfológica, sintáctica y léxica), y sobre la estructura conceptual vehiculada por la lengua (análisis de discurso, análisis semántico), en la perspectiva de utilización de estos datos por autómatas.
2. El desarrollo de interfaces adecuadas entre las máquinas y las personas, que tengan en cuenta las características de las lenguas a tratar: teclados completos, cadenas de caracteres, lenguajes de interrogación y consulta.
3. La formación de personal especializado de alto nivel tanto en el ámbito de la investigación-desarrollo como en el de la utilización de máquinas y programas.

4. APARICIÓN Y DESARROLLO DE LAS INDUSTRIAS DEL LENGUAJE

Cuáles fueron en su día los motivos fundamentales primeramente de la creación del concepto «industrias del lenguaje» y, más tarde, de la importancia creciente que este ámbito ha ido adquiriendo?

En primer lugar, debemos destacar que el término «industrias del lenguaje» nace en 1986 en una reunión de tipo político: la primera cumbre de jefes de estado de países francófonos; y con una finalidad precisa: «asegurar que en el campo de la cultura y de la comunicación, los países de lengua francesa puedan beneficiarse de las tecnologías de punta actualmente en acelerado proceso de desarrollo, con vistas a ofrecer a las jóvenes generaciones un lugar de primera fila en el mundo moderno» (documento preparatorio de la cumbre de París de 1985: 1985: 2). Y,



asimismo, la noción «IL» se va delimitando, ampliando, perfilando y materializando en las siguientes cumbre de países de la llamada francofonía (París: febrero de 1986; Tours: febrero 1987; Quebec: setiembre 1987; Dakar: mayo 1989).

En segundo lugar, no podemos menospreclar, porque de no ser así difícilmente se habrían desarrollado las IL, el grado de desarrollo que la técnica (fundamentalmente de la tecnología) ha alcanzado en la actualidad; sobretodo las tecnologías de la información y la comunicación.

Finalmente, tampoco podemos olvidar que las características psicosociales del mundo actual han favorecido claramente el desarrollo y la explosión posterior de este nuevo ámbito de actividad.

La sociedad actual se caracteriza por una serie de rasgos que la diferencian notablemente de la sociedad del primer tercio de siglo. En ese sentido los antropólogos hablan de la entrada en una nueva civilización post-industrial emergente, marcada por un dominio tecnológico que permite incluso modificar las leyes más profundas de la naturaleza: la ingeniería genética o la biotecnología son muestras harto representativas de esa nueva era.

De la sociedad rural de los dos primeros tercios del siglo XIX, y de la cultura industrial de los últimos treinta años del siglo XIX y de la primera mitad del XX, hemos pasado a la llamada cultura post-industrial de la segunda mitad de nuestro siglo. Y podemos hablar de un cambio de civilización cuando un determinado número de rasgos característicos adquiere en una sociedad un prestigio que se va generalizando lentamente, hasta imponerse por encima de otro conjunto de rasgos que hasta entonces había sido preponderante.

El paso, por ejemplo, de la sociedad rural a la industrial supuso un cambio de civilización, manifiesta en unas nuevas formas económicas predominantes, en una distribución geográfica distinta de la población, o en una concepción de la familia, de la religión o del trabajo claramente diferenciada de la anterior. La sociedad rural se caracterizaba por su dedicación al cultivo de la tierra, por su dispersión en pequeños núcleos de población, por una concepción de la economía basada simplemente en la subsistencia, por la escasa escolarización de la población y el alto índice de analfabetismo, por un sentido jerarquizado y autoritario de las relaciones intragrupalas, por una importancia destacada de la religión, etc. Por el contrario, las sociedades desarrolladas integradas en el mundo industrializado concentran la población en grandes ciudades, donde se suelen instalar grandes complejos industriales que son la fuente de subsistencia de las familias que viven en ellas. La economía de mercado basa su desarrollo en el fomento de la producción y del consumo, lo que favorece la pérdida de la concepción de la familia amplia concebida como unidad económica y de protección. La concentración de la población —con la consiguiente aparición de la conciencia de clase— y la generalización de la enseñanza son factores que determinan la democratización de las relaciones interpersonales y de los regímenes políticos.



