

Moure, T., & Llisterri, J. (1996). Lenguaje y nuevas tecnologías. El campo de la lingüística computacional. In M. Fernández Pérez (Ed.), *Avances en lingüística aplicada* (pp. 147-228). Santiago de Compostela: Servicio de Publicacións e Intercambio Científico.

[http://liceu.uab.cat/~joaquim/publicacions/Llisterri\\_Moure\\_96\\_Linguistica\\_Computacional.pdf](http://liceu.uab.cat/~joaquim/publicacions/Llisterri_Moure_96_Linguistica_Computacional.pdf)

# LENGUAJE Y NUEVAS TECNOLOGÍAS: EL CAMPO DE LA LINGÜÍSTICA COMPUTACIONAL

**Teresa Moure**

Universidade de Santiago de Compostela

**Joaquim Llisteri**

Universitat Autònoma de Barcelona

## 1. EL ÁMBITO DE LA LINGÜÍSTICA COMPUTACIONAL

Bajo la denominación de *Lingüística computacional* es posible agrupar un conjunto relativamente heterogéneo de teorías, métodos, herramientas, aplicaciones y productos que tienen en común la consideración de la lengua como un objeto susceptible de ser tratado mediante procedimientos informáticos. Adoptando una perspectiva muy amplia, entrarán dentro de este ámbito desde las ayudas proporcionadas por los ordenadores a la investigación lingüística en disciplinas como la filología, la estilística, la estadística lingüística o la lexicografía hasta los sistemas capaces de llevar a cabo automáticamente la traducción de un texto o de ofrecer servicios telefónicos sin la intervención de un operador humano, pasando por los correctores que habitualmente se integran en los programas de tratamiento de texto o por los sistemas de recuperación automática de la información que hacen uso de datos lingüísticos. Puede observarse, pues, que se trata de un campo extremadamente diverso, dentro del cual se definen diferentes necesidades y objetivos. Por ello, tal vez sea útil empezar este capítulo intentando delimitar, aunque sea de forma somera y un tanto artificial, algunas de las aproximaciones actuales a la interrelación entre la lingüística y la informática.

## **1.1. Lingüística e informática: algunos enfoques**

### *1.1.1. Herramientas informáticas para la investigación lingüística*

El lector de formación filológica estará sin duda familiarizado con lo que en la tradición anglosajona se conoce como *Computers and the Humanities*, tendencia bien representada por la revista del mismo nombre publicada por la *Association for Computers and the Humanities* (ACH), o por *Literary & Linguistic Computing*, editada a su vez por la *Association for Literary and Linguistic Computing* (ALLC). Manuales como los de Hockey (1980) o Butler (1985) son también representativos de esta corriente, que agrupa a un amplio abanico de especialistas en lingüística sincrónica y diacrónica, filología, literatura, estilística, métrica, lexicografía, enseñanza de lenguas y traducción, entre otras áreas, interesados todos ellos en la aplicación de la informática al estudio de su propio campo. Existen programas que permiten, por ejemplo, obtener información estadística y distribucional sobre la aparición de unidades lingüísticas (Cf. 3.1.3), lo que es útil tanto en la descripción de la lengua como en la atribución de autor o fecha a un texto, o en la selección del vocabulario y las construcciones más usuales para la elaboración de programas de enseñanza de lenguas. La redacción de diccionarios, la constitución de repertorios terminológicos o la elaboración de atlas lingüísticos requieren la organización de los materiales en bases de datos, para lo cual existen también numerosos programas; con objeto de ayudar a la transcripción de conversaciones o de habla infantil se han elaborado igualmente programas específicos que facilitan una tarea de por sí compleja. En todos los casos, se trata, sin embargo, de un uso de herramientas informáticas que constituyen elementos auxiliares, aunque actualmente del todo imprescindibles, para llevar a cabo una investigación, sea para ordenar, clasificar y verificar la coherencia de los datos, sea para procesarlos y extraer información nueva que permita establecer generalizaciones o determinar tendencias estadísticas.

### *1.1.2. El procesamiento del lenguaje natural*

Frente a esta utilización instrumental de la informática, existe un campo cuyo objetivo es realizar automáticamente transformaciones entre distintas representaciones u objetos lingüísticos: pasar de un texto a una representación con información sobre la categoría gramatical de cada palabra, su estructura de constituyentes o su significado, traducir de una lengua a otra, resumir el contenido de un texto, extraer la información necesaria para recuperarlo después de haberlo introducido en un sistema de archivo, o escribirlo a partir de los conceptos básicos que forman su

estructura. Por ello se habla de *procesamiento o tratamiento del lenguaje natural* (natural language processing, NLP), usándose el término ‘natural’ para distinguir el lenguaje humano de los lenguajes de programación comunes en informática. Este ámbito está representado por publicaciones como *Computational Linguistics*, por asociaciones por –ejemplo, la *Association for Computational Linguistics* (ACL) con su capítulo europeo EACL o la *Sociedad Española para el Procesamiento del Lenguaje Natural* (SEPLN)– y por congresos como *COLING* (International Conference on Computational Linguistics). Panorámicas como la de Grishman (1986) ofrecen una buena perspectiva del campo que estamos describiendo, y que forma el núcleo de lo que se denomina específicamente *Lingüística computacional*. Los temas de trabajo propios de este ámbito se abordan en el apartado 3 de este capítulo.

### 1.1.3. Las tecnologías del habla

Aunque potencialmente la lingüística computacional abarcaría también el tratamiento de la lengua hablada, cuyo vehículo es la onda sonora, los trabajos en este terreno se incluyen más bien dentro de las denominadas *tecnologías del habla* (speech technologies). Ello obedece principalmente a razones históricas, puesto que si el procesamiento del lenguaje natural se inició desde la óptica de la informática, el del tratamiento del habla comenzó en el marco de la ingeniería de telecomunicaciones. Existen por ello asociaciones específicas como, en el ámbito europeo, la *European Speech Communication Association* (ESCA) que desde 1991 organiza bianualmente la *European Conference on Speech Communication and Technology* (EUROSPEECH) y publicaciones periódicas como *Speech Communication* dedicadas a las tecnologías del habla. En conjunto, se trata de un campo cuyos intereses se centran principalmente en la conversión de textos escritos en su equivalente oral –síntesis–, en la transformación del habla en texto –reconocimiento–, en la traducción automática de conversaciones, la identificación o verificación de hablantes en servicios telefónicos o en el desarrollo de sistemas que permitan el diálogo oral entre personas y máquinas. En el apartado 2 del presente capítulo se describen las principales líneas de trabajo en tecnologías del habla.

### 1.1.4. Las industrias de la lengua

Finalmente, cabe considerar las llamadas *industrias de la lengua*, denominación usada a menudo junto con la de *ingeniería lingüística* para referirse a las aplicaciones del procesamiento del lenguaje natural y del habla en el desarrollo de

productos comerciales, destinados a usuarios finales, que incluyen una parte importante de conocimientos sobre la lengua. Ejemplos de productos de este tipo serían los sistemas de conversión de texto a habla, de dictado automático, de traducción automática o asistida, los correctores ortográficos y gramaticales, las ayudas para el tratamiento multilingüe de textos, los programas de enseñanza de la lengua asistida por ordenador, o los sistemas de recuperación de la información a partir del análisis de textos. Las industrias de la lengua constituyen pues un ámbito orientado a las realizaciones prácticas y a la consecución de productos que cubran necesidades específicas de grupos de usuarios no especializados. Como puede deducirse, conllevan la implicación de empresas responsables del desarrollo y la comercialización de tales productos. Este campo dispone de revistas especializadas como el *Language Industry Monitor*, *Language Technology* o el *Journal of Natural Language Engineering* y de foros como la *Language Engineering Convention* que se celebra anualmente desde 1994. Por otra parte, existen también organismos dedicados al análisis del sector como la red francófona de observatorios de industrias de la lengua que incluye al *Observatoire Français des Industries de la Langue* (OFIL) editor de *La Tribune des Industries de la Langue* o el recién creado *Observatorio Español de Industrias de la Lengua* (OEIL). Publicaciones como Carré et alii (1991), Vidal Beneyto (1991), Sager (1992) o Hearn y Button (1994) presentan un panorama global de las industrias de la lengua.

#### 1.1.5. Hacia la integración

Es importante destacar después de esta breve panorámica que, pese a las diversas tendencias señaladas, existen numerosos elementos comunes en el amplio territorio que estamos describiendo. En primer lugar, se observa desde hace ya tiempo una integración cada vez mayor entre el procesamiento del lenguaje natural y las tecnologías del habla (Cf. 4); esta integración, objeto de proyectos europeos como *ELSNET* (European Network for the Integration of Language and Speech), se realiza tanto en lo que se refiere a las metodologías utilizadas –por ejemplo, el uso común de técnicas de modelado estadístico del lenguaje– como en el desarrollo de sistemas de traducción oral automática, que implican a la vez el procesamiento de la lengua hablada y la escrita. En el campo de las publicaciones, se encuentran revistas como *Computer Speech and Language* que abordan conjuntamente ambos aspectos. Congresos como los dedicados al procesamiento de la lengua oral (*ICSLP*, International Conference on Spoken Language Processing, organizado bianualmente desde 1990) contribuyen también a una mayor colaboración entre especialistas de ambos campos.

Por otra parte, las fronteras entre el procesamiento del lenguaje natural y del habla y las industrias de la lengua son relativamente difusas. En principio, puede pensarse que el primero desarrolla las técnicas mientras que las segundas producen las aplicaciones comerciales integradas en otros sistemas; en la práctica, existen empresas del sector que cuentan con departamentos de I+D (Investigación y Desarrollo) en los que se lleva a cabo el ciclo completo y en los que trabajan lingüistas computacionales o especialistas en tecnologías del habla junto con informáticos e ingenieros que desarrollan e implementan los productos.

Finalmente, existe un punto común entre las aproximaciones mencionadas anteriormente, que es la necesidad de los llamados *recursos lingüísticos*. Por recursos lingüísticos se entienden recopilaciones de textos escritos, de grabaciones de lengua hablada, diccionarios, terminologías especializadas o gramáticas, todos ellos en un formato digital que posibilite su acceso y tratamiento informáticos. Los recursos son imprescindibles para la creación de las aplicaciones que hemos ido mencionando, y buena muestra del interés que generan es la reciente creación de la *European Language Resources Association* (ELRA) dedicada a la difusión de recursos, fruto de una acción paralela a la del *Linguistic Data Consortium* (LDC) americano.

Al margen de su utilidad en el desarrollo de sistemas, los recursos lingüísticos son también poderosas herramientas en la descripción de la lengua (Cf. 3.1.3.). Por ello hay que añadir a los ámbitos anteriormente mencionados el de la *lingüística de corpus*, que se ocupa de la constitución y explotación de grandes muestras de uso real de la lengua tanto en su vertiente textual como oral. Presentaciones como las de Leech (1991) y Leech y Fligelstone (1992) ofrecen una buena perspectiva del campo<sup>1</sup>, que cuenta también con revistas como el *International Journal of Corpus Linguistics* y con congresos propios de la especialidad (Svartvik 1992, Leitner 1992, de Haan y Oostdijk 1993). De lo que acabamos de exponer se deduce que la lingüística de corpus puede ser considerada una disciplina transversal a los ámbitos mencionados, por lo que las nociones necesarias se irán presentando a lo largo del capítulo.

## 1.2. Lingüística computacional y Lingüística aplicada

Tal como hemos intentado mostrar en los apartados precedentes, bajo la denominación de Lingüística computacional se engloban una serie de campos muy

---

<sup>1</sup> Vid. también Aarts y Meijs (1984, 1986, 1990), Meijs (1987), Kyts et alii (1988), Aijmer y Altenberg (1991), Johansson y Stenström (1991), Aarts et alii (1993) y Baker et alii (1993). Para el español puede consultarse Alvar y Villena (1994).

diversos que tienen en común la integración de conceptos y procedimientos informáticos en el tratamiento del lenguaje y del habla. De aquí el término *Lingüística informática* que otros autores utilizan. Aunque hasta el momento hayamos hecho referencia principalmente a sistemas, no por ello debe pensarse que la lingüística computacional es un ámbito carente de base teórica. Por una parte, algunos de los conceptos y herramientas fundamentales se toman de las ciencias de la computación –por ejemplo, las estructuras de bases de datos–, de la lógica –algunos de los formalismos gramaticales–, del procesamiento de señales –por ejemplo, en las tecnologías del habla– y, finalmente, muchos de ellos provienen de la lingüística. Por otra, la lingüística computacional ha desarrollado sus propios formalismos –por citar dos casos, en el ámbito de la sintaxis y de la descripción del léxico– y sus propios conceptos teóricos. Tales conceptos están en la base del desarrollo de sistemas que realizan operaciones concretas –traducción, conversión de texto a habla, generación de textos, etc.– y de herramientas para el tratamiento de los datos –etiquetado de partes de la oración, análisis sintáctico automático, alineación de corpus paralelos, etc.–. Los sistemas pueden convertirse en productos comerciales, sea incluidos en una aplicación más general –por ejemplo el control vocal del sistema operativo de un ordenador o los correctores ortográficos y gramaticales en un procesador de textos– o en forma de productos independientes –dictado automático, traducción automática, conversión de texto a habla, etc.–. También los recursos lingüísticos como los diccionarios o los repertorios terminológicos son susceptibles de comercializarse –típicamente en CD-ROM<sup>2</sup>– una vez elaborado un buen sistema de acceso a la información que contienen. El conocimiento sobre el lenguaje se integra pues, junto con elementos procedentes de otras disciplinas, tanto en la teoría de la lingüística computacional como en la concepción de sistemas y herramientas y en la creación de recursos lingüísticos. Por ello, la lingüística computacional en el sentido amplio en que la venimos describiendo puede considerarse como una de las aplicaciones de la lingüística –al igual que puede considerarse como una aplicación de las ciencias de la computación al tratamiento del lenguaje y del habla– en tanto que toma de ella conceptos, unidades, procedimientos de análisis y modelos de descripción para incorporarlos en un nuevo marco (Cf. 5). En el momento actual, los investigadores en procesamiento del lenguaje natural proceden tanto de la lingüística como de las ciencias de la computación, al igual que los que se dedican a las tecnologías del habla provienen del campo de la fonética y de la ingeniería de telecomunicaciones; también en estos campos encontramos personas especiali-

---

<sup>2</sup> Los CD-ROM (Read-Only-Memory Compact Disk) son discos para ordenador similares a los discos compactos para sistemas audio, que pueden contener enormes volúmenes de texto o de señales sonoras ya que almacenan unos 600 MegaBytes.

zadas inicialmente en psicología, en matemáticas o en otras disciplinas. Es preciso resaltar que cada vez es mayor el número de universidades que ofrecen una formación específica en lingüística computacional o en tecnologías del lenguaje y del habla –en algunos países existen Licenciaturas y en otros, como en el caso de España, Cursos de postgrado–, integrando junto con la lingüística otras materias de los ámbitos relevantes para una formación adecuada (ELSNET 1993; Bloothoof et alii 1995). Con ello se está consolidando un campo de trabajo caracterizado principalmente por la interdisciplinariedad y por ofrecer la posibilidad de convertir las teorías en realidades, materializadas en productos que, en última instancia, tienen como objeto ayudar a las personas en aquellas tareas en las que el lenguaje juega un papel preponderante.

## **2. EL TRATAMIENTO DEL HABLA**

### **2.1. Las tecnologías del habla: objetivos y problemas**

Podría decirse que el objetivo global de las tecnologías del habla es llegar a disponer de procedimientos que permitan la comunicación natural entre las personas y los ordenadores mediante el habla<sup>3</sup>. Para ello es preciso que un sistema informático sea capaz de transformar un enunciado oral en una representación simbólica y convertir a su vez el habla en un conjunto de símbolos. Tal necesidad obedece a que, mientras que el habla humana es una señal sonora continua que varía a lo largo del tiempo en –otras palabras, una señal analógica–, las aplicaciones informáticas trabajan sobre cadenas de símbolos discretos –es decir, sobre señales digitales–.

El paso de un tipo de representación a otra no está exento de dificultades, dada la naturaleza variable de la onda sonora que producimos al hablar. Por una parte, una misma unidad lingüística –pongamos por caso el fonema– no tiene idéntica manifestación sonora en todos los contextos en que puede aparecer y, a su vez, en un determinado segmento de la onda sonora se encuentra información acústica correspondiente a más de una unidad lingüística. Este es el problema descrito en fonética como la ausencia de invariación en las manifestaciones acústicas de las unidades lingüísticas<sup>4</sup>. Por otra parte, cada hablante posee características

---

<sup>3</sup> Vid. O'Shaughnessy (1987), Holmes (1988), Calliope (1989), Saito (1991) y Keller (1994) para una presentación de las tecnologías del habla.

<sup>4</sup> Lindblom (1986:496) resume claramente la cuestión:

For a given language there seems to be no unique set of acoustic properties that will always be present in the production of a given unit (feature, phoneme, (demy) syllable, etc.) and that will reliably be found



individuales –algunas de ellas intrínsecas como el sexo, otras variables como la edad, el estado físico o el estado emocional– y rasgos que señalan su pertenencia a un grupo social y a una zona geográfica, al tiempo que puede llegar a dominar una amplia gama de estilos o diferencias en el habla asociadas a una situación comunicativa particular. A todo ello se alude en fonética cuando se mencionan las variaciones inter- e intra-locutor. Mientras que el sistema humano de procesamiento del habla está preparado para responder a tales variaciones y también para realizarlas, la producción y el reconocimiento automáticos por parte de un ordenador suponen la capacidad de tratar adecuadamente los tipos de variación que acabamos de mencionar.