La nueva civilización que vivimos actualmente se diferencia netamente de la anterior. Sus características están todavía poco sistematizadas en conjunto, pero puede observarse un amplio abanico de manifestaciones cotidianas que nos permiten considerarlas al menos como indicativas de una nueva etapa. La pérdida de las ideologías tradicionales ha conducido, por ejemplo, a una sociedad basada en la búsqueda del triunfo material, en el fomento del individualismo y de la competitividad, en la mitificación del poder y del éxito.

En esa nueva fase el cambio cultural se percibe especialmente en dos grandes ámbitos: la tecnificación de la sociedad y el valor de la información. Esos dos nuevos pilares culturales han venido a sustituir otros elementos que caracterizaban la sociedad industrial que la precedía. Algunos de los cambios producidos han afectado de forma esencial el lenguaje y las formas de comunicación interpersonal, y han generado así nuevas formas de organización comunicativa, nuevas profesiones en torno al lenguaje y en consecuencia nuevos productos lingüísticos.

Así, en la civilización industrial, la generalización de la enseñanza y la importancia creciente de la comunicación escrita (frente al dominio de la comunicación oral en la cultura rural) hacía necesaria la codificación de las lenguas y la fijación de un registro estándar funcional. De este modo, en gran parte de las lenguas de cultura, los usos lingüísticos experimentaron un proceso de jerarquización, y llegó a establecerse «la norma»; y gracias a esa unificación, las lenguas dominantes se fueron consolidando.

En cambio, en la presente etapa, se produce un fenómeno paradójico en cuanto al estatus de las lenguas: por un lado se tiende hacia el monolingüismo, justificado por la necesidad de una comunicación directa y eficaz; por otro lado, se reivindican las lenguas nacionales como instrumentos naturales de comunicación, aptos para todos los usos, tanto para los generales como los especializados. La reivindicación de la identidad cultural justifica la defensa de las lenguas no dominantes frente al colonialismo cultural y lingüístico de las más poderosas. En ese contexto, se llevan a cabo importantes esfuerzos de racionalización en forma de proyectos de planificación lingüística organizados por los gobiernos y asumidos ampliamente por las sociedades. Esos planes, si son efectivamente racionales, conjugan a la vez el respeto a la identidad cultural y el fomento de las relaciones internacionales. Un proceso de normalización lingüística debe respetar las idiosincrasias culturales y formales de cada lengua, y debe permitir a su vez que la sociedad se integre en grupos más amplios y no se encierre en un aislamiento estéril.

La civilización actual está marcada por distintos elementos que, en tanto que civilización o cultura, la caracterizan específicamente y la distinguen de culturas anteriores; entre otros cabe destacar los siguientes:

- a) La ciencia y la técnica conocen un desarrollo sin precedentes, lo que provoca la aparición de una gran cantidad de conceptos nuevos e incluso de campos conceptuales nuevos, que requieren nuevas denominaciones.



b) La tecnología crece aceleradamente y se difunde de forma generalizada en la sociedad. Este hecho provoca la aparición de nuevos campos de actuación económica, como el de las llamadas «industrias del lenguaje». El mismo desarrollo tecnológico en el campo de la información y comunicación genera la necesidad de nuevas formas comunicativas que exigen una renovación constante.

c) Las relaciones internacionales, tanto políticas como culturales y económicas, crecen de forma espectacular. De un mercado restringido se pasa al mercado universal. Aparecen las empresas multinacionales. La producción en masa de bienes de consumo exige un mercado cada vez más extenso.

d) La masificación de la producción es consecuencia y a la par motor de la importancia exclusiva del producto estándar. El concepto «artesanal» se torna caduco.

e) La transferencia de conocimientos y de productos, considerada como uno de los aspectos más relevantes de la sociedad actual, provoca la aparición de nuevos mercados de intercambio científico, técnico, cultural y comercial; además, obliga a plantear y resolver el multilingüismo de los nuevos foros de intercambio; y, finalmente, provoca la necesidad de normalizar los elementos que canalizan el intercambio: los sistemas y las unidades básicas de transferencia.

f) La información pasa a tener una importancia capital, y se multiplica extraordinariamente. Esa gran masa de datos reclama soportes potentes y eficaces. Así, se crean bases de datos de todo tipo que requieren una actualización continua, que deben ser de fácil acceso y que además deben poder utilizarse multidimensionalmente. Como consecuencia, surge una nueva necesidad de normalizar los sistemas y los elementos de almacenamiento y recuperación de la información, así como la de estandarizar los sistemas de transferencia automatizada y los contenidos de los grandes almacenes de datos, cada vez más sofisticados.

g) El desarrollo de los medios de comunicación de masas permite una difusión generalizada de la terminología con la consiguiente interacción entre el léxico general y especializado. Gracias a los medios de comunicación, los términos específicos se banalizan.

h) La intervención de los gobiernos en materia lingüística a través de la aplicación de planes de normalización y la creación de organismos oficiales para gestionar las lenguas. La concentración casi exclusiva de la creación



científico-técnica en las potencias económicamente dominantes produce una transferencia unidireccional de conocimientos y nuevos productos que conlleva una entrada en masa de préstamos, sobre todo técnicos y comerciales, en países que hasta hace poco se consideraban culturalmente potentes. Una forma de hacer frente a esa nueva situación es la puesta en marcha de políticas lingüísticas, a menudo defensivas, que favorecen la aparición de nuevas profesiones en esos campos. Países pequeños, con o sin estado, con lenguas en situación inestable en algún grado, elaboran también planificaciones de la lengua que requieren un buen número de profesionales del lenguaje. Uno de los terrenos de importancia capital para la normalización de la lengua es el de los lenguajes de especialidad, y en su seno, el de la terminología.

Así pues, y a modo de síntesis, podemos afirmar que la emergencia de las IL en el mundo moderno se explica por varios fenómenos convergentes de naturaleza diversa:

- a) Factores de tipo psico-sociocultural (las características de la sociedad moderna);
- b) Factores de tipo científico y tecnológico (el avance de la investigación en todos los campos aplicados y, en particular, en el de las tecnologías de la información con la difusión de la microinformática);
- c) Factores de tipo político (la defensa de las lenguas por parte de los gobiernos);
- d) Factores de tipo económico (la generalización de mercados y la potenciación creciente de un «mercado lingüístico» [cabe no olvidar la celebración de ferias y exposiciones dedicadas específicamente a las lenguas; por ejemplo, Expolingua, que se convoca para casi todas las lenguas]).

La aparición y desarrollo de las llamadas IL tienen como objetivo primordial el de superar un reto múltiple: a) un reto económico (conurrencia en el mercado de la edición, de la traducción, de los bancos de datos, de los programas de diálogo persona-máquina, etc.); b) un reto cultural (supervivencia de las lenguas como lenguas vehiculares aptas para todo tipo de comunicación y tratamiento); c) y, finalmente, un reto político (adaptación por parte de los diferentes países a la modernidad, inserción en el multilingüismo, articulación en proyectos internacionales, cooperación en el eje norte-norte y participación en el desarrollo de los países del sur).

Que estos retos son los objetivos a conseguir lo explicita el texto del informe de síntesis de la Segunda Cumbre de Jefes de Estado y de Gobierno de los países francófonos celebrada en Québec del 2 al 4 de setiembre de 1987:



Le programme des industries de la langue est surtout orienté vers la défense et le développement de la langue française. La langue française doit s'inscrire rapidement dans le mouvement actuel d'industrialisation des langues; autrement, elle deviendra de moins en moins apte au développement de la recherche et de la production dans les secteurs de pointe et, à long terme, se marginalisera par rapport aux autres grandes langues de communication internationale dans ces champs d'activité essentiels à l'avenir de la francophonie. (Rapport de synthèse, 1987 : 173)

5. LAS IL EN SU DIVERSIDAD: INTENTO DE CLASIFICACIÓN

Para proceder a una clasificación sistemática de las IL, vamos a partir del supuesto de que las industrias del lenguaje se conciben como un campo todavía abierto que se desarrolla progresiva y simultáneamente por múltiples vías, y se restringen conceptualmente a las actividades que utilizan un material lingüístico para dar solución a problemas industriales, ergonómicos y comunicacionales haciendo uso de las nuevas tecnologías de la información.