Por ello, aunque el campo de las tecnologías del habla se haya desarrollado tradicionalmente ligado a la ingeniería de telecomunicaciones, cada vez es mayor la presencia de especialistas en fonética y en comunicación oral en los equipos de investigación. Tanto los problemas de la invariación y la variabilidad ligados a la segmentación del continuum sonoro como el desarrollo de modelos de producción y percepción del habla están en el centro de las preocupaciones de la teoría fonética, y desde un buen conocimiento del proceso de comunicación humano pueden realizarse aportaciones al desarrollo tecnológico (Llisterri 1990). Pero por otra parte, la necesidad de desarrollar aplicaciones informáticas que, de alguna manera, realicen las dos operaciones básicas de producir y reconocer, obliga también a responder a una serie de preguntas que estimulan nuevas investigaciones en el campo de la fonética<sup>5</sup> y a formalizar en sistemas de reglas explícitos los conocimientos ya adquiridos. Por este motivo, la relación entre la fonética –o el estudio de la comunicación oral– y las tecnologías del habla puede concebirse, como sugiere Fant (1984), como una separación entre teoría y aplicaciones, más que como una división real en dos campos distintos. En la práctica actual, es la orientación hacia uno u otro aspecto lo que guía las diversas líneas de trabajo que conviven y se fertilizan mutuamente en el interior de grupos cada vez más interdisciplinarios.

A pesar de las dificultades anteriormente mencionadas, la posibilidad de comunicarse oralmente con los ordenadores ofrece una serie de ventajas (Witten 1982, Nadeu y Mariño 1985): la libertad de movimientos y la facilidad para llevar

---

in all conceivable contexts [...] The signal cannot be unambiguously segmented into temporally, non overlapping chunks corresponding to e.g. phonemes, syllables and words.

<sup>5</sup> Así lo plantea, por ejemplo, Allen (1985:549):

While we usually see the goals of research in speech synthesis and recognition as the satisfaction of important technological needs, this research will also be of great value for the insight it provides into complex human behavior and its modelling through the utilization of imaginative and comprehensive computational systems.

a cabo otras tareas al no tener que depender de una pantalla y un teclado, la potencial simultaneidad del habla con otros sistemas de comunicación, la posibilidad casi universal de acceso remoto mediante el teléfono y, especialmente, el hecho de que el habla constituye nuestro sistema de comunicación más natural y, por lo tanto, su utilización contribuye a acercar el mundo de las tecnologías de la información a una amplia gama de personas, incluyendo aquellas que, por razones diversas como la edad, el nivel cultural o determinadas discapacidades, tienen dificultades con la lectura, la escritura o con el uso de teclados y monitores.

Atendiendo al objetivo general anteriormente descrito, las tecnologías del habla pueden subdividirse en dos grandes ámbitos: la síntesis, mediante la cual se transforma un conjunto de símbolos discretos en una señal sonora del habla, y el reconocimiento, en el cual se realiza la operación inversa. Dentro de cada uno de ellos existen diversas aproximaciones, que se exploran brevemente en los apartados 2.2 y 2.3. Por otra parte, existe también la posibilidad de combinar las tecnologías del habla con las tecnologías orientadas al tratamiento de la lengua escrita, abriendo las perspectivas que se presentan en el apartado 4.

## **2.2. La síntesis del habla**

### *2.2.1. Características generales de los sistemas de síntesis*

Como hemos mencionado anteriormente, un sistema de síntesis debe generar automáticamente una señal sonora que sea interpretada como un mensaje oral por un oyente humano. Para conseguir este resultado existen diversas técnicas, como veremos en el apartado 2.2.2. Sin embargo, al margen del procedimiento utilizado, los sistemas de síntesis pueden caracterizarse mediante una serie de parámetros generales: la naturalidad e inteligibilidad del habla sintetizada, la versatilidad del sistema, la complejidad del procesamiento y del equipo requerido, y la relación con la aplicación prevista<sup>6</sup>. En lo que se refiere al primer factor, a la hora de evaluar los resultados de un sistema de síntesis suele distinguirse entre la inteligibilidad y la naturalidad. Si la inteligibilidad es relativamente fácil de cuantificar por medio de una serie de pruebas más o menos estandarizadas en las que se consideran los porcentajes de vocales, consonantes, palabras o frases correctamente identificados por un grupo de oyentes, la medida de la naturalidad es más compleja, ya que

---

<sup>6</sup> Para una visión general de la síntesis del habla vid. Flanagan (1972), Mariño, Nadeu y Llisterri (1987) o Llisterri (1988). Son útiles para profundizar en el tema Bristow (1984), Keller (1994) y van Santen et alii (1995). Una perspectiva histórica del campo se halla en Flanagan y Rabiner (1973) o en Martí (1987).

depende de un conjunto de factores entre los cuales juegan un papel preponderante la entonación, la duración de los sonidos, la asignación de pausas o la calidad de la voz<sup>7</sup>. En cuanto a la flexibilidad, se entiende aquí la capacidad del sistema de producir una lista cerrada de palabras o frases o de generar enunciados sin restricciones. También puede incluirse en ese factor la posibilidad de variar automáticamente las voces, la variante geográfica, o el estilo de habla utilizados. Naturalmente, cuanto más flexible es un sistema, más aumenta su complejidad interna y los requisitos informáticos necesarios para su buen funcionamiento.

Finalmente, las decisiones que se tomen sobre las características del sistema en el momento de su diseño dependen de la aplicación final que se le desee dar. Por ejemplo, un sistema de síntesis concebido como herramienta de trabajo para usuarios invidentes debe cubrir una amplia gama de requisitos. De entrada, deberá ser lo bastante natural para evitar el cansancio que, de otro modo, produciría su uso continuado. Pero además, ha de ser inteligible, capaz de leer cualquier tipo de texto y de funcionar con un ordenador personal. En cambio, un sistema incorporado a un automóvil y destinado a facilitar información sobre el estado del vehículo únicamente tiene que producir un conjunto muy limitado de enunciados, lo que reduce esos requisitos de eficiencia, pero a cambio exige que su diseño no encarezca excesivamente el precio final del producto.

### *2.2.2. Las técnicas de síntesis*

En conjunto las técnicas de síntesis pueden dividirse en tres grandes categorías: por un lado, la codificación del habla, que es esencialmente un proceso de grabación y reproducción; por otro, las que se basan en la concatenación de unidades previamente parametrizadas y almacenadas y, finalmente, la síntesis por reglas, en la que se formalizan conocimientos fonéticos para la creación de enunciados.

Los parámetros anteriormente descritos permiten ofrecer una somera caracterización de estas tres técnicas. Con la codificación podemos obtener una excelente calidad de habla, pero la flexibilidad queda limitada a la reproducción de los enunciados almacenados. Mediante la parametrización de un conjunto de unidades se llega a reducir las necesidades de memoria del sistema, puesto que los enunciados se forman mediante la concatenación de estas unidades, pero a cambio

---

<sup>7</sup> Sobre la evaluación de los sistemas de síntesis vid. van Bezooijen y Pols (1990) y Pols (1991). Aguilar et alii (1994) presentan pruebas de evaluación para el español.

la calidad de la síntesis disminuye al perder información directa sobre la onda sonora. Finalmente, la síntesis por reglas requiere un procesamiento fonético muy complejo, que se traduce también en una disminución de la calidad del habla generada; en contrapartida, la flexibilidad del sistema es total, y el número de unidades mínimo, ya que se parte del inventario fonético de la lengua.

#### 2.2.2.1. La codificación del habla

Las técnicas de síntesis mediante codificación del habla se basan en el principio del almacenamiento y reproducción de los mensajes requeridos para una determinada aplicación. No difieren pues, en su idea básica, de una grabación convencional como la que puede hacerse con un sistema reproductor de audio, aunque en el caso de la síntesis se trabaja sobre un formato digital. Es claro pues que el número de mensajes disponibles está limitado a los que previamente se hayan recogido, y que la calidad de la síntesis depende por completo de la aplicación de técnicas de procesamiento de señales digitales. Con los procedimientos adecuados puede obtenerse una calidad muy alta, aunque en detrimento de la cantidad de memoria requerida para almacenar los mensajes. Por ello, los esfuerzos que se realizan en este campo se concentran en reducir al máximo la cantidad de memoria necesaria, acudiendo a técnicas de codificación de la onda sonora. Con la codificación se persigue guardar el mínimo de información necesaria sobre la señal a la vez que se mantiene la máxima calidad de reproducción, efectuando operaciones que podríamos denominar de «compresión» del habla.

Este procedimiento es útil cuando la aplicación requiere únicamente un número reducido de enunciados sintetizados, por ejemplo un «chip» sonoro que nos informe sobre las operaciones que realiza un electrodoméstico en un momento dado, sobre los problemas que puedan surgir durante la conducción de un vehículo o, en un nivel algo más complejo, un contestador automático de una centralita telefónica. Para este tipo de aplicaciones, suelen codificarse bien los enunciados completos, bien una lista de palabras que posteriormente se concatenan o se insertan en una frase-marco también codificada –por ejemplo «La próxima parada es...»–. En tal caso, es necesario considerar el mantenimiento adecuado de las características prosódicas.

Sin embargo, en el diseño y la implementación de los sistemas de síntesis por codificación no interviene casi ningún conocimiento fonético ni lingüístico, ya que se requiere únicamente un buen dominio del procesamiento digital de señales, propio de la ingeniería de telecomunicaciones. La intervención del lingüista puede limitarse, como máximo, a la selección del hablante o hablantes que realizarán las grabaciones y a la evaluación de la calidad del resultado final.

#### 2.2.2.2. La concatenación de unidades parametrizadas

Una alternativa a la grabación y codificación de mensajes completos o de listas de palabras es la concatenación de unidades menores previamente parametrizadas y guardadas en un diccionario de unidades de síntesis. En este caso se utilizan unidades menores que la palabra y se conservan para cada una de ellas una serie de propiedades que, como veremos, mantienen una estrecha relación con un modelo de producción del habla. Su interés lingüístico es, por tanto, mucho mayor y requiere la intervención del conocimiento fonético de un experto.

##### 2.2.2.2.1. Las unidades de síntesis

En principio, cualquiera de las unidades habitualmente consideradas en el análisis lingüístico podría utilizarse en el desarrollo de un sistema de síntesis. Sin embargo, la selección de la unidad viene condicionada por la aplicación, la cual a su vez determina también la técnica de síntesis elegida. En un sistema de codificación del habla suelen emplearse, como hemos visto, frases o palabras. En cambio, cuando la flexibilidad del sistema debe ser mayor, es habitual recurrir a unidades inferiores.

Una manera de reducir la memoria necesaria y de incrementar simultáneamente las posibilidades de un sistema de síntesis consiste en partir de un inventario de unidades que puedan almacenarse conservando la información esencial para cada una de ellas –es decir, parametrizándolas– y combinándolas –o concatenándolas –para formar un número infinito de enunciados. En la selección de la unidad más idónea intervienen de forma decisiva criterios fonéticos relativos al conocimiento sobre la variabilidad contextual de diversos tipos de elementos. Por ejemplo, aunque la situación ideal sería poder concatenar elementos simples y que existan en número reducido como los alófonos, es claro que la simple unión de tales elementos no llega a producir un mensaje natural si no se reproducen adecuadamente los efectos coarticulatorios que se dan entre ellos en el habla continua. La modelización de tales efectos aún no está completamente lograda en la síntesis, y resulta extraordinariamente compleja dado el elevado grado de variabilidad contextual de las unidades segmentales. Por este motivo, en los sistemas basados en la concatenación se utilizan unidades que incluyan ya en su interior las zonas de mayor variabilidad acústica en el paso de un segmento a otro –las transiciones– y que puedan yuxtaponerse en los puntos en que la onda sonora es más estable, es decir, en la parte central de una vocal o de una consonante como una lateral, fricativa o aproximante o en el silencio de las oclusivas. Tales unidades son los *difonemas* y las *semisílabas* (Hess 1995).

Un difonema –denominado a veces difono– es un segmento acústico que incluye la transición entre dos alófonos consecutivos, formado por la parte estacionaria del primero, la transición del primero al segundo y la parte estacionaria del segundo; en la palabra castellana «casa» se distinguen los siguientes: /k/-/ka/ -/as/-/sa/-/a/. Para facilitar la concatenación de segmentos acústicamente muy poco estables como pueden ser las consonantes vibrantes o las semivocales y semiconsonantes, se introducen en los sistemas de síntesis unidades como los trifonemas o los cuatrifonemas –llamadas genéricamente polifonemas o polifonos–, en el interior de los cuales se sitúan los segmentos que ofrecen mayores dificultades de concatenación. En cambio, la semisílaba es el fragmento de sílaba comprendido entre su inicio y el centro de la vocal o entre el centro de la vocal y su final.

Puede observarse que éstas no son unidades habituales en el análisis fonético o fonológico de la lengua, pero en cambio responden perfectamente a los requisitos de un sistema de síntesis: se concatenan con relativa facilidad, no es necesario un número extraordinariamente elevado para la producción de cualquier tipo de enunciado y responden a determinadas propiedades acústicas bien conocidas por la fonética, por lo que su parametrización puede realizarse en función de modelos ampliamente utilizados. La necesidad de desarrollar aplicaciones ha tenido como resultado la aparición de nuevas unidades propias de las tecnologías del habla.

#### 2.2.2.2. La parametrización de las unidades

Una vez seleccionadas las unidades, es necesario buscar un procedimiento para almacenarlas eficazmente en la memoria del sistema. La alternativa a guardar las unidades enteras es conservar únicamente una serie de parámetros de cada una. A la hora de seleccionar estos parámetros, se ha tenido en cuenta una de las aportaciones más cruciales para el desarrollo tanto de la fonética como de las tecnologías del habla: el modelo de la fuente y el filtro<sup>8</sup>. Esencialmente, podemos dividir el aparato fonador en dos partes: aquella en la que se origina la onda sonora necesaria para la producción de un sonido –las cuerdas vocales en los sonidos sonoros o el tracto vocal en los sordos– y aquella en la que se realiza una modificación de las características de la onda –el tracto vocal, incluyendo las cavidades faríngea, bucal y nasal–. Denominamos «fuente» a la primera, mientras que de la segunda se dice que tiene un efecto de «filtro». Por ejemplo, una vocal oral tiene su fuente en las cuerdas vocales y su filtro en la cavidad bucal, mientras

<sup>8</sup>La primera descripción formal de este modelo se encuentra en Fant (1960). Para una presentación más elemental véase, por ejemplo, Lieberman y Blumstein (1988).

que una oclusiva bilabial sorda tiene su fuente en la explosión labial y su filtro en la cavidad bucal. La fuente y el filtro pueden modelarse matemáticamente y sus efectos son susceptibles de reproducirse digitalmente en un sintetizador. Por ello, la parametrización de unidades de síntesis como el difonema y la semisílaba suele realizarse analizando primero y almacenando después las características de la fuente y del filtro. La fuente se relaciona con la sonoridad, mientras que el filtro responde a la forma del tracto vocal y, por lo tanto, sus características se reflejan en la estructura formántica.

En la actualidad se utilizan principalmente dos técnicas para la parametrización de las unidades de síntesis en términos de fuente y filtro: la predicción lineal (Linear Predictive Coding, LPC) y el análisis en formantes. En ambos casos se requiere en primer lugar un análisis de las unidades y un almacenamiento de los parámetros obtenidos. En segundo lugar, es preciso diseñar un sintetizador, que no es más que un sistema digital mediante el cual, a partir de los valores de los parámetros, se reconstruye la forma sonora original de las unidades. Como es lógico, la calidad de la síntesis depende tanto de una parametrización correcta como de la buena reconstitución llevada a cabo por el sintetizador.

En un sistema de síntesis por LPC las unidades se parametrizan en una serie de coeficientes que reflejan las variaciones en las propiedades acústicas del tracto vocal en la producción de los sonidos que configuran la unidad, y se reconstruyen posteriormente mediante un sintetizador LPC, que interpreta estos coeficientes para generar la onda sonora correspondiente. Desde la perspectiva de la fonética, los coeficientes LPC son difíciles de relacionar directamente con los parámetros habitualmente usados en la descripción articulatoria y acústica, aunque los resultados de la síntesis permitan desarrollar aplicaciones útiles en diversos campos, dado el ahorro de memoria que supone un sistema de estas características.

En cambio, la parametrización y la síntesis de las unidades en términos de formantes ofrece la posibilidad de interpretar el proceso partiendo de la base del conocimiento fonético actual sobre la producción y las características acústicas de los sonidos del habla. Por ello, la síntesis por formantes es un modelo apto tanto para la investigación en fonética como para el desarrollo de aplicaciones; el sistema desarrollado por Klatt (1980) es quizás uno de los más utilizados en ambos casos.

Otra posibilidad, empleada recientemente con éxito, es la codificación de las unidades en el dominio del tiempo, utilizando técnicas de procesamiento de señales como el *PSOLA* –Pitch Synchronous Overlapp and Add– (Moulines y Charpentier 1990).

### 2.2.2.3. La síntesis por reglas

Las técnicas de síntesis que hemos visto hasta ahora parten de enunciados o palabras codificadas o de un diccionario de unidades parametrizadas; en ambos casos el objetivo es recuperar y reconstruir información previamente almacenada. Mientras que en la codificación se recurre a técnicas de procesamiento de señal, la síntesis paramétrica hace uso de sintetizadores controlados por los valores de los parámetros, que, en el caso de la síntesis por formantes, se relacionan directamente con modelos de análisis en fonética.

En cambio, la síntesis por reglas parte del conocimiento de los valores de los parámetros acústicos que caracterizan a cada uno de los alófonos de la lengua y de la formalización mediante reglas de las variaciones acústicas provocadas por su aparición en diversos contextos. De modo que el conocimiento almacenado no consiste en diccionarios de unidades parametrizadas, sino en tablas de datos que precisan los valores frecuenciales y temporales de cada uno de los alófonos de la lengua y de sus transiciones, acompañadas de reglas de concatenación o modelos de transición de un segmento a otro. Naturalmente, también en este caso se requiere un sintetizador –habitualmente por formantes– que convierta esas cadenas de valores de parámetros en una onda sonora.

Las unidades utilizadas en este caso son pues fonemas o alófonos, pudiéndose usar también tramas (frames) para representar segmentos acústicos de características estables más breves que los alófonos. La reducción de la información es importante, aunque a cambio aumenta extraordinariamente la complejidad de las reglas; pero por otra parte, la síntesis por reglas es un modelo basado en conocimiento fonético, sobre el que se empezó a trabajar ya en los años cincuenta en los Laboratorios Haskins y que ha continuado refinándose hasta las implementaciones actuales en diversos sistemas.

Finalmente, en el ámbito de los sistemas de síntesis basados en reglas, debe mencionarse la síntesis articuladora (Gabioud 1994). En este caso, los parámetros que controlan el sintetizador no son conjuntos de valores acústicos, sino reglas articuladoras que describen la posición del tracto vocal en la producción de cada sonido y las estrategias de coarticulación. En el momento de la síntesis se parte de esta información para calcular el resultado acústico de una sucesión de configuraciones del tracto vocal. Sin embargo, la modelización articuladora es compleja, debe basarse en datos extraídos del análisis de la producción del habla y considerar la interacción entre la articulación y su resultado acústico. Por este motivo es difícil producir automáticamente habla sintetizada de calidad controlada únicamente por parámetros articulatorios, y estos sistemas aún no se han implementado en



aplicaciones comerciales, quedando de momento reducidos al ámbito de la investigación.

### *2.2.3. Las aplicaciones*

#### *2.2.3.1. La conversión de texto a habla*

La aplicación principal de la síntesis es la creación de sistemas de conversión de texto a habla (text-to-speech, TTS) que, partiendo de un texto escrito sin ningún tipo de restricción, permitan llegar a su realización sonora<sup>9</sup>. Para ello, es preciso disponer de un sistema modular que opere una serie de transformaciones entre diferentes niveles de representación, desde la cadena de caracteres que forma el texto hasta llegar a la onda sonora. A continuación se describe brevemente cada uno de los módulos.

Un texto escrito como el que podemos encontrar en un periódico, por ejemplo, no contiene únicamente caracteres directamente interpretables, sino que presenta también abreviaturas, siglas, cifras, números romanos, fechas, horas, y algunos símbolos especiales. Todo ello tiene que ser convertido en texto escrito deletreado, y ésta es precisamente la función del módulo de pre-procesado lingüístico de un conversor. Si bien en algunos casos, como en las abreviaturas o las siglas, la tarea es relativamente fácil si se dispone de un diccionario, la conversión de cifras requiere, por ejemplo, mantener la concordancia en los casos de «un» - «una» o «doscientos» - «doscientas». Por ello un módulo de pre-procesado lingüístico incluye tanto diccionarios como conjuntos de reglas para la transformación de los elementos mencionados en cadenas de caracteres.

El siguiente paso necesario para la conversión de texto a habla es la transcripción fonética automática de la cadena de caracteres obtenida del módulo de pre-procesamiento lingüístico. Esto puede llevarse a cabo mediante un conjunto de reglas que definen la correspondencia entre grafías y sonidos, pero también puede intervenir eficazmente un diccionario que registre las excepciones. En esta misma fase suele asignarse el acento fonético a las palabras y, de ser necesario, se realiza también una división en sílabas. En algunas lenguas como el inglés, el proceso es más eficiente si se lleva a cabo previamente un análisis morfológico de las palabras, aunque para el español, por ejemplo, este proceso no es estrictamente necesario<sup>10</sup>.

---

<sup>9</sup> Para una presentación general de la conversión de texto a habla vid. Allen et alii (1987), Klatt (1987) y Carlson et alii (1990).

<sup>10</sup> Algunos aspectos de la transcripción fonética automática en español se tratan en Enríquez (1991) y en Ríos (1993).

Obtenida la transcripción fonética, cabe incorporar información propia de otros niveles de análisis lingüístico como pueda ser la parte de la oración o un análisis sintáctico<sup>11</sup>. Este último es fundamental si pretendemos que la lectura del texto se realice con características prosódicas que la acerquen al habla natural, siendo especialmente importante para una correcta asignación de pausas. También en este módulo de análisis lingüístico sería necesario poder introducir información de tipo semántico y pragmático, pensando siempre en el objetivo final de conseguir una lectura próxima a la de un hablante humano. Aunque para ello sería preciso que este aspecto estuviera mucho más avanzado en el campo del procesamiento del lenguaje natural (Cf. 3.2.2.2 y 3.2.2.3), la mejora de los sistemas actuales de conversión de texto a habla pasa indudablemente por la incorporación de técnicas desarrolladas en este ámbito.

Llegados al punto en que disponemos de una adecuada transcripción fonética del texto y de toda la información lingüística que hayamos incorporado, puede actuar ya el módulo prosódico. Su primera función es asignar adecuadamente las pausas en el texto, basándose tanto en los signos de puntuación como en el análisis sintáctico; en defecto de este último, puede usarse también información sobre la parte de la oración a la que pertenece cada palabra pues sabemos que, por ejemplo, en español difícilmente se introducirá una pausa entre el artículo y el nombre que le sigue. La segunda función del módulo prosódico es determinar las modificaciones a la duración de cada uno de los elementos segmentales en función del contexto en que aparezcan. Esto se realiza en general mediante una tabla con la duración intrínseca de cada segmento y un conjunto de reglas que la modifican: por ejemplo, en español una vocal acentuada es más larga que una no acentuada, un segmento en posición prepausal se alarga en relación a otro que no se encuentre en esta posición, o determinado tipo de consonantes alarga la duración de la vocal precedente. En tercer lugar, el módulo prosódico asigna los movimientos entonativos propios de cada grupo fónico –es decir, de cada unidad situada entre dos pausas– combinando para ello información sobre la modalidad oracional, la posición del grupo en el enunciado y su patrón acentual<sup>12</sup>. Es preciso a continuación iniciar el proceso de concatenación de las unidades elegidas para la síntesis, si se trata de un sistema basado en este principio. Para ello, se recuperan las unidades del diccionario, y se modifican los valores de los parámetros acústicos almacenados para cada una

<sup>11</sup> Vid. Monzón et alii (1993) como ejemplo de un módulo de análisis sintáctico aplicado a la conversión de texto a habla en español.

<sup>12</sup> Diversos aspectos relacionados con la prosodia en la conversión de texto a habla en español han sido abordados por Puigví et alii (1994), Santos et alii (1988) y Marín, (1994), Moreno et alii (1988), Garrido, (1991) y López et alii (1994a).

de ellas en función de las variaciones introducidas por el módulo prosódico. Las unidades –por ejemplo los difonemas anteriormente descritos– se recuperan a partir de su transcripción fonética, y se concatenan mediante un conjunto de reglas que ajustan las modificaciones que deben producirse en las fronteras entre ellas. En un sistema completamente basado en reglas se buscan los valores acústicos de cada segmento y se ajustan en función de las indicaciones del módulo prosódico y de los modelos de transición disponibles.

Finalmente, la cadena de valores de parámetros acústicos –sean valores de formantes o conjuntos de coeficientes LPC– se envían al sintetizador, que los convierte en una onda sonora, resultado final de todo el proceso.

Como puede observarse, en la conversión de texto a habla interviene de forma relevante información fonética y lingüística, generalmente formalizada en las reglas y los diccionarios que componen los distintos módulos. Por ello, es imprescindible en este caso contar con la participación de equipos multidisciplinares en los que colaboren, junto a especialistas en informática y en procesamiento de señal que se ocupan del diseño y la implementación del sistema, fonetistas y lingüistas capaces de proporcionar los conocimientos necesarios para el funcionamiento correcto de los distintos módulos.

#### 2.2.3.2. Las aplicaciones de la síntesis

Cada una de las técnicas descritas anteriormente es apta para determinadas aplicaciones. Hemos mencionado ya que la codificación del habla es adecuada cuando se requiere un conjunto limitado de enunciados y se desea disponer de un método de bajo coste que pueda ser incorporado a otros sistemas. Este es el caso de los «chips hablantes» incorporados a automóviles, máquinas distribuidoras o juguetes. También se usa la codificación en los sistemas de información como los que se encuentran en los transportes públicos o los que ofrecen las compañías telefónicas –información sobre números de abonados, por ejemplo–. Finalmente, los llamados servidores vocales (voice response servers) pueden efectuar las funciones de una operadora controlando el tráfico de llamadas en una centralita telefónica, de modo que sirven de contestador automático, informan de los desvíos de llamadas, indican la forma de dejar mensajes, etc.

Por otra parte, los sistemas de síntesis por concatenación ofrecen un número ilimitado de mensajes y pueden integrarse fácilmente en sistemas informáticos que requieran una salida vocal de inteligibilidad media con un procesamiento no demasiado complejo. Por último, la conversión de texto a habla está disponible ya

en el mercado en forma de productos comerciales, utilizables en ordenadores personales, diseñados para transformar en habla cualquier texto que se encuentre en formato digital. La naturalidad e inteligibilidad de estos sistemas suele ser directamente proporcional a la complejidad del procesamiento, la cantidad de reglas y el tamaño del diccionario de unidades, por lo que en general, puede decirse también que es directamente proporcional a su coste. Algunos de estos sistemas están desarrollados en español (Martínez et alii 1986, Rodríguez et alii 1993, Bullón y Pérez 1994, López Gonzalo et alii 1994b) o en catalán (Martí 1987), y se prevé disponer pronto de prototipos en vasco (Hernández et alii 1994) y en gallego.

La conversión de texto a habla es útil en una serie de aplicaciones relacionadas, en primer lugar, con la consulta telefónica de información almacenada en ordenadores; para el usuario, esto puede traducirse tanto en servicios individuales –como la lectura a distancia del correo electrónico– como en el acceso a servicios públicos de información –como los ofrecidos por los ayuntamientos, farmacias de guardia o cartelera de espectáculos (Gagnoulet et alii 1991)– o las propias compañías telefónicas –noticias, partes meteorológicas, etc.–. Un segundo ámbito en el que la conversión de texto a habla resulta imprescindible es en las ayudas a personas con discapacidades visuales o vocales (Damper 1990). En el primer caso, es posible acoplar un reconocedor óptico de caracteres (Optical Character Recognizer, OCR) a un sistema de conversión de texto a habla para tener acceso a cualquier publicación en papel, o utilizar simplemente un sistema comercial de conversión para leer los textos almacenados en soporte digital. La difusión cada vez mayor de documentos en este formato –tanto en CD-ROM como en el World Wide Web (WWW) a través de Internet– abre nuevas vías de acceso a la información a este tipo de usuarios hasta ahora confinados al Braille o a las cintas previamente grabadas. En personas con discapacidades vocales, la síntesis puede ofrecer una ayuda importante para producir habla con una buena inteligibilidad –el caso paradigmático es Stephen Hawking– a partir de una entrada simbólica o de un texto escrito.

Finalmente, cabe mencionar la utilización de la síntesis en sistemas de enseñanza asistida por ordenador que requieran un refuerzo vocal (Hiller et alii 1993), o como complemento de cualquier sistema informático que proporcione un acceso multimedia a la información.

### **2.3. El reconocimiento del habla**

#### *2.3.1. Características generales de los sistemas de reconocimiento*

Antes de abordar el campo del *reconocimiento del habla* (automatic speech recognition, ASR) es necesario realizar una primera diferenciación entre «recono-

cimiento» y «comprensión». En el campo de las tecnologías del habla, suele trabajarse en el nivel del reconocimiento, es decir, en la conversión de una señal acústica –el habla– en una cadena de símbolos o de caracteres correspondientes a esta señal. Así puede generarse un texto escrito o controlar el funcionamiento de un sistema informático, pero no se consigue la comprensión, si por ello se entiende la extracción de información semántica (Cf. 3.2.2).

Por otra parte, es necesario también distinguir el reconocimiento del habla de la *identificación o verificación de hablantes* (speaker identification / verification) a partir de sus voces (Chollet, 1994). La tarea en este caso es completamente distinta, ya que se trata de decidir automáticamente si la voz de una persona corresponde con alguna de las que se encuentran almacenadas de una base de datos a fin de comprobar la identidad que dice tener o de averiguar si está autorizado para acceder a determinados servicios. Contrariamente al reconocimiento, en este tipo de tarea no se encuentran implicados conocimientos lingüísticos, sino los propios del procesamiento de señal.

Un campo igualmente relacionado con el reconocimiento del habla es la *identificación automática de la lengua* (language identification) que utiliza un determinado hablante. Los sistemas que llevan a cabo esta tarea se basan en el uso de modelos estadísticos sobre la aparición de unidades en el inventario fonético de la lengua, y tienen su aplicación en servicios telefónicos multilingües en los que, a fin de ofrecer la información en la lengua del usuario, debe primero determinarse cuál es la que está hablando.

Al igual que un sistema de síntesis, un sistema de reconocimiento puede caracterizarse mediante una serie de parámetros: el locutor, el vocabulario, el tipo de enunciado y el entorno. A continuación presentaremos sucintamente cada uno de ellos<sup>13</sup>. Si diseñamos un sistema de reconocimiento que permita, por ejemplo, efectuar consultas telefónicas sobre la cotización de divisas, está claro que cualquier persona debe poder utilizarlo; en cambio, un sistema que sustituya al ratón y al teclado en el manejo de un ordenador personal puede estar concebido para un usuario único. Los sistemas del primer tipo se definen como *independientes del locutor*, mientras que en el caso de los segundos se habla de *sistemas dependientes del locutor*. Estos últimos requieren una fase de entrenamiento –en el caso que hemos mencionado consistiría en una lectura de los comandos que aparecen en los menús–

---

<sup>13</sup> Como presentaciones generales del reconocimiento del habla véanse Levinson y Liberman (1981), Casacuberta y Vidal (1987), Ainsworth (1988), Holmes (1988), Pardo (1988), Haton et alii (1991), Rabiner y Huang (1993) y Chollet (1994). Resultan útiles las compilaciones de Bristow (1986), Laface (1990) y de Waibel y Lee (1990).

para adaptarse a las características de la voz y del habla de cada usuario, mientras que tal operación es impracticable cuando consideramos aplicaciones de uso general en las que no es posible prever quién va a acceder a ellas.

Por otra parte, al igual que en la síntesis, podemos plantearnos la flexibilidad del sistema en términos de los elementos de vocabulario que va a ser capaz de reconocer. En las dos aplicaciones mencionadas anteriormente, el vocabulario es obviamente limitado –nombres de las divisas, cifras, algunas preguntas o los textos que aparecen en los menús–, mientras que en un sistema capaz de convertir en texto cualquier enunciado hablado, el vocabulario aumenta y es prácticamente ilimitado. Entre estos dos extremos, existirían casos de vocabulario restringido, como por ejemplo en una aplicación, ya existente actualmente, que permite a los médicos dictar comentarios mientras examinan una radiografía a fin de disponer automáticamente de un informe almacenado en forma de texto en la memoria de su ordenador.

En un sistema de reconocimiento cabe considerar también el tipo de enunciados que deben reconocerse. Podemos pensar que para manejar un ordenador es suficiente con que dispongamos de un reconocedor de palabras aisladas –«abrir», «cerrar», «cortar», «copiar», etc.–, mientras que en un sistema de dictado automático es preciso que el reconocedor sea capaz de tratar enunciados pronunciados de forma fluida o continua. El estado actual de la tecnología no permite aún esta última posibilidad, pero sí permite en cambio el reconocimiento de enunciados en los que se introducen breves pausas entre palabras –habla segmentada–.

Finalmente, debe tenerse en cuenta el entorno del reconocedor. Si se pretende elaborar un sistema para marcar números de teléfono mediante el habla en el interior de un coche, es obvio que el ruido ambiental va a ser importante; lo mismo sucede en el control vocal de un ordenador en cierto tipo de oficina como por ejemplo la redacción de un periódico. Sin embargo, el sistema que utilice un abogado dictando informes en un despacho individual probablemente podrá ser menos resistente al ruido ambiental. En relación con este aspecto, suele hablarse de sistemas robustos cuando están diseñados para operar en entornos muy ruidosos. Los reconocedores que operan a través del teléfono debe ser también suficientemente robustos para no tener interferencias del entorno ni de la propia línea telefónica.

### *2.3.2. Las técnicas en reconocimiento del habla*

#### *2.3.2.1. Estructura general de un sistema de reconocimiento*

Al igual que en la síntesis, en el diseño de un sistema de reconocimiento de habla pueden adoptarse diversas estrategias. También, en este caso, las decisiones

tomadas estarán en función de la aplicación prevista. Sin embargo, existen algunas características comunes a la mayoría de los sistemas de reconocimiento, que se describen brevemente a continuación.

La primera fase en la creación de un sistema de reconocimiento es el *entrenamiento*. En realidad, muchos sistemas funcionan a partir de la comparación entre unidades almacenadas –cuya función es análoga a la de los diccionarios de unidades de síntesis– y los enunciados que produce el hablante usuario del sistema. Por ello, el primer paso suele ser la creación de «plantillas» de referencia con las que comparar los enunciados nuevos. La creación de estas plantillas constituye el proceso al que nos hemos referido como entrenamiento. Para llevarlo a cabo es preciso disponer de un conjunto de realizaciones de la lengua –o corpus de entrenamiento–, y es necesario también haber decidido previamente cuáles van a ser las unidades de reconocimiento. Como hemos visto anteriormente, estas unidades pueden corresponder a cualquiera de los niveles del análisis lingüístico, desde palabras enteras hasta segmentos de señal sonora; es habitual también el uso de difonemas, semisílabas o alófonos. Puesto que la tendencia actual es un entrenamiento con base estadística, las unidades de reconocimiento tienen que estar convenientemente representadas en el corpus.

Las unidades que constituyen las plantillas de referencia deben estar también adecuadamente parametrizadas, es decir, analizadas en términos de su estructura acústica; para ello se utilizan criterios similares a los que se consideran en la síntesis y en el análisis acústico del habla –coeficientes LPC, valores de formantes, características temporales–, aunque la selección de los parámetros se realiza en función del rendimiento que proporcionan en el reconocimiento y de la dificultad de extraerlos automáticamente mediante un análisis de la señal sonora.

Una vez creadas las plantillas de referencia para cada una de las unidades, estamos ya en condiciones de realizar el reconocimiento. La onda sonora producida al hablar es, como se ha indicado al principio, un continuum acústico en el que no es sencillo detectar fronteras entre unidades. El análisis de la señal realizado a la entrada del reconocedor permite detectar aquellas zonas en las que la coincidencia de valores en una serie de parámetros permite localizar las unidades y compararlas con las referencias que se han almacenado en la fase de entrenamiento. Si el sistema trabaja con palabras aisladas, la detección de las mismas y la comparación con los modelos de referencia no es una operación excesivamente compleja y entra dentro de las posibilidades de la tecnología actual. Por el contrario, si se trata de secuencias de habla conectada o continua, es preciso realizar una segmentación en términos de las unidades elegidas para que la comparación pueda llevarse a cabo.

Puede observarse pues que, tal como se indicaba anteriormente, el reconocimiento es un proceso que nada tiene que ver con la comprensión, ya que se trata de comparar una señal de entrada analizada por el sistema y un modelo previamente creado en la etapa de entrenamiento. El resultado de tal operación puede ser una cadena de caracteres que representan el texto correspondiente al enunciado reconocido, o la ejecución de determinados comandos verbales. En cambio, no se llega en ningún caso a la interpretación semántica del mensaje emitido por el hablante.

#### 2.3.2.2. Aproximaciones al reconocimiento del habla

Las etapas de entrenamiento y de reconocimiento expuestas en el apartado anterior constituyen un modelo genérico, en el marco del cual se utilizan diversas aproximaciones en función del tipo de enunciado y de la amplitud del vocabulario que requiera cada aplicación. En conjunto, como se expone a continuación, suelen distinguirse las técnicas aplicadas al reconocimiento de palabras aisladas de las que se utilizan en el reconocimiento de grandes vocabularios.

En el reconocimiento de palabras aisladas se crean patrones o plantillas para cada una de las palabras que forman el vocabulario del sistema. El reconocimiento consiste entonces en la comparación entre ambos elementos, calculando la probabilidad de semejanza entre la palabra que aparece en el diccionario del reconocedor y la que se ha pronunciado realmente. Para ello, es preciso realizar lo que se conoce como una *alineación* o normalización temporal, ajustando la duración de la palabra que debe reconocerse a la del modelo de que dispone el sistema.

El reconocimiento de grandes vocabularios, imprescindible en aplicaciones como el dictado automático, requiere la utilización de otro tipo de técnicas, entre las cuales se encuentra la decodificación acústico-fonética. Mediante este procedimiento, la señal de entrada al sistema es segmentada en elementos fonéticos gracias a un análisis acústico que extrae los parámetros más relevantes para llevar a cabo la segmentación y la identificación de cada unidad. Las unidades utilizadas se caracterizan como unidades subléxicas, ya que son menores que la palabra. A partir de la coincidencia de un número determinado de rasgos fonéticos, se genera una primera transcripción aproximada del enunciado, con diversas alternativas y probabilidades. La decisión final sobre el enunciado reconocido se lleva a cabo en muchos sistemas con la ayuda de lo que en reconocimiento de habla suele denominarse un *modelo de lenguaje*.

Un modelo de lenguaje, simplificando mucho su definición, consiste en un conjunto de probabilidades de transición entre palabras. Para obtenerlas, es preciso



partir de un corpus escrito utilizado como corpus de entrenamiento. Si se desea trabajar a partir de clases de palabras, el corpus debe estar convenientemente etiquetado con esta información para cada una de las palabras. Con este enfoque puramente estadístico, pueden solventarse algunos de los errores que comete el reconocedor en el primer análisis fonético, ya que algunas palabras propuestas como candidatas se descartan en función de la probabilidad de aparecer en un determinado contexto. Es posible también recurrir en este punto a la ayuda de un diccionario que contenga representaciones fonéticas para complementar las decisiones sobre el enunciado finalmente reconocido.

Por supuesto, para el reconocimiento de grandes vocabularios se requiere un conocimiento de la lengua que va más allá del análisis acústico y fonético de la señal sonora y que entra ya de lleno en el ámbito del procesamiento de las lenguas. La incorporación de analizadores sintácticos (Cf. 3.2.2.2) a los sistemas de reconocimiento de habla abre nuevas vías de colaboración entre dos campos que tradicionalmente han estado separados.

Las técnicas de decodificación acústico-fonética complementadas con diccionarios y modelos de lenguaje son también necesarias en el reconocimiento de habla continua. Para este propósito, puede utilizarse además información prosódica –extraída del análisis acústico de la señal– como los movimientos e interrupciones de la curva melódica, que constituyen indicios para realizar la segmentación (Waibel, 1986). La interacción entre la prosodia y la sintaxis constituye también en el caso del reconocimiento uno de los campos en los que se manifiesta una mayor necesidad de integración entre las tecnologías del habla y el procesamiento de las lenguas.

### *2.3.3. Las aplicaciones del reconocimiento*

La gama de aplicaciones del reconocimiento del habla depende, esencialmente, de las posibilidades de los sistemas desarrollados en lo que se refiere a los parámetros discutidos anteriormente. El reconocimiento de palabras aisladas y dependiente del locutor es útil, por ejemplo, en aplicaciones relacionadas con el control de sistemas informáticos, máquinas, herramientas, robots, o para la introducción de datos mediante el habla. En el terreno del control de sistemas, se han llevado también a cabo experiencias en el uso de reconocedores de habla para el pilotaje de aviones especialmente en contextos militares; en estos casos, el sistema no sólo debe poder actuar con un elevado ruido ambiental, sino que debe poder responder a las alteraciones que las situaciones de estrés provocan en la voz.

El reconocimiento independiente del locutor, en cambio, es importante en todas aquellas utilidades que pueden proporcionarse a través del teléfono: consultas

a servicios de información, transacciones bancarias, reservas de billetes, reservas de fechas para determinadas actividades como la revisión de vehículos, etc. Aunque en la actualidad muchas de estas aplicaciones se basan en el reconocimiento de dígitos aislados o conectados que se utilizan como un menú de servicios disponibles, algunas de ellas requieren una interacción compleja entre el usuario y el sistema, tema que se aborda en el apartado 4.1.

Sin embargo, tal vez la aplicación más ambiciosa del reconocimiento es el dictado automático o lo que también se ha denominado la máquina de escribir activada por la voz (Baker 1993a, 1993b; Steinbiss et alii 1995). Se trata de conseguir que el texto dictado por el usuario se guarde en forma escrita y pueda ser posteriormente editado utilizando un procesador de textos. Actualmente existen productos comerciales desarrollados para alcanzar este objetivo, algunos de ellos para el español como el de Dragon Systems y el Voice Type de IBM. Tales sistemas necesitan un entrenamiento por parte del usuario, y requieren que se realicen breves pausas entre palabras.

Cabe mencionar que las técnicas de reconocimiento de habla están siendo utilizadas en sistemas de enseñanza de lenguas extranjeras asistida por ordenador (Janot-Giorgetti 1983), ya que pueden proporcionar para un conjunto de enunciados previamente determinado en los sistemas actuales una medida del grado de diferencia entre la pronunciación nativa y la del estudiante. Finalmente, los sistemas de reconocimiento pueden proporcionar medios de controlar sistemas mediante el habla a personas con discapacidades físicas (Fisher 1986).

### **3. EL TRATAMIENTO DEL TEXTO**

#### **3.1. El tratamiento del texto: problemas**

##### *3.1.1. El procesamiento de la lengua escrita*

El lenguaje humano es, en esencia, un medio de comunicación oral. Pero las limitaciones naturales de una señal de desvanecimiento tan rápido como el habla desencadenaron la aparición de medios de transmisión alternativos, del tipo de la escritura. Con este precedente, las dificultades que el habla plantea a un tratamiento informatizado (Cf. 2.1) nos pueden llevar a creer que el texto es el cauce natural para el procesamiento de las lenguas. Sin embargo, su condición secundaria frente al discurso oral y las características propias de la lengua escrita confieren al procesamiento del discurso textual problemas específicos. En primer lugar, en todas las lenguas es habitual la homografía, esto es, la identidad en la escritura de signos lingüísticos diferentes. En español, por ejemplo, «canto» o «ama» pueden ser

sustantivos o verbos, y en lenguas con una morfología más pobre, como el inglés, los casos se multiplican: la forma «cut» puede corresponder a un sustantivo, un verbo o un adjetivo y la forma «right» puede encerrar un sustantivo, un adjetivo o un adverbio. Como es de esperar, este sincretismo en las formas léxicas de las lenguas produce enormes problemas a la hora de traducir, interpretar o generar automáticamente un texto escrito.

Y las dificultades no acaban ahí. Incluso si no hay dudas sobre la categorización adecuada para una forma escrita, la polisemia o pluralidad de significados que podría albergar exige el uso de reglas contextuales muy finas. En español el verbo «pegar» puede significar ‘azotar’ o ‘adherir’ y, sólo el esquema sintáctico en que se incluya dará la clave para resolver el conflicto: «pegar» seguido de un complemento indirecto significa ‘azotar’; seguido de un complemento directo, ‘adherir’. Un sistema automático sólo podría interpretar verbos de este estilo estableciendo una primera versión sobre su significado básico y volviendo luego atrás para efectuar las modificaciones oportunas. Así el verbo inglés «to grow» equivale al español «crecer» pero, si el sistema se encuentra con que va seguido de un complemento directo animado, deberá traducirlo al español por «criar»; si ese complemento directo fuese inanimado y del tipo planta, «to grow» se identificaría con «cultivar». Ejemplos como éstos explican la evolución de la lingüística computacional. Aunque en los comienzos se optó por la estrategia ingenua de tratar un texto palabra por palabra, en seguida se comprendió que el análisis y la delimitación de las relaciones sintácticas latentes en él era un paso ineludible.

Por otra parte, las lenguas humanas son códigos afectados de fuertes dosis de imprecisión. En realidad, un código dispuesto para el intercambio cotidiano no necesita ser muy preciso. Si se admite que el fin último del lenguaje es la comunicación, se comprenderá que el hablante urgido por la necesidad de comunicarse tiende a eliminar de sus mensajes lo que le parezca accesorio. Toda la información que su interlocutor pueda deducir del contexto o la situación en que se produce el discurso, o todo lo que se desprenda de la experiencia estará ausente de los mensajes lingüísticos. Pero al aligerar el mensaje, el hablante lo hace más impreciso y se arriesga a propiciar la colisión de interpretaciones distintas. Este tipo de imprecisión, denominado ambigüedad, no es más que un accidente que tiene su origen en las imperfecciones del simbolismo utilizado, pues «ambiguo» se opone, de algún modo, a «unívoco». Una de sus caras más frecuentes en las lenguas es el sincretismo o capacidad de un significante para canalizar dos significados divergentes, como en el caso de las homografías ya vistas. No obstante la ambigüedad puede tener mayor alcance. Así, de esa dimensión terminológica se pasa a una ambigüedad estructural. Un caso típico es el de cláusulas como «La tirada

fue reducida», capaces de fomentar una lectura pasiva –si se entiende que el editor redujo la tirada de una edición– y otra atributiva –si, simplemente, se trata de consignar que la tirada contó con un escaso número de ejemplares–.

Ante un discurso ambiguo, el receptor indaga en el contexto la información necesaria para decodificarlo convenientemente. Pero para los tratamientos computacionales el problema radica en construir sistemas capaces de acceder al contexto. Además de las dificultades técnicas para lograr ese objetivo, hay que valorar el hecho de que buena parte de esas ambigüedades se resuelven en el procesamiento humano de la información a partir de conocimientos extralingüísticos. Una secuencia como «Todas las jovencitas están enamoradas de un cantante» permite suponer que todas aman a uno en particular, o bien que todas aman a alguno sin que sus preferencias tengan que coincidir absolutamente en la misma persona. Aunque la primera interpretación es perfectamente legítima, la segunda es la que cualquier hablante tiende a dar por buena en condiciones no-marcadas. La elección, sin embargo, no depende del contexto sintáctico o semántico sino, más bien, del conocimiento del mundo que nos da una cierta lógica de la vida cotidiana. Por supuesto, el modo en que efectuamos esos razonamientos ordinarios es difícil de formalizar. Pero difícil no significa imposible. Y de ahí que la investigación más reciente sobre el tratamiento computacional del texto se oriente hacia análisis pragmáticos, provistos de depuradas modelizaciones del conocimiento que el ser humano tiene sobre el mundo que le rodea.

### 3.1.2. *El análisis computacional del texto escrito*

La mayor parte de los productos generados por las industrias de la lengua escrita requieren que los sistemas informáticos proyecten sus análisis en los distintos estratos del estudio lingüístico. En primer lugar, a partir de las formas gráficas pueden inferirse unidades fonológicas. El éxito de la tarea depende de la posibilidad de proporcionar al sistema reglas explícitas que establezcan esa correspondencia. En español, por ejemplo, el asunto es relativamente fácil, pero el grado de complicación aumenta en lenguas donde la escritura se haya alejado más del principio fonémico. Aunque el campo de aplicación de estos análisis se dirige prioritariamente a los sistemas de conversión de texto a habla (Cf. 2.2.3.1), pueden ser también bastante útiles en áreas como la métrica.

Una vez conseguido el elenco de palabras del texto, una investigación puede exigir que se atienda, no ya a su forma gráfica, sino a los lexemas o lemas<sup>14</sup> a los

<sup>14</sup> Aunque guarden cierta similitud, los términos *lexema* y *lema* no pueden verse, en absoluto, como sinónimos. Mientras que el lexema es una unidad morfológica reconocida -v. gr. un tipo de morfema

que remiten las variantes morfélicas. Así, el lexema [tranquil-] aparece en el sustantivo «tranquilidad», en el adjetivo «tranquilo», en el también adjetivo «tranquilizador» o en el verbo «tranquilizar». El proceso de agrupación de todas las variantes de una unidad léxica, la lematización, es difícil de automatizar, dada la dificultad de especificar reglas suficientemente explícitas y generales para explicar la formación de palabras. La lematización es una de las cuestiones más problemáticas para el análisis computacional del texto y la mayor parte de los proyectos de lexicografía computacional han tenido que combinar análisis automáticos y procedimientos manuales. La técnica más usual consiste en lematizar un texto manualmente y, al mismo tiempo, almacenar ese resultado en un diccionario automático. En una fase posterior ese diccionario se usará para lematizar un segundo texto, añadiendo los lemas que no estuvieran incluidos en la primera versión. De este modo se va enriqueciendo el diccionario inicial.

También a partir de la secuencia de letras y signos de puntuación los sistemas informáticos deben establecer los límites de las unidades sintácticas y la clasificación de las palabras y de las unidades de alto nivel. Este objetivo pasa por formalizar detalladamente la organización de la lengua en cuestión mediante una gramática automática. Por supuesto, no es sencillo confeccionar un programa que distinga las distintas relaciones presentes en (1) y (2), habida cuenta de la similitud de sus unidades:

- (1) [[ *la respuesta a [la exageración de Juan]*]]
- (2) [[ *la tendencia a la exageración] de Juan*]]

Por fin, los sistemas computacionales exigen un tratamiento semántico, que examina el contenido del texto. La aproximación más simple consiste en producir una lista de palabras y, de su análisis cuantitativo, deducir el tema del texto, con la idea de que los conceptos más frecuentemente mencionados serán indicativos del tema general. Por supuesto, la investigación debe afrontar en este sentido problemas como los ya mencionados de homografía o ambigüedad, además de tener en cuenta la posible variación semántica en función del contexto.

### 3.1.3. Acopio de datos

La recopilación masiva de datos ha sido una constante histórica en el desarrollo de la lingüística descriptiva. Sin embargo, con la llegada de los

---

dotado de significado léxico-, el lema no tiene ese estatus en gramática. Se trata, simplemente, de una etiqueta introducida por las aplicaciones computacionales con la que se hace referencia a las distintas formas flexivas que puede tomar una palabra.

ordenadores, esta técnica de trabajo acabó consolidando una disciplina independiente, la llamada *lingüística de corpus*, de cuyo desarrollo dependen buena parte de las aplicaciones computacionales. Esta orientación empírica nace como una reacción contra la excesiva importancia que la gramática generativa había dado a la introspección del lingüista, considerada como el único recurso útil para validar las propuestas emanadas de la pura teoría.

Desde un punto de vista metodológico, los productos generados por la lingüística computacional pueden dividirse en dos grandes grupos. Por un lado, se sitúan aquéllos relacionados con el procesamiento de las lenguas –síntesis y reconocimiento del habla, traducción automática o generación y comprensión automática de textos–. Sea en la vertiente oral o en la textual en todos estos casos el producto resulta de un trabajo lingüístico estrechamente vinculado a los procedimientos de la inteligencia artificial. Por otro lado, hay que contar con el uso de sistemas computacionales para dar forma a los textos en áreas como la estilística, la lexicografía, la edición textual, la enseñanza de lenguas o la propia lingüística descriptiva. Pero tanto en los productos –indudablemente más llamativos– del primer grupo, como en estas aplicaciones «menores» del segundo grupo la organización de los datos y el recuento de apariciones de formas particulares cobran una singular importancia. Desde luego, ambos procesos pueden ser realizados por el ser humano pero el ordenador aporta su enorme velocidad –millones de cálculos por segundo en una máquina poderosa–, su capacidad de almacenamiento y su grado de fiabilidad.

A partir de recursos gráficos mínimos, como la separación entre palabras, el ordenador produce listas de palabras, índices y concordancias. Las listas de palabras proporcionan la frecuencia de cada forma, ordenada según el criterio más conveniente<sup>15</sup>. Los índices aportan como información adicional una referencia sobre el lugar –obra, página, línea– en que se consigna cada aparición. Finalmente, las concordancias añaden a las informaciones anteriores un pequeño contexto. En un principio, la lingüística de corpus se contempló como una valiosa fuente de información sobre la frecuencia y distribución de las unidades léxicas o el tamaño de las cláusulas. Pero los problemas metodológicos derivados de una óptica que trabaja con datos en bruto no se harían esperar; la falta de información gramatical impedía distinguir las homografías o trabajar con unidades polisémicas<sup>16</sup>. Será entonces, a medida que se persevere en la tarea de depurar sus procedimientos,

<sup>15</sup> En este sentido, incluso criterios aparentemente tan caprichosos como el orden alfabético invertido han prestado una ayuda de inestimable valor en morfología o estilística.

<sup>16</sup> Junto a estas consideraciones teóricas, algunos aspectos prácticos desempeñan un papel relevante en el proceso de recopilación de datos. Aunque no haya espacio aquí para tratarlas con más detalle,

cuando la lingüística de corpus adquiriera una notable pujanza. Con una perspectiva ligeramente superior a la que se observaba en las primeras aplicaciones, la lingüística de corpus se ha convertido en la vía idónea para describir las lenguas teniendo presente el uso real que los hablantes hacen de ellas. Los proyectos de corpora – los más importantes realizados sobre inglés, francés, alemán, italiano y holandés– han producido gramáticas que disponen de amplios conjuntos de materiales procedentes de textos reales, lo que permite llevar a cabo descripciones bastante afinadas de los fenómenos lingüísticos registrados en cada una de esas lenguas y elaborar teorías gramaticales mejor fundamentadas que las precedentes.

#### *3.1.4. Sistemas de anotación*

La necesidad de completar los datos almacenados con algún tipo de información gramatical motiva la aparición de corpora anotados. En síntesis, un sistema de anotación no es más que un conjunto de convenciones por las que se asocia cada palabra de un corpus a la información gramatical y semántica que pueda ser relevante para su análisis una vez que se ha extraído del texto.

El desarrollo de la lingüística de corpus ha sido realmente importante en la última década, dado que una serie de innovaciones técnicas permiten hoy manipular, almacenar y ordenar gran cantidad de datos con bajo coste y a gran velocidad. Los problemas de espacio han ido superándose con la introducción en el mercado del soporte CD-ROM que facilita el acceso a ordenadores personales y aumenta el interés comercial de los productos resultantes. Con este auge, el terreno ha experimentado una importante expansión de sus intereses, objetivos y métodos iniciales. Si en la versión tradicional se trata, simplemente, de una técnica que busca la corroboración empírica de las teorías, últimamente el corpus ha comenzado a contemplarse como una fuente de información con la que obtener estrategias de procesamiento, ya que muchas de las aplicaciones de la lingüística computacional –por no decir todas– dependen de su desarrollo. Esta evolución, sin embargo, ha producido algunos motivos de inquietud. Desde el punto de vista de la lingüística, la tendencia observada en los últimos años parece orientarse exclusivamente a la rentabilidad de los sistemas computacionales, aún si esa rentabilidad se consigue a expensas del conocimiento lingüístico. Uno de los efectos más llamativos de este creciente interés por las estrategias de procesamiento es el desmedido crecimiento del tamaño de los corpus. Un corpus clásico, como el LOB (Lancaster-Oslo-

---

cuestiones como los derechos de propiedad, la capacidad de almacenamiento del sistema, el tratamiento de la puntuación y la ortografía o la posibilidad de corregir errores no son en absoluto triviales.

Bergen), contiene un millón de palabras, mientras que la tendencia de los últimos años es construir corpora de varios millones. En realidad, el tamaño del corpus viene condicionado por la orientación a que se dedica. Los lexicógrafos, por ejemplo, precisan grandes cantidades de datos para observar la presencia de neologismos y, en general, establecer las pautas que rigen el cambio en las lenguas. Para este fin, por ejemplo, el Birmingham Main Corpus compilado por el diccionario COBUILD (Sinclair et alii 1987) con sus veinte millones de palabras se ha quedado pequeño. La creación del British National Corpus (Leech et alii 1994) sugiere que cuanto mayor es un corpus, mejor resulta para estos fines. Pero, a medida que crece el tamaño del corpus, aumentan los problemas de anotación. La enorme inversión en tiempo y en esfuerzo humano que estos proyectos suponen obliga a reducir las anotaciones. Los corpora destinados a aplicaciones en el procesamiento de las lenguas acusan aún más esa pérdida de detalle porque sus anotaciones cumplen únicamente fines estadísticos. El peligro es que el volumen de esfuerzo invertido no se compense con un progreso del conocimiento y, en definitiva, que esos enormes listados dejen de ser interesantes para la teoría lingüística.

### 3.1.5. *Formalismos gramaticales*

El enfrentamiento entre los requisitos de la teoría lingüística y las exigencias de eficacia propias de los sistemas computacionales es un tema recurrente en este campo. La lingüística computacional es, al fin, una disciplina de corte aplicado y por tanto, no sólo sus objetivos se orientan a la consecución de productos materiales, sino que, además, debe lograrlos en un plazo de tiempo razonable.

La ingeniería informática denuncia con frecuencia la debilidad de las herramientas aptas para procesos automáticos pero los lingüistas ponen en entredicho si, ya en este estadio, se está haciendo investigación lingüística y critican el recurso a soluciones locales. Como indica Rouault (1989:146), el problema fundamental que debe afrontar la formalización procede del propio estado de nuestros conocimientos. Sabemos muy poco de los mecanismos con los que el cerebro humano produce o comprende el lenguaje, así se explica que no podamos simular el proceso automáticamente. Obviamente, el tratamiento automatizado de las lenguas debe surgir de un cuidado equilibrio entre teoría lingüística y formalización.

Quizá uno de los más serios obstáculos para el progreso de la lingüística computacional viene dado por el escaso interés de buena parte de las teorías lingüísticas en producir gramáticas que puedan ser procesadas. En los lenguajes de la lógica se planteó, hace ya algunas décadas, la posibilidad de que la interpretación



semántica de los enunciados pudiera seguirse de un conjunto de reglas que se aplicaran directamente al análisis sintáctico. La expansión de este método a los sistemas informáticos no constituyó, sin embargo, una línea de investigación fecunda porque los instrumentos de la lógica se orientaban exclusivamente a captar ciertos tipos de argumento. En general, el papel de los formalismos gramaticales en este desarrollo ha sido subestimado. A ello contribuyó que la gramática generativo-transformacional, bajo la influencia de Noam Chomsky, destacara el papel de la sintaxis como disciplina independiente y se concentrara en estudiar la gramaticalidad o correcta construcción de las cláusulas al margen de su contenido semántico. El modelo más poderoso de interpretación semántica surge a partir de los trabajos de Montague (1974), que desarrolló una lógica intensional y formuló reglas precisas para la interpretación del inglés, particularmente para los cuantificadores.

Los productos resultantes de la aproximación computacional son a menudo una mezcla heterogénea de procedimientos que recurren a modelos lingüísticos distintos. Sólo en los últimos años parece notarse una cierta tendencia a seguir de modo más estricto y coherente las orientaciones de determinadas líneas teóricas. El declinar de los modelos de inspiración generativa ha ido parejo a una creciente alza de los modelos lexicalistas y de las semánticas formales. Aunque esta dirección sea prometedora, su avance sólo será posible cuidando al máximo las exigencias de rigor. La formalización puede ser un requisito imprescindible para los sistemas automáticos pero también tiene un precio: muchas de sus leyes no se ven sustentadas por una fuerte motivación empírica, sino por la coherencia interna del propio sistema formal. Como contrapeso, una formalización estricta puede hacer surgir nuevos problemas o reconsiderar algunos análisis de fenómenos ya conocidos. Además, los formalismos gramaticales constituyen el único medio para asegurar el progreso de las herramientas computacionales en lingüística. La lingüística de corpus, por ejemplo, nos ha enseñado que para obtener una muestra de determinada lengua suficientemente representativa de su uso real por parte de los hablantes es necesario almacenar ingentes masas de datos. Si un corpus, por definición, debe dar cuenta del mayor número posible de entradas léxicas de una lengua, esa amplitud todavía se dilata al observar la enorme cantidad de textos que se deben recoger para superar un corto inventario constituido por las palabras más frecuentes. Con tales problemas, la posibilidad de conseguir sistemas de anotación y análisis automáticos no debe descuidarse. En los últimos años se han obtenido importantes progresos en esta dirección. Un equipo de la Universidad de Nimega dirigido por Kees Koster ha desarrollado el formalismo AGFL (Affix Grammar over a Finite Lattice) y el grupo ChLoE del INRIA (Institut National de Recherche en Informatique et en Automatique) dirigido por Bernard Lang ha desarrollado el formalismo LPDA

(Logic Push-Down Automata) que ha conseguido excelentes resultados a la hora de evitar los problemas de completitud y corrección presentes en la mayoría de los sistemas de tratamiento automático de las lenguas<sup>17</sup>.

### 3.2. El tratamiento del texto: productos

#### 3.2.1. Traducción automática

##### 3.2.1.1. Breve historia del campo

El primer contacto entre dos mundos aparentemente ajenos como son informática y traducción se produce a mediados de siglo, en un intento de dar al ordenador horizontes más amplios que el terreno militar en el que se encontraba confinado. A finales de los años cuarenta Warren Weaver, un matemático americano interesado en teoría de la comunicación, propone utilizar las técnicas del desciframiento criptográfico para traducir textos mecánicamente<sup>18</sup>. Nace así la traducción automática, un mal nombre para una actividad que todavía no ha logrado la solvencia que el adjetivo promete. De hecho, la complejidad de la información contenida en las lenguas, su rica organización interna y su diversidad invita a considerar altamente improbable el que algún día llegemos a conseguir traducciones de alta calidad mediante procesos totalmente automatizados. Pero el término *traducción automática* se usa aún hoy, a pesar de las connotaciones que arrastra sobre el optimismo con que se abrazaron las nuevas tecnologías<sup>19</sup>. En cualquier caso, la propuesta no era consecuencia del avance tecnológico, pues los ingenios informáticos disponibles carecían de la operatividad requerida para ejecutar traducciones aceptables. Al contrario, era la necesidad la que imponía esta estrategia. Se estaba descubriendo el valor de la información, un bien propio del siglo XX, y con su trasvase automático de una lengua a otra se buscaba reducir costes en material y tiempo. En 1952 se celebra en el Instituto de Tecnología de Massachusetts el congreso que inauguraría todo un nuevo campo de trabajo. El entusiasmo con que fueron acogidos los primeros ensayos en esta dirección se explica en el contexto político de la guerra fría y la carrera armamentística. De hecho, el primer sistema piloto en Estados Unidos, el G.A.T. –Georgetown Automatic Translation–, se

<sup>17</sup> Vid. Rodríguez et alii (1995).

<sup>18</sup> A pesar de la aparente originalidad de Weaver, en realidad la traducción no es una actividad inusual para un ordenador. Siempre que un programador da una determinada instrucción a la máquina lo hace a través de algún lenguaje informático como BASIC o PROLOG que resulta más cómodo que la verdadera orden en lenguaje de máquina, limitada a una secuencia de ceros y unos.

<sup>19</sup> Vid. Niremburg (1987), Slocum (1988), de Sopeña (1990) y Lewis (1992) para una presentación general.

dedicaba exclusivamente a la traducción de textos científicos del ruso al inglés. Las esperanzas suscitadas por estas primeras investigaciones se desvanecieron rápidamente porque el proyecto se había concebido desde planteamientos excesivamente ingenuos. En opinión de sus promotores bastaría con dar cabida en el ordenador a diccionarios de cierta sofisticación, que además realizasen algunos procesos de ordenación y reajuste, de acuerdo con las peculiaridades internas de las lenguas concernidas.

Ante esta simplicidad conceptual los resultados son muy pobres. En los años sesenta, se comienza a cuestionar las posibilidades de éxito de la empresa y, en general, se acusa la excesiva precipitación con que se estaba abordando el asunto. El problema de fondo radica en la propia naturaleza de los sistemas utilizados en este período, la llamada primera generación en traducción automática. Se trata de complejos informáticos que pretenden conseguir una traducción aceptable trabajando palabra por palabra a partir de un diccionario, sin construir una representación sintáctica ni semántica de los mensajes y sin considerar las pertinentes diferencias entre las lenguas que manejan. La reflexión lingüística quedaba marginada de estos sistemas porque la investigación contaba únicamente con especialistas en otros campos, fundamentalmente matemáticos e ingenieros informáticos.

En 1965 las agencias gubernamentales que habían financiado la investigación durante veinte años exigen la formación de un comité de expertos que evalúe los progresos obtenidos hasta el momento. El resultado es el famoso informe ALPAC –Automatic Language Processing Advisory Committee–, que vertía resultados desoladores. En primer lugar, criticaba la escasa calidad y el alto coste de unas traducciones que precisaban importantes revisiones. Pero, además, llegaba a cuestionar la viabilidad y la promoción futura de semejantes proyectos. Como consecuencia las inversiones decayeron hasta el punto de que algunos equipos tuvieron que abandonar su tarea. La innovación quedaba relegada a marcos universitarios donde parecía adquirir dimensiones puramente especulativas. Visto desde la perspectiva actual, el informe era oportuno en la medida en que denunciaba las deficiencias de un modo particular de llevar a cabo el proyecto. Pero eso no significa que el proyecto mismo tuviera que desestimarse, sino que debería comenzar de nuevo su andadura a partir del marco teórico de la lingüística computacional<sup>20</sup>. Durante los diez años que siguieron a la publicación del informe ALPAC apenas

---

<sup>20</sup> Frente a la ingenuidad en que se mueve la traducción automática en el contexto americano, en esa misma época dentro de la antigua URSS aparecen proyectos dotados de una mayor carga teórica. El problema que abordan los investigadores soviéticos es el de la construcción de una lengua intermedia universal a partir de modelizaciones lógicas o matemáticas. Esta tendencia, representada por lingüistas como I. Mel'cuk, anuncia los sistemas de segunda generación.

unos cuantos grupos aislados pudieron mantenerse en torno a una empresa que parecía abocada al fracaso. En ese momento, una serie de acontecimientos independientes van a asegurar, azarosamente, la continuidad de la investigación. De un lado, Noam Chomsky formulaba una metodología lingüística que prometía dar el tratamiento informático adecuado para las lenguas. Un conjunto de reglas, al estilo de la lógica simbólica, parecía condensar la gramática y dotarla de la cobertura formal exigida por el ordenador. Del otro, aparecen nuevos lenguajes informáticos, como ALGOL y LISP, mejor dotados que sus predecesores para trabajar con lenguas. Por fin, un acontecimiento de otra índole, la guerra de Vietnam, iba a fomentar el interés por traducir las lenguas del sudeste asiático, escasamente conocidas a causa de la distancia geográfica y genética que las separa de Occidente y de su nula repercusión en el mercado mundial. Con una mínima apoyatura teórica y a la sombra de la Secretaría de Defensa de Estados Unidos, surge la segunda generación de técnicas de traducción automática, representada por los sistemas METAL y SYSTRAN en Estados Unidos, el TAUM-METEO en Canadá o el GETA en Francia. El esquema general seguido por todos ellos supone la superposición de etapas de trabajo sobre el eje de una lengua-puente. En una primera fase, un mensaje escrito en la lengua objeto de traducción (lengua-fuente) es almacenado y convertido a un estadio intermedio, más o menos artificial, que acoja tanto los significados referenciales como las peculiaridades formales del sistema. Este puente permitirá, en una segunda fase, recorrer el camino inverso, simplificando los problemas de versión a la lengua a la que se quiere traducir (lengua-destino). La técnica, que ha admitido distintos tipos de perfeccionamiento conceptual, admite la intromisión de otras variadas fases, sin excluir el tratamiento desde el exterior por parte de lingüistas y traductores. De hecho, el grado de intervención externa entabla relación directa con la calidad final del producto.

La segunda generación de sistemas de traducción automática deja entrever una fuerte toma de conciencia sobre la naturaleza de los problemas inherentes a la traducción. Los investigadores parecen ya convencidos de la necesidad de realizar análisis diferenciados de las lenguas fuente y destino, así como de proyectar la traducción a un nivel más profundo que en la etapa anterior, comenzando por un análisis sintáctico que después se enriquece con conocimientos semánticos y contextuales. En un segundo plano, debe constatarse también que el fracaso anterior impulsa un modelo más realista de investigación. En lugar del ambicioso plan de traducir una lengua a otra, se plantean proyectos restringidos, que traduzcan textos técnicos relativos a un dominio bien delimitado. Esta estrategia aminora el vocabulario –un objetivo interesante en sí mismo habida cuenta de las limitaciones técnicas de los sistemas de la época– pero, y esto es más importante, también

disminuye la combinatoria gramatical y, en consecuencia, la ambigüedad. Sin embargo, los conocimientos lingüísticos van estrechamente vinculados al tipo de texto, de modo que estos sistemas son únicamente válidos para traducir grandes cantidades de materiales homogéneos, como manuales de instrucciones. El mejor ejemplo de esta restricción a dominios locales es el sistema TAUM-METEO que traduce exclusivamente boletines meteorológicos para las dos comunidades lingüísticas aglutinadas por el estado de Canadá.

En los últimos veinte años el campo ha experimentado una nueva vitalidad alentada por la creciente necesidad de traducciones en las empresas públicas y privadas y por el progreso de las herramientas informáticas. Los sistemas de segunda generación perfeccionados cuentan con lenguajes de programación más sofisticados, capaces de acceder de forma recursiva a rutinas que detectan caminos erróneos y vuelven atrás para tomar una nueva ruta. El sistema ARIANE, desarrollado por GETA (Francia), consigue traducciones aceptables del ruso o del alemán al francés, y METAL (Siemens, Alemania) trabaja en ambas direcciones con alemán, inglés, francés y español. Los años ochenta supusieron el lanzamiento de proyectos nacionales e internacionales en este dominio, como EUROTRA, que surge en 1982 para todas las combinaciones posibles de lenguas oficiales de la Comunidad Europea. En Japón, el proyecto MU incrementa la traducción automática con sistemas de visualización multilingüe, bancos de datos terminológicos, síntesis y reconocimiento del habla para la elaboración de teléfonos traductores. Esta vía ya anuncia una tercera y una cuarta generación de sistemas de traducción.

### 3.2.1.2. Traducción automática frente a traducción asistida

Puesto que ningún sistema de traducción automática puede, al menos en el estado actual de la investigación, prescindir de toda intervención humana, se ha ido imponiendo la denominación de traducción asistida. Sin embargo, habría que distinguir, de acuerdo con el peso que adquiere cada uno de los agentes, entre (i) traducción (humana) asistida por el ordenador y (ii) traducción automática asistida por el ser humano. Aunque en ambos casos se registra la participación humana, ésta tiene lugar en distintas fases del trabajo. La *traducción asistida por ordenador* requiere, durante el propio proceso de traducción, la intervención de un traductor que tome las decisiones más comprometedoras –por ejemplo, solventar ambigüedades o confirmar el uso de una estructura poco frecuente–. En definitiva, la traducción asistida por ordenador implica un sistema interactivo que, al encontrar cualquier dificultad, reclama al ser humano para que la resuelva. En formas más limitadas, se habla también de traducción asistida por ordenador en referencia a una

serie de productos de ayuda técnica para la traducción, en especial, los bancos de datos terminológicos bilingües o plurilingües, aunque limitar la técnica a esta somera intervención de la máquina no parece realmente apropiado<sup>21</sup>. En cambio, se habla de *traducción automática* (asistida por el ser humano) cuando el sistema exige una fase previa de preparación del texto en lengua-fuente o una revisión del texto de salida en lengua-destino. Por supuesto, existen notables diferencias en lo que respecta a la naturaleza y cantidad del trabajo humano necesario en estas colaboraciones. La preparación del texto puede limitarse a simples anotaciones con las que se marcan los límites de cada estructura sintáctica o suponer una nueva redacción. Del mismo modo, la revisión va desde los retoques tipográficos y las correcciones de estilo hasta la eliminación de contrasentidos.

El hecho de que el traductor esté siempre presente no debe interpretarse, en ningún caso, como un síntoma del fracaso de la empresa. Situando las cosas en su lugar, el traductor es un especialista en la tarea y, por tanto, un componente imprescindible, aún cuando su papel se limitase en un estadio ideal al de mero supervisor. Nadie espera que los sistemas de traducción automática funcionen sin la intervención de un ingeniero informático, así que prescindir del traductor sería tanto como relegar a un segundo plano, una vez más, el papel de la lingüística en la tarea. Como línea de investigación, la traducción automática no precisa hoy en día justificación alguna pero, al menos en ciertos círculos, todavía se ataca su rentabilidad práctica. En este sentido conviene apuntar que cualquier valoración de los logros en este campo debe tener en cuenta que las necesidades de traducción son mayores cada día y que el coste de los traductores humanos es considerable. Los avances que experimente la investigación en esta línea tienen una aplicación inmediata garantizada.

Por otro lado, la selección de una de las dos alternativas presentadas –traducción asistida por ordenador o traducción automática asistida por el traductor–, depende de las necesidades que la traducción deba satisfacer. Obviamente, la traducción de textos técnicos o científicos persigue como fin primordial la inteligibilidad de un producto cuyo contenido debe conocerse con exactitud y rapidez y cuyo estilo importa bastante menos. Este tipo de producto, la traducción para el especialista, constituye el marco idóneo para la traducción automática. Los textos administrativos o jurídicos, así como los manuales de instrucciones tienen exigencias más fuertes y pueden ser abordados por un sistema automático, siempre que en una

---

<sup>21</sup> Los bancos terminológicos son bases de datos textuales que almacenan informaciones sobre el uso de las palabras destinadas específicamente al traductor. Como ejemplo podría citarse EURODICAUTOM, la base de datos terminológica de la CE, o el TERMIUM de Montréal.

etapa posterior sean revisados por un traductor. Es el campo de la traducción para el revisor<sup>22</sup>. Obviamente, la revisión no exige el mismo esfuerzo ni el mismo tiempo<sup>23</sup> que la traducción completa y la máquina habría simplificado el proceso y el precio final del producto, además de librar al ser humano de las fases más repetitivas y menos motivadoras de su trabajo. Por fin, los textos literarios exigen una calidad formal hoy inalcanzable con medios mecánicos, por lo que reclaman una traducción para el traductor, es decir, un uso moderado de las herramientas de traducción asistida que haga del especialista humano el autor inequívoco del producto.

### 3.2.1.3. Organización de los sistemas de traducción automática

En términos generales, cualquier sistema de traducción automática debe analizar el texto de entrada y generar su equivalente en una lengua distinta. Eso implica un trabajo lingüístico amplio, que abarca todos los estratos del estudio. Se trabaja en morfología para identificar unidades mínimas, en sintaxis para delimitar los grupos de palabras que funcionan en bloque y sus relaciones con el todo de que forman parte y, por fin, se trabaja en semántica para consignar la naturaleza de esas relaciones y determinar el significado del texto. Estos pasos se cubren a partir de los datos lingüísticos contenidos en diccionarios o en gramáticas. Habitualmente el trabajo en cada estrato se estructura jerárquicamente, como en EUROTRA, pero en algunos casos las informaciones morfológica, sintáctica y semántica interactúan en una perfecta cooperación –sistemas ARIANE, METAL o LOGOS–, de modo que las ambigüedades sintácticas se resuelven a partir de los conocimientos semánticos. En conjunto, todos los sistemas de traducción automática se componen de varios diccionarios, un analizador y un sistema de conversión de la lengua-fuente a la lengua-destino.

El funcionamiento de los sistemas de traducción automática depende, en buena medida, de una serie de diccionarios perfectamente articulados. De hecho, el diccionario era la única herramienta lingüística de los primeros sistemas. En principio, las palabras del texto de entrada se ponen en relación con la información gramatical contenida en un diccionario monolingüe, a fin de facilitar el proceso de análisis que va a tener lugar. Aunque existen fuertes variaciones en este punto, los diccionarios monolingües usados en esta primera fase de la traducción suelen

---

<sup>22</sup> Algunos sistemas de segunda generación, como LOGOS o METAL han sido específicamente concebidos para hacer traducciones para el revisor.

<sup>23</sup> En la actualidad, un sistema de traducción para el especialista traduce unas diez páginas por hora y un sistema de traducción para el revisor alrededor de una página por hora y precisa una revisión de unos veinte minutos.

contener la entrada léxica, información relativa a la categorización gramatical y alguna información –sintáctica el tipo de complemento usual en el caso de los verbos, o la preposición regida por el verbo o por un adjetivo–. A continuación, un diccionario bilingüe convierte las unidades de la lengua-fuente en sus equivalentes en la lengua-destino. Este segundo diccionario contiene una información similar al anterior, además de algunas indicaciones sobre el contexto en el que se usan las palabras de la lengua-fuente o sobre su dominio semántico –por ejemplo, uso legal, o uso técnico–. En esta fase las frases hechas y las locuciones, que no pueden dividirse en componentes, al igual que los homógrafos son registrados como tales.

Una vez establecidos los límites entre palabras, un analizador morfológico descompone el texto en entidades mínimas y lo prepara para la fase de análisis sintáctico, que requiere representaciones más complejas. Por supuesto, el análisis no es una fase exclusiva de la traducción automática y se advierte más estrechamente vinculada a la comprensión de textos (vid. 3.2.2.). Además, tampoco es imprescindible reconocer grupos de constituyentes de alto nivel para traducir una secuencia. Es posible, por ejemplo, reconocer una determinada clase de palabras a partir de una entrada léxica y luego predecir la siguiente con cierto grado de certeza. En términos generales, sin embargo, la traducción automática se realiza sobre análisis que identifican al menos los constituyentes de alto nivel. Siguiendo el modelo de la gramática de dependencias se suele aceptar que los verbos actúan como núcleos en torno a los cuales se articula un número variable de argumentos que saturan la predicación. Algunos sistemas, como METAL o el japonés ATLAS, extienden esta perspectiva considerando que el verbo es un marco amplio de significado concretado después por las distintas funciones semánticas. De este modo, se supone que las relaciones semánticas subyacentes constituyen un fundamento universal del análisis que luego se realiza con estructuras de superficie específicas de cada lengua –como activas o pasivas, una secuencia de preposición y sustantivo o un sustantivo en dativo, etc. (Hutchins 1986:197)–. En todo caso puede notarse la estrecha relación existente entre los sistemas de traducción automática y las representaciones sintácticas de la lingüística estructural y generativa.

Los primeros sistemas de traducción automática consistían únicamente en la conversión de cada unidad de la lengua-fuente en su equivalente en la lengua-destino. Los problemas derivados de semejante concepción no son difíciles de prever: ni se respetan las relaciones que las unidades mantienen con el todo superior en el que se integran, ni se supera una visión local del texto, que hace de la palabra la unidad de traducción por excelencia. Además de los lógicos problemas de solvencia, la traducción con este mecanismo sólo es posible si se restringe a un par de lenguas determinado como en la primera versión de SYSTRAN, dedicada



exclusivamente a la conversión del ruso al inglés. Estos problemas obligaron a optar por procedimientos indirectos, que trasvasasen el texto de una lengua a otra teniendo en cuenta representaciones intermedias. Como primera alternativa surge la *traducción por pivote o interlengua*, articulada sobre dos módulos: un módulo de análisis, que produce una representación del texto de entrada en una lengua-pivote presuntamente independiente de cualquier lengua particular, y un módulo de generación, que construye a partir de esa representación un texto de salida.

texto en lengua-fuente > MÓDULO DE ANÁLISIS > representación  
del texto en lengua-pivote > MÓDULO DE GENERACIÓN > texto en lengua-destino.

El cometido de la lengua-pivote es, por tanto, servir de puente entre el análisis y la generación, garantizando una representación suficientemente abstracta como para preservar el contenido del texto con independencia de las restricciones propias de cada una de las dos lenguas. Al situarse en un nivel puramente conceptual, la lengua-pivote admite información extralingüística relativa al dominio particular del texto. Este mecanismo fue adoptado en los años setenta por algunos de los grandes sistemas de segunda generación como GETA. La solución parecía particularmente interesante para traducciones entre lenguas tipológicamente alejadas pero planteaba problemas teóricos –llevaba al extremo la hipótesis de los universales semánticos y exigía análisis muy profundos, a veces imposibles de llevar a efecto<sup>24</sup>–. La segunda alternativa consta de tres módulos. Esto supone que al análisis y la generación se añade un nivel intermedio o de *transferencia*, que transforma la salida del módulo de análisis en una entrada para el de generación:

texto en lengua-fuente > MÓDULO DE ANÁLISIS > representación  
R del texto > MÓDULO DE TRANSFERENCIA > representación  
R' del texto > MÓDULO DE GENERACIÓN > texto en lengua-destino.

Los *sistemas de transferencia* son bastante más realistas porque las representaciones R y R' son distintas y, por tanto, están situadas en un nivel más concreto que la lengua-pivote. Los grandes sistemas de traducción automática han optado, desde los años ochenta por este mecanismo. Así, podrían citarse los proyectos GETA (desde 1971), METAL (desde 1978), MU (desde 1980) y

---

<sup>24</sup>La elaboración de una lengua-pivote conserva, sin embargo, adeptos en el campo de la inteligencia artificial, donde se aspira a elaborar maquetas de traducción en universos cerrados. Como veremos más adelante (vid. 3.2.2.) esta perspectiva pretende que la representación del texto consiga el nivel de profundidad necesario para su comprensión. Sin embargo, la construcción de verdaderos sistemas a tamaño real supondría un trabajo considerable y exigiría la sistematización rigurosa de una gran cantidad de conocimiento del mundo.

EUROTRA (desde 1982)<sup>25</sup>. A partir de los años noventa se han ido abriendo nuevas vías de investigación en este campo. El desarrollo camina hacia cálculos probabilísticos o sistemas expertos sustentados por un diálogo entre hombre y máquina donde el sistema aparece dotado de ciertas capacidades de aprendizaje.

### 3.2.2. *Comprensión automática de textos*

#### 3.2.2.1. El ámbito de la comprensión automática

La mayor parte de los productos resultantes de las industrias de la lengua exigen que, en alguna fase de la tarea, el ordenador pueda extraer y manejar el significado contenido en un texto (Verdejo 1987). La traducción automática, las interfaces en lengua natural o los procesadores de texto de cierta entidad no lograrían sus objetivos sin una comprensión automática del texto. Esta denominación genérica del campo puede parecer un tanto ampulosa, sobre todo si se tiene en cuenta el tipo de trabajo real que los sistemas llevan a cabo. En este sentido, convendría hacer dos precisiones.

En primer lugar, parece más adecuado utilizar la etiqueta *comprensión del texto* que otra, probablemente más extendida, como *comprensión del lenguaje natural*. Es bastante habitual en el terreno de la lingüística computacional utilizar la denominación de *lenguaje natural* para aglutinar las lenguas humanas por oposición a los lenguajes informáticos. Aunque esta oposición resulte muy cómoda y goce de gran éxito, desde la perspectiva lingüística resulta poco afortunada. Los lenguajes informáticos son códigos sustitutivos que reemplazan a las lenguas humanas en la comunicación entre hombre y máquina. Así que el riesgo de confusión no existe. La etiqueta *lenguaje natural* subraya una precisión innecesaria –lenguaje es, por excelencia, el lenguaje natural; lo extraño, lo marcado es que exista un lenguaje técnico de programación–. En este sentido, parece preferible hablar del procesamiento de las lenguas y no del procesamiento del lenguaje natural<sup>26</sup>. Pero la precisión terminológica no concluye ahí, porque no podemos pretender una

---

<sup>25</sup> Además de ser el proyecto más ambicioso, EUROTRA es particularmente interesante desde el punto de vista lingüístico porque produce representaciones que son accesibles a los lingüistas y que pueden modificarse a la luz de la experiencia o de las propiedades de una lengua concreta. Diversos artículos describiendo el proyecto se encuentran en el vol. 6/2 de la revista *Machine Translation*, publicado en 1991.

<sup>26</sup> Salvando las distancias, las diferencias entre *procesamiento del lenguaje* y *procesamiento de las lenguas* serían similares a las establecidas por Coseriu (1974) en el tema de los universales para distinguir entre *universales del lenguaje* y *universales de las lenguas*.

comprensión automática de las lenguas, a lo sumo observaremos sus efectos sobre textos, esto es, sobre porciones muy limitadas del discurso emitido en las lenguas.

En segundo lugar, la denominación todavía puede propiciar fuertes equívocos por el carácter inherentemente humano que se asocia al concepto de comprensión. Como suele ocurrir en estos temas, la etiqueta favorece ciertas suspicacias. En la actualidad resulta pueril el debate sobre si términos como *comprensión* o *inteligencia* deben aplicarse a los sistemas informáticos. La investigación en ciencias cognitivas es una vía prometedora que intenta emular muchos de los procedimientos del cerebro humano pero, en último caso, la comprensión de un sistema informático no tiene que ser similar a la del hombre: basta con que el ordenador procese información lingüística y extraiga de ella las inferencias oportunas, del mismo modo en que actúa en los procesos de cálculo. Con todo, el problema de la comprensión de textos es uno de los más delicados y complejos a los que se han dedicado los ordenadores. Y su evidente utilidad no se corresponde con un alto nivel de desarrollo. Como ocurría en el caso de la traducción, todavía sólo es posible construir analizadores más o menos eficientes para dominios muy restringidos.

#### 3.2.2.2. Análisis sintáctico

Los analizadores automáticos de datos lingüísticos –parsers– son sistemas informáticos que incorporan, de algún modo, el conocimiento que los seres humanos tienen de la lengua o lenguas que manejan<sup>27</sup>. Para trabajar con un ordenador, ese conocimiento debe formalizarse del modo adecuado. Aunque el estudio de los sistemas basados en conocimiento pertenece a la rama de la computación conocida como *inteligencia artificial*, el futuro de la investigación en el campo depende, en buena medida, del desarrollo de formalismos gramaticales compatibles con los sistemas de procesamiento. La comprensión del texto tiene, como primera fase, el análisis sintáctico, esto es, la detección de las unidades que lo componen y de las relaciones que esas unidades entablan. Ese análisis puede enfocarse, bien en sentido descendente –top-down parsing–, bien en sentido ascendente –bottom-up parsing–. El análisis descendente comienza por las unidades sintácticas de nivel superior –oraciones o cláusulas– para distinguir en su interior unidades menores que realizan funciones del tipo de sujeto, complemento directo o indirecto. A su vez, en el interior de estas nuevas unidades se distinguirán las relaciones pertinentes hasta llegar a la palabra aislada. Las representaciones arbóreas de la estructura sintáctica –propiciadas por la gramática generativa pero asumidas como herramienta de trabajo

---

<sup>27</sup> Vid., por ejemplo, King (1983) o Dowty et alii (1985).

por la práctica totalidad de los modelos vigentes en teoría gramatical– utilizan esta estrategia descendente. Se trata, en definitiva, de calibrar diferencias de comportamiento entre los constituyentes o unidades que configuran el discurso en cada nivel para identificar la poderosa red de relaciones que media en todo mensaje lingüístico. En el terreno computacional, sin embargo, se ha desarrollado también un análisis ascendente, que inicia la ruta en la palabra para, desde ahí, inferir el grupo de palabras que, juntas, desempeñan una misma función hasta llegar a la cláusula o al texto en su conjunto. Por supuesto, cada una de estas vías tiene ventajas y desventajas. El análisis descendente es más simple tanto a nivel conceptual como por la facilidad de escribir el programa para ordenador y, además, favorece la eliminación de ambigüedades. En una cláusula como (3), la unidad «canto» nunca podría ser considerada un verbo:

(3) *Tu canto nos ensordecía*

El problema más importante que esta aproximación debe afrontar es la necesidad de ir continuamente atrás hasta obtener el análisis idóneo. La estrategia ralentiza el proceso y malgasta los recursos disponibles. Para reconocer un verbo como tal habría que tener en cuenta la posibilidad de que se realizase como forma simple (4a), compuesta (4b) o perifrástica (4c), además de contemplar los posibles complejos verbales o las formas pronominales (4d):

(4a) *Juan dió un libro a su hermano*  
VERB

(4b) *Juan ha dado un libro a su hermano*  
VERB=AUX + VERB

(4c) *Juan anda diciendo que no vendrás*  
VERB=AUX + VERB

(4d) *Juan no se da cuenta de nada*  
VERB= se + VERB + SUST

El análisis ascendente, por su parte, tropieza con dificultades similares. Su objetivo consiste en señalar las fronteras entre palabras y asignar cada una de ellas a una clase determinada. Pero en una cláusula tan sencilla como (5) casi todas las palabras pueden incluirse en varias clases:

(5) *El niño pinta los dados de naranja*  
*el*: ART / PRON  
*pinta*: VERB / SUST  
*los*: ART / PRON  
*dados*: SUST / VERB  
*naranja*: SUST / ADJ

Es cierto que con esta técnica los constituyentes se construyen de una vez para siempre, sin posibles vueltas atrás pero, a cambio, muchas de las agrupaciones son totalmente incorrectas. Ante la dificultad de decidir entre ambas posibilidades, se han ido desarrollando técnicas mixtas que pretenden evitar los problemas de cada una de ellas. Lo más habitual es utilizar condicionales de tipo predictivo del estilo de «Si el verbo SER aparece seguido de participio, ambas unidades se analizan como un solo verbo de significado pasivo; en otro caso, SER se trata como verbo copulativo». Este modelo es bastante eficaz pero indudablemente lento<sup>28</sup>. Durante algún tiempo se confió en que los análisis transformacionales sirviesen como alternativa. En la versión estándar del modelo generativo (Chomsky 1965) el significado de cada cláusula se preserva en un árbol de estructura profunda, basado en la sintaxis, una estrategia que pretende llegar a la representación total del texto. Sin embargo, los analizadores de base transformacional se encontraron con serios problemas. El modelo generativo derivaba estructuras de superficie a través de una representación de la estructura profunda y no al revés, de modo que era difícil invertir el proceso para reconstruir la información de la estructura profunda de una cláusula sobre la cadena superficial. Como indica Lewis (1992:83) el modelo era una hipótesis de caracterización y no un método para procesar en el ordenador. Con el tiempo, la teoría de rección y ligamiento propondrá un sistema más general de descripción, que limita el alcance de las transformaciones –en la primera etapa excesivamente poderosas y, a menudo, demasiado complejas para los fenómenos lingüísticos que describían– especificando condiciones locales de aplicación. Sin embargo, esta versión no se aplicó al tratamiento informático, a pesar de ser bastante más simple. Al final, muchos sistemas han tenido que elaborar modelos gramaticales propios, como la gramática de estructura de frase generalizada de Gazdar (1987, 1989a, 1989b).

### 3.2.2.3. Análisis semántico

Una vez conseguido el análisis sintáctico, el tratamiento automático de las lenguas exige un análisis semántico, que asocie cada secuencia a una representación interna de su contenido. El procedimiento más habitual en la construcción de sistemas de este tipo pasa por separar la información gramatical de la meramente semántica en bases de datos independientes y, a la vez, independientes del algoritmo de análisis, para que sea más fácil hacer correcciones o añadir información nueva.

---

<sup>28</sup> Existen, por supuesto, otras posibilidades, como las *redes de transición* en las que sin embargo no parece oportuno entrar en una visión meramente panorámica.

Sin embargo, las distintas fuentes de información deben interactuar cuando el programa está corriendo.

Como el Estructuralismo nos ha enseñado, comprender un texto es algo más que comprender el significado de sus palabras aisladas –el dominio clásico de la semántica léxica–, es también comprender el significado de las relaciones que esas unidades entablan –semántica gramatical–. Y ni siquiera esta conjunción de unidades y relaciones compendia la totalidad de la tarea; únicamente explica el significado de cláusulas aisladas, algo bastante diferente del significado completo de un texto tal como el ser humano lo aprehende luego de un proceso de comprensión. La semántica cognitiva ha avanzado poderosamente en los últimos años, pero ni siquiera es necesario un marco teórico determinado para percibir, de modo intuitivo, que la comprensión del texto supera la simple decodificación del contenido literal del mismo para abarcar también la reconstrucción del mensaje en el momento en que fue emitido, las connotaciones que arrastra y lo no-dicho pero implícito en el enunciado. Comprender un texto, al fin, es interpretarlo a la luz de la situación y las condiciones que lo produjeron.

En este sentido, los sistemas de comprensión automática de textos escritos deben afrontar dos tipos de problemas. Por un lado, deben apoyarse en una sólida gramática del texto, un terreno poco explorado, que atiende a las relaciones supra-oracionales puesto que cada secuencia de un texto se apoya sobre las precedentes y/o anuncia las que van a sucederle. Por otro lado, debe atender a la identificación de los interlocutores, del momento en que se emite el discurso y de las condiciones de la enunciación. La gramática tradicional y la retórica utilizaban las denominaciones de anáfora y catáfora para dar cuenta de algunos de los fenómenos concernidos por esta visión amplia de la comprensión. En su forma más simple, se trata de elementos gramaticales que envían a un antecedente previamente expresado –catáfora– o que tardará algún tiempo en aparecer –anáfora–. Los pronombres personales, los demostrativos, los posesivos o los artículos determinados tienen casi siempre estos valores y el problema se limita a identificar ese antecedente/consecuente, especialmente cuando el contexto permite una cierta pluralidad de soluciones (6a). Más difíciles de tratar son los casos de anáfora asociativa (6b) donde el antecedente no está explícitamente presente en el contexto anterior, sino que debe reconstruirse a partir de lo que se dice, o los casos de elipsis catafórica habitual en las coordinaciones (6c). Por fin, la correferencia también puede realizarse mediante procedimientos léxicos, como ocurre con el uso de sinónimos o perífrasis (6d):

- (6a) *Al cruzar la calle Juan se encontró con su hermano. ¿Adónde vas? -le dijo*  
(¿Quién lo dijo, Juan o su hermano?)

- (6b) *Llegó al pueblo. La plaza tenía un aspecto triste*  
(La plaza del pueblo)
- (6c) *Juan compró y Pedro usó el coche*  
(Lo que Juan compró fue el coche)
- (6d) *Cervantes es universalmente conocido como autor de El Quijote. Pero el manco de Lepanto escribió otras obras*  
(El manco de Lepanto = Cervantes)

El tratamiento computacional de base consiste en establecer una lista de antecedentes potenciales, imponer restricciones morfológicas, semánticas o pragmáticas a la anáfora y verificar si esas restricciones se satisfacen en el caso particular. Estos problemas apuntan la necesidad de ciertas herramientas de representación que permitan efectuar los cálculos necesarios para la construcción progresiva del sentido de un texto. Por supuesto estos útiles pueden ser más o menos finos en función de la tarea que se quiera realizar pero inevitablemente surge la relación entre la información extraída del texto y los conocimientos extralingüísticos.

#### 3.2.2.4. Conocimiento del mundo

Cuando un ser humano interpreta un texto no depende exclusivamente de sus conocimientos lingüísticos en tanto que hablante de determinada lengua. Una fuente de información crucial para la comprensión procede de la cultura o las propiedades de los objetos del mundo real que se conocen por la experiencia, las creencias, presuposiciones o recuerdos personales. Esos conocimientos extralingüísticos filtran la información y seleccionan la posibilidad adecuada de entre las diversas significaciones potenciales. De ahí que, para interpretar correctamente los textos, los sistemas de comprensión deban disponer de ese conocimiento del mundo que sobrepasa el contenido lingüístico propio de la semántica. Para comprender una declaración judicial, por ejemplo, hay que saber algo del sistema de valores de una sociedad, del concepto que ésta tenga de la propiedad o de sus creencias morales. Y para entender la descripción de un accidente automovilístico hay que tener una idea, siquiera mínima, del código de circulación. En definitiva, hace falta disponer de una organización de los conocimientos que no esté ligada exclusivamente a las palabras sino también a los objetos, a los sucesos y a las leyes del dominio que se trata. Algunos proyectos de investigación han pretendido modelizar el conjunto de conocimientos de un ser humano, como el proyecto americano CYC, aunque sus posibilidades de éxito parecen bastante remotas. Por mucho que se persevere en la tarea de almacenar el conocimiento enciclopédico, siempre habrá que hacer frente al hecho de que las relaciones entre los signos lingüísticos y la realidad varían continuamente en las comunidades, dependen del contexto y se establecen de modo

subjetivo; lo que explica el carácter creativo del lenguaje y sus posibilidades metafóricas.

La representación del conocimiento enciclopédico o conjunto de saberes de que el hombre dispone por estar vivo y actuar en una sociedad y un contexto determinado no es un problema específico del tratamiento automático de las lenguas; al contrario, es la cuestión central sobre la que la inteligencia artificial se ha ido cimentando en los últimos veinte años. La elaboración de sistemas de segunda generación especialmente concebidos para la traducción automática impulsó también la investigación en comprensión automática de textos en la perspectiva de la inteligencia artificial. El interés de los investigadores en este dominio se centra en la articulación de semántica, pragmática y conocimiento del mundo (cf. Winograd 1983). En esta perspectiva se construyeron sistemas de comprensión limitada a universos restringidos y susceptibles de integrarse en procesos automáticos. La mayor parte de los sistemas utilizan, bajo variantes diversas, una forma de representación introducida por Minsky (1975), los esquemas –orig. *frames*–. Un esquema es una red semántica que asocia información a un atributo dado. La coherencia del sistema se debe a controles locales que permiten verificar si es verosímil asignar un determinado valor a un concepto. En esta misma línea Schank (1975; Schank y Abelson 1977) elaboró la teoría de dependencias conceptuales que proponía un nivel de representación semántica independiente de las lenguas particulares para que hiciese el papel de lengua intermedia en un sistema de traducción denominado MARGIE. Pero el mundo de la lingüística computacional y el de la inteligencia artificial permanecieron distantes durante bastante tiempo pues sus marcos teóricos eran bien diferentes, y sus objetivos no llegaban a coincidir<sup>29</sup>. No obstante, en los últimos años han ido experimentando un tímido acercamiento. La idea de fondo es integrar en los programas de comprensión automática sistemas expertos que permitan gestionar grandes bases de datos de conocimiento del mundo y efectuar las inferencias necesarias para establecer el significado de un texto. Este conocimiento del mundo no se utiliza tanto para simular una comprensión total del texto, cuanto para lograr determinados efectos prácticos. En ejemplo de Fuchs (1993:198), para comprender que en «La flor está compuesta de sépalos del cáliz y de la corola», la frase preposicional «de la corola» es coordinada a «sépalos» y no a «cáliz» hay que tener conocimientos de botánica pero esa perfecta comprensión no es necesaria para traducir la secuencia al francés, que conoce también esa ambigüedad estructural.

<sup>29</sup> En términos generales, se diría que el objetivo de la lingüística computacional radica en la obtención de sistemas operativos a escala real. Por su parte, la inteligencia artificial, al ocuparse de estos asuntos, únicamente persigue la elaboración de maquetas experimentales en las que procesos como comprensión, traducción o generación intervienen únicamente como tests.



### 3.2.2.5. Perspectivas de futuro

Durante mucho tiempo los sistemas de comprensión automática no fueron más que maquetas de laboratorio, muy útiles para el desarrollo de la investigación, pero incapaces de producir aplicaciones verdaderamente representativas. En realidad, este paso es difícil de conseguir porque el control del sistema se hace muy delicado cuando no es posible asegurar manualmente la coherencia de las bases de conocimiento. El futuro exige una planificación rigurosa de los objetivos y la colaboración activa entre los sectores industriales implicados en la explotación comercial de las aplicaciones y los investigadores encargados del desarrollo teórico del campo. Un ejemplo idóneo de esa colaboración es el sistema PALME sobre las ofertas de empleo que aparecen en el periódico francés *Le Monde* realizado por GSI-ERLI. Con quince años de uso a sus espaldas, PALME se apoya sobre un servicio telemático en videotexto que permite al usuario acceder a las ofertas publicadas por el periódico, bien poniendo una pregunta de formulación libre, bien describiendo sus méritos profesionales. El sistema selecciona las ofertas de empleo adaptadas al perfil del candidato a partir de un índice de las palabras clave aparecidas en los anuncios y su proximidad semántica con la información que el usuario ha dado.

### 3.2.3. Generación automática de textos

#### 3.2.3.1. Procesos implicados en la generación

En el marco de la comunicación entre hombre y máquina, la generación automática de textos se ocupa de producir las respuestas de la máquina en la lengua deseada por el usuario. Como ámbito de estudio, se trata de una de las ramas más jóvenes de la lingüística computacional y también una de las más heterogéneas. La naturaleza de su cometido hace que comparta intereses con la teoría lingüística, la inteligencia artificial y la psicolingüística. Posee, por tanto, una dimensión teórica ligada a los ingredientes psicológicos y formales implicados en la producción de los mensajes lingüísticos, además de su clara proyección práctica: facilitar al usuario de sistemas informáticos algún tipo de aplicación en la forma más simple posible, es decir, en la lengua que éste use. El lugar del lingüista en este dominio ha sido muy modesto, ya que el grueso de la investigación descansa en manos de la psicología cognitiva y la inteligencia artificial. Pero la naturaleza de los problemas implicados en este terreno obliga, cada vez más, a reivindicar el papel del lingüista como un agente indiscutible del proceso.

La tarea de generar textos exige atender tres cuestiones diferenciadas: (i) la realización gramatical del discurso –cómo se dice algo–; (ii) la selección de los

contenidos pertinentes al texto –qué se dice– y (iii) la motivación para producir el texto –por qué se dice–. El campo comenzó su andadura histórica intentando satisfacer las dos primeras cuestiones; la primera se apoyaba en una amplia y antigua tradición gramatical y la segunda había recibido un fuerte empuje por parte de la filosofía del lenguaje (Austin 1962, Searle 1969 y Grice 1971). Sin embargo, no ha sido fácil aunar ambos tipos de orientación. Los programas que construyen emisiones correctas no son capaces de seleccionar el material informativo y organizarlo en texto coherente y, a su vez, los sistemas que logran esa selección del contenido tienen malos resultados en gramática. Los primeros proyectos en generación automática nacieron bajo la influencia de la gramática generativa<sup>30</sup> y, por tanto, se ocupaban exclusivamente de los aspectos sintácticos a despecho de los semánticos y pragmáticos. Su objetivo era validar la coherencia de la teoría antes que expresar contenidos semánticos en contextos determinados. Ni siquiera en los años siguientes, cuando la generación empezó a atender a la producción de frases aisladas (Goldman 1975, McDonald 1981) fue posible obtener discursos coherentes. Sólo los sistemas más recientes, como PAULINE (Hovy 1988), han tomado en consideración que el texto busca incidir directamente en el interlocutor al que va dirigido y que, en consecuencia, ese objetivo inmediato puede determinar no sólo su forma, sino también su contenido<sup>31</sup>.

### 3.2.3.2. Métodos de generación

A pesar de que comprensión y generación de textos pueden verse como dos procesos estrechamente conectados, su desarrollo computacional no ha ido en paralelo. Mientras que los primeros analizadores hacen su aparición a partir de los años cincuenta, los programas de generación no surgen hasta finales de los setenta. Ese retraso –hoy a punto de superarse– se explica por varias razones. La vocación descriptiva que tradicionalmente han tenido los estudios lingüísticos justifica que la comprensión y el análisis de los textos se vieran privilegiados sobre la generación. Hubo que esperar al progreso teórico de la lingüística del texto y la pragmática para contar con una base lingüística que respaldase este proceso. Pero más que esta razón, de orden teórico, debió de influir la propia naturaleza de la tarea. Desde el punto de vista computacional la generación de textos, al menos si nos contentamos con

<sup>30</sup> El término *gramática generativa* puede propiciar, en este contexto, algún equívoco. El adjetivo *generativa*, aplicado a la corriente lingüística abanderada por N. Chomsky, tiene un sentido puramente matemático y el modelo es neutro en cuanto a la problemática de comprensión y producción de textos.

<sup>31</sup> Para más información sobre el desarrollo pragmático de la generación automática de textos, vid. Kempen (1987) y Bateman y Hovy (1992).

una calidad mínima, es mucho más fácil que su comprensión. De hecho, el método más sencillo de generar un texto por ordenador consiste en reproducir un mensaje previamente registrado, del estilo de las indicaciones frecuentes en los procesadores de texto, como «asegúrese de que la impresora está conectada» o en los cajeros automáticos de los bancos, que emiten mensajes del tipo de «bienvenido, por favor inserte su tarjeta». La máquina sólo tiene que buscar, de entre el conjunto de textos registrados, el que se ajuste a la situación concreta. Pero la evidente simplicidad de la operación se consigue a expensas del poder expresivo. Este método sólo es satisfactorio cuando el número de respuestas posibles es muy restringido.

El primer intento de ampliar el modelo consiste en introducir variables vacías que puedan rellenarse con datos diferentes. Así se solventan problemas de memoria aunque la limitación combinatoria sigue manteniéndose. Cuando el sistema cuenta con pocas variables, éstas pueden recuperarse mediante una simple lista; de otro modo, deben asociarse a unidades léxicas particulares. El primer sistema que exploró esta posibilidad, BABEL (Goldman 1975), seleccionaba las entradas léxicas por medio de árboles de discriminación que se habían asociado a unidades primitivas de representación. Así, INGEST –esp. «ingerir»– era la variable utilizada para incluir procesos de ingestión de alimento (comer), líquido (beber), aire (respirar) o humo (fumar) entre otras posibilidades. Este tipo de dependencias conceptuales (Schank 1975), que fueron desarrolladas por otros generadores, buscan la palabra apropiada antes de producir ninguna emisión. La aproximación contraria se encuentra en los generadores que se organizan sobre módulos interdependientes, cada uno de los cuales cumple una función, de modo que primero se construye la estructura de la cláusula y luego se cubre con las palabras oportunas. Para ello se introducen en la máquina las estructuras más elementales de una gramática, haciendo de la generación un proceso de expansiones repetidas. Así, por ejemplo, en español una frase nominal se definiría con una regla de este tipo:

FN → DET - NUC - MOD

Para generar un texto, el sistema empieza con el símbolo FN que, por medio de la regla dada, se expande en sus constituyentes. De utilizar unidades más complejas que la frase, el número de expansiones tendría que repetirse en cascada hasta obtener el efecto deseado. Este sistema tiene su primera limitación en el orden fijo de constituyentes, que no respeta la realidad de la lengua. Para continuar con el ejemplo anterior, no todas las frases nominales comienzan con determinantes ni, mucho menos, todas tienen modificadores. Además, los sistemas de construcción en cascada deben afrontar el hecho de que, cuanto más rica sea la gramática usada por el generador, más cuidadoso tendrá que ser el usuario a la hora de especificar

la información pertinente de entrada para que el programa seleccione la alternativa correcta en la cascada. Esto exige que el usuario tenga conocimientos lingüísticos.

En los modelos siguientes las unidades léxicas se reemplazaron por haces de rasgos identificadores que permitían, por ejemplo, registrar las variantes morfológicas. El generador opera con una colección de rasgos que luego traslada al texto preciso. Desde luego, los sistemas basados en rasgos constituyen el método de generación más sofisticado pues consiguen representar y manipular explícitamente cada variante gramatical. La generación se convierte en el ensamblaje de un conjunto de valores, expresados en rasgos mínimos, que contienen el significado deseado. El problema básico consiste en decidir qué rasgos son los pertinentes y cómo representarlos.

### 3.2.3.3. Los problemas de la generación

Habitualmente, el proceso de generación se descompone en dos módulos; el primero se ocupa del terreno conceptual –por ejemplo, del orden y secuencia de las informaciones– y el segundo de las cuestiones puramente lingüísticas, como la selección léxica o sintáctica. La tarea está jerarquizada, ya que se considera que las operaciones conceptuales se sitúan en el nivel superior, mientras que las cuestiones sintácticas serían de nivel y urgencia menor; y las operaciones morfológicas se insertarían en el nivel inferior del análisis. Ahora bien, esta aproximación estratificacional, que también apareció en los primeros tiempos en la comprensión de textos, se ha mostrado muy limitada pues no puede dar cuenta de la interdependencia que media entre estratos distintos.

La aproximación estratificacional reclama un orden fijo de operaciones. En primer lugar se selecciona una estructura de discurso, a continuación un predicado, luego una construcción sintáctica y, por fin, se aplican las reglas morfológicas oportunas. Así, una situación del tipo TRANSACCIÓN permite elegir en español el predicado «vender» que selecciona la estructura sintáctica SUJ-PRED-CDIR-CIND, cubierta luego con algo del tipo «Juan le vendió un coche a María». Esta técnica resulta convincente desde el punto de vista computacional porque permite construir un algoritmo modular en el que intervengan secuencialmente semántica, sintaxis y morfología, pero tiene obstáculos insalvables desde el punto de vista lingüístico. La estructura del discurso anterior puede expresarse en español con cualquiera de los esquemas siguientes:

- (a) SUJcomprador-PREDact(comprar)-CDIRmercancía[-CINDvendedor]
- (b) SUJmercancía-PREDpas(comprar)-CAGcomprador[-CINDvendedor]
- (c) SUJvendedor-PREDact(vender)-CDIRmercancía[-CINDcomprador]
- (d) SUJmercancía-PREDpas(vender)-CAGvendedor[-CINDcomprador]

El generador debe elegir la mejor de las posibilidades, de acuerdo con el predicado y su contexto. Si el comprador no estuviese especificado en la información contenida, debería eliminar la opción (a) y si la mercancía apareciera en cantidad indefinida, elegiría una estructura sintáctica activa en vez de una pasiva. Así que el generador debe incluir una léxico-gramática que indique las propiedades sintácticas de cada entrada léxica y clarifique, por ejemplo, si hay posibilidad de verificar determinadas transformaciones como la pasiva. Para evitar esta secuencialidad de las aproximaciones estratificacionales se impone (Danlos, 1985, 1993) un modelo de dos módulos:

representación del texto-fuente > COMPONENTE ESTRATÉGICO >  
esquema del texto > COMPONENTE SINTÁCTICO > texto-salida

El componente estratégico seleccionaría las estructuras de discurso, los elementos predicativos que expresan los predicados y las construcciones sintácticas, respetando la interdependencia entre estas operaciones. Para ello se apoyaría sobre una gramática del discurso, que incluiría informaciones conceptuales y lingüísticas, y una base léxica nutrida de información gramatical. Por supuesto, tanto una como otra dependen del tipo de texto que se quiera generar y, por tanto, se elaboran sobre un corpus de textos del dominio correspondiente. La idea de fondo es que todo texto posee una estructura determinada por el tipo de información que transmite, algo que sólo una gramática del discurso puede captar. Los sistemas de generación producen productos fiables, con un porcentaje de errores nulo. Esta situación contrasta con la de los sistemas de comprensión automática que no pueden conseguir estos resultados porque dependen del usuario y no de la máquina. Tanto la lingüística teórica como la lingüística computacional pueden extraer importantes progresos de un acercamiento a la generación automática. Como vemos, esta disciplina obliga a realizar finos análisis gramaticales y exige una fuerte reflexión sobre las cuestiones fundamentales de la teoría de la comunicación.

### *3.2.4. Dicionarios automáticos*

#### *3.2.4.1. La evolución del diccionario*

En el sentido usual del término, un diccionario es un inventario de las formas léxicas de una lengua determinada asociadas a una definición de su significado o de las condiciones en que se usan. Pero, más allá de esta visión superficial, el diccionario es una herramienta lingüística y, como tal, ha ido alcanzando un creciente grado de sofisticación. En la actualidad un buen diccionario contiene, al menos, las siguientes informaciones:

a) entrada léxica o forma externa de la palabra con variantes ortográficas si fuera preciso;

b) información gramatical, del tipo *clase de palabras*, *uso transitivo* o *intransitivo* en el caso de los verbos, *contable* o *no contable* en el caso de los nombres, *uso atributivo* en el caso de los adjetivos, etc;

c) variantes flexivas, a través del registro de formas compuestas, derivadas o irregulares;

d) pronunciación en una notación estándar, normalmente el alfabeto fonético internacional;

e) información semántica, esto es, indicaciones sobre el significado de la palabra correspondiente, enriquecidas con sinónimos y notas de uso;

f) ejemplos que ilustran el uso de la palabra en varias acepciones, expresiones fijas en que aparece, usos especiales, registro y variación regional.

La profundidad de esta información varía, por supuesto, en función del destinatario al que el diccionario va dirigido. Manejamos diccionarios monolingües y bilingües, etimológicos o puramente sincrónicos, de uso general o restringidos a la terminología particular de una disciplina, y de cada uno de ellos reclamamos cuestiones específicas.

En todo caso, la expansión que las nuevas tecnologías han experimentado en los últimos años ha alcanzado también a la lexicografía. La intervención de procedimientos informáticos ha transformado sensiblemente la tarea de elaborar un diccionario y su propia organización interna. A partir de los años setenta comienza a plantearse la posibilidad de que el ordenador maneje las informaciones contenidas en un diccionario. Los *diccionarios legibles por ordenador*, el primer paso dentro de la lexicografía computacional, debían tomar la forma adecuada para cumplir ese objetivo. El paso siguiente consiste en utilizar la propia máquina para elaborar *diccionarios asistidos por ordenador* mediante programas que no sólo procesen sus datos, sino que también aseguren que las definiciones de cada entrada léxica incluyen únicamente palabras contenidas por el propio diccionario. Por fin, los *diccionarios diseñados por ordenador* explotarían la intervención informática en todas las fases de realización: en la compilación de datos, la selección de entradas, el diseño de cada entrada y en su disposición interna. Meijs (1992) ilustra esta evolución con las primeras versiones automáticas de tres diccionarios de calidad producidos en Gran Bretaña. En los años setenta el OALD –Oxford Advanced Learner’s Dictionary– (Hornby 1974) se convertía en uno de los primeros diccionarios que tomaban la forma adecuada para ser manejados por un ordenador. Unos cuantos años

más tarde, LDOCE –Longman Dictionary of Contemporary English– (Procter 1978) aparecía ya como un diccionario asistido por ordenador y, finalmente, COBUILD –Collins COBUILD English Language Dictionary– (Sinclair et alii 1987) puede verse como un diccionario totalmente diseñado por ordenador. En este caso, el proceso de elaboración comienza con la compilación de un gran corpus de datos lingüísticos y su almacenamiento en una base de datos. A continuación, el trabajo puramente lexicográfico –asignación de la clase de palabras, pronunciación, comportamiento flexivo, etc.– recibe tratamiento computacional por medio de un programa que proporciona tipos y formatos particulares para cada tipo de información. Al facilitar así el acceso al corpus, los encargados de escribir cada entrada disponen de concordancias que ayudan a discriminar sentidos diferentes o proporcionan ejemplos asegurando la consistencia y el acabado final del producto.

#### 3.2.4.2. Diccionarios y Lingüística de corpus

Aunque los lexicógrafos siempre han coleccionado datos, el ordenador ha dado nuevas dimensiones a este trabajo. Las capacidades analíticas y estadísticas de los sistemas informáticos permiten manejar la frecuencia relativa de las unidades léxicas como criterio para decidir o no su inclusión en un diccionario.

A pesar de esta clara pertenencia al dominio de la lingüística de corpus, la lexicografía computacional ha venido a conciliar dos tradiciones siempre vigentes en lingüística: el acopio de datos y la reflexión teórica. Aparentemente se diría que las herramientas informáticas, al ampliar los datos de referencia, dan un sesgo marcadamente empírico a la investigación. Pero cuando el tamaño de esos datos adquiere unas dimensiones respetables, se plantea inevitablemente el problema de equilibrar el crecimiento del corpus y su cobertura léxica. Como ya se ha indicado (vid. 3.1.3), el aumento de un corpus no se traduce en un aumento de unidades léxicas diferentes en su interior. El corpus LOB, por ejemplo, almacenó un millón de palabras procedentes de quinientos textos escritos pero sólo cincuenta mil de ellas eran palabras distintas. En general, en cualquier corpus de textos, aproximadamente la mitad de las palabras sólo aparecen una vez y la mayoría de las restantes aparecen menos de diez veces. Desde el punto de vista lexicográfico ni las palabras muy usuales ni las muy raras son problemáticas, ya que aparecen en número suficiente o, simplemente, no aparecen nunca. El verdadero problema radica en el estrato intermedio; en aquellas palabras que no son suficientemente usuales como para aparecer con seguridad en un corpus aleatorio, ni suficientemente raras como para considerar adecuado el corpus que no las contenga. La mayoría de los corpus elaborados recientemente aseguran ese equilibrio restringiendo sus miras a una

variedad sincrónica, sintópica y sinestrática y nutriéndose de textos de distintos géneros, que incluyan formas periodísticas y, si es posible, registros orales. A modo de ilustración, el COBUILD resulta de un corpus de veinte millones de palabras procedentes de textos variados incluyendo más de un millón de palabras del inglés hablado, procesadas mediante un escáner óptico desde 1980. La muestra pretende ser representativa de la variante de inglés estándar, desprovista de usos técnicos y, por tanto, correspondiente al uso natural de un hablante culto del inglés.

#### 3.2.4.3. Organización de los datos

Una vez almacenados los datos, conviene relacionarlos de modo económico y eficiente. El método más común consiste en un programa de concordancias, que busca en el corpus las apariciones de determinada palabra y produce ficheros o muestras de pantalla con esas apariciones y un ligero contexto. Además, el examen de las apariciones de cada palabra sobre un contexto original ayuda al lexicógrafo a decidir el número de acepciones que conviene distinguir y le facilita una fuente de ejemplos reales para ilustrarlas. El ordenador simplifica un trabajo forzosamente pesado y permite al lexicógrafo concentrarse en su verdadera tarea, asegurándole además que sus decisiones son coherentes, al menos dentro de los límites predefinidos.

En la actualidad los esfuerzos se centran en coordinar y estandarizar la investigación de este rico recurso de conocimiento léxico. Un objetivo básico es obtener un modelo de entrada léxica computacional que pueda generalizarse para facilitar un acceso sistemático a los distintos diccionarios y posibilitar la cooperación entre ellos. En este sentido, la CE financia el proyecto ACQUILEX (Acquisition of lexical knowledge for natural language processing systems)<sup>32</sup>. Este proyecto agrupa a investigadores de Amsterdam, Barcelona, Cambridge, Dublín y Pisa, que procesan una serie de diccionarios y los transforman en un formato estándar basado en una particular concepción de cómo debe organizarse una entrada léxica. El objetivo de esta tarea es comparar la organización conceptual de cada red semántica en los distintos diccionarios y enriquecerlos con las informaciones que aparezcan en los demás.

La construcción de este tipo de diccionarios necesita información de subcategorización sensible al contexto. El sistema ASCOT, por ejemplo, fue específicamente diseñado para obtener este tipo de información léxica «enriquecida»

---

<sup>32</sup> Vid. Boguraev y Briscoe (1989) y Martí et alii (1994).



a partir de un proyecto de la Universidad de Nimega, llamado TOSCA (Aarts y van den Heuvel 1985). Se trataba de producir ficheros que contuvieran sólo datos sintácticos correctos y refinados para volver sobre ellos en la forma de una base de datos jerarquizada, que asociaría automáticamente esta información a una lista de palabras. En el sistema TOSCA las palabras de un corpus reciben códigos relativos a cierta información gramatical –normalmente la clase de palabras– a través de un lematizador desarrollado en la Universidad de Lancaster. A continuación ASCOT proporciona información subcategorial adicional, que resulta esencial para extender los afijos. De este modo se busca, por ejemplo, si el sujeto concuerda en número con el verbo, teniendo en cuenta casos particulares del tipo de sustantivos que en su forma singular pueden tomar verbos en plural. También usa información sobre los complementos regidos por un verbo para decantarse por un determinado análisis en casos de ambigüedad estructural. El resultado de TOSCA es un corpus de cláusulas totalmente analizadas desde el punto de vista sintáctico que se almacenan en una base de datos lingüísticos para la investigación.

En los últimos años la mayor parte de los esfuerzos se han orientado hacia el tratamiento semántico de la información contenida en los diccionarios. La estructura argumental del predicado o conjunto de funciones semánticas que pueden saturar la predicación es un proyecto basado en diccionarios desde las primeras gramáticas de valencias alemanas (Helbig y Schenkel 1969). Pocos o ningún diccionario de uso general dan información formalizable en este sentido