Pero sin embargo, dado el carácter bastante reciente de este campo de actividad y su relativa falta de precisión como ámbito delimitado, cualquier intento unidimensional de clasificar las IL se ve abocado a dar de ellas una imagen parcial e incompleta. Así, procederemos a una clasificación multidimensional que tenga en cuenta sus diferentes facetas, fundamentalmente cuatro:

1) el grado de manipulación y transformación a que se somete el material lingüístico (la materia prima de las IL es el lenguaje natural). Podemos distinguir tres niveles de uso del texto:

a) Utilización directa sin ningún tipo de tratamiento previo (ej. los sistemas de tratamiento de texto, corrección, sistemas de edición, reconocimiento de caracteres, ayudas a la redacción, [tratamiento de texto, correctores, editica], programas de composición de textos, impresoras láser, etc.)

b) Tratamiento del lenguaje para una posterior selección de datos

*tratamiento del lenguaje oral:



contestadores automáticos
autómatas de llamadas
aplicaciones telemáticas
redes de transmisión de datos
correo oral
prótesis vocales
control industrial vocal
identificadores de voz
interfaces orales
máquinas de escribir con entrada de voz
sistemas de información vocal
sistemas de orientación vocal
interpretación
conducción vocal de máquinas y vehículos
electrodomésticos «inteligentes»
sintetizadores de voz a partir de texto
sistemas de reconocimiento de voz
etc.

*tratamiento del lenguaje escrito:

tratamiento de la información escrita
sistemas de correo electrónico
interfaces en lenguaje natural para diversos fines [gestión de bases de datos, diálogo persona/máquina, interpretación de las lenguas naturales, etc.]
formación asistida
sistemas de redacción automática y de generación multilingüe de correo y de documentos
sistemas de generación automática de textos
sistemas de búsqueda de información por ordenador (acceso y selección de la información contenida en las bases de datos)
analizadores de texto
sistemas automáticos de resumen y de producción de resúmenes (*abstracts*)
sistemas de ayuda a la confección de definiciones
tratamiento de las lenguas naturales (algoritmos de representación de los LN, generadores, correctores, lematizadores, etc.)
sistemas de traducción asistida y automática, etc.



c) Tratamiento de los datos lingüísticos:

- aplicaciones telemáticas
- bancos de datos
- informática aplicada a la lexicografía
- informática aplicada a la terminología
- producción de diccionarios electrónicos
- informática documental
- sistemas de indización
- sistemas de clasificación temática
- sistemas de ayuda a la confección de registros de bases de datos
- sistemas de lectura y comprensión de textos, etc.

d) Formalización de los datos lingüísticos para su inserción en un sistema complejo de inteligencia artificial (ej. traducción automática, etc.)

2) Según los ámbitos temáticos que más explotan las IL, distinguimos nueve polos de desarrollo:

- *comunicación oral (reconocimiento y síntesis)
- *traducción (asistida y automática)
- *interfaces en lenguaje natural
- *tratamiento del texto escrito (impreso y manuscrito)
- *análisis, comprensión y generación de textos (asistido y automático)
- *informática documental (creación de recursos)
- *edición electrónica y publicación asistida
- *tratamiento avanzado del texto escrito (ej: codificación en TEI)
- *inteligencia artificial (sistemas expertos inteligentes en un determinado nivel, capaces de realizar operaciones propias de los seres humanos, tanto a nivel de resolución de procesos como de comprensión e interpretación de datos y situaciones).

3) Según el tipo de investigación que se lleva a cabo, debemos distinguir entre:

- *la investigación fundamental y la aplicada
- *la investigación sobre lingüística, sobre informática (hard y soft), y sobre cognición.

d) De acuerdo con los campos en que se utilizan los resultados de las investigaciones y productos de las IL, estableceremos seis ámbitos profesionales de aplicación:



*ámbito de la informática: conviviabilidad de los programas, diálogos persona-máquina, creación de software, sistemas de gestión de bases de datos, etc.

*ofimática: tratamientos avanzados de texto, edítica, programas de ayuda a la redacción, la supervisión, etc.

*formación: enseñanza asistida por ordenador, sistemas de autoaprendizaje, etc.

*lingüística: asistencia a actividades diversas: traducción, reconocimiento de unidades, análisis de los datos lingüísticos, lematización, confección de glosarios y diccionarios, etc.

*gestión documental: sistemas de almacenamiento, indización y recuperación de la información; confección automática de abstracts, etc.

*sector medios de comunicación y mundo editorial: corrección de textos y edición automáticas.

d) A partir de las actividades básicas que los autómatas pueden realizar, establecemos cinco operaciones básicas:

*tratar el lenguaje humano: correctores, lematizadores, etc.

*manipular el lenguaje humano: lectores de texto (OCR), bases de datos, codificadores, etc.

*generar el lenguaje humano: generadores de texto, sintetizadores de voz, etc.

*comprender el lenguaje humano: reconocimiento de voz, interfaces, sistemas de clasificación, etc.

*realizar actividades lingüísticas complejas: reconocedores de unidades, traductores, etc.



6. A MODO DE CONCLUSIÓN

Con esta compleja clasificación de las llamadas imprecisamente «industrias del lenguaje» en una mezcla poco noble de la ciencia lingüística y del mercado de las lenguas, damos fin a este texto, que, como nos proponíamos, no pretende más que explorar horizontalmente este nuevo campo de investigación y actividad que tanto parece atraer —a juzgar por la priorización del tema en la mayoría de programas internacionales de investigación— a los gobiernos tanto de los países industrializados como a los que están en vías de industrialización.

Pero esta priorización que observamos no resulta gratuita. Parece bastante razonable considerar que las lenguas sólo sobreviven si se adaptan al tiempo, a las sociedades en que se hablan y a las necesidades de los individuos que las utilizan. Sólo pues si cumplen los requisitos de modernidad, flexibilidad e intercambio tienen mínimamente asegurada su supervivencia si los procesos históricos se producen de manera natural, sin sobresaltos inesperados.

Las lenguas deben ser útiles para la comunicación en todos los ámbitos, generales y especializados, formales e informales, sobre cualquier temática y canal y para cualquier propósito. Una lengua anclada en el rígido monolitismo de un solo registro es una lengua condenada a morir. Una lengua no actualizada al ritmo que se actualiza la sociedad está destinada a desaparecer por falta de eficacia. Una lengua aislada del panorama multilingüe que impone nuestra época es también una lengua con pocas posibilidades de sobrevivir.

Y, finalmente, en un mundo en el que la tecnología se encuentra en el primer plano de las actividades de comunicación e intercambio, y en el que cada vez más las máquinas reemplazan a los seres humanos en las actividades mecánicas o que fuerzan su comportamiento natural, las lenguas deben estar al nivel de ser tratadas tecnológicamente para poder participar de esta actualización. Una lengua que no esté a la altura de las demás lenguas concurrentes en el plano tecnológico, no figurará entre los sistemas de intercambio del futuro.

Así, en este complicado panorama social y político, marcado por los intercambios internacionales a todos los niveles y en todos los ámbitos, y en el que los distintos países se mueven entre la tensión y la cooperación norte-sur y entre la dicotomía entre la diversidad lingüística y la tendencia a reducir a una sola lengua la comunicación internacional, las industrias del lenguaje representan un punto clave para garantizar la supervivencia, la equidad y la dignidad funcional de todas las lenguas. Ya que, por principio y de manera natural, todas las lenguas son aptas para comunicarse en todos los ámbitos de intercambio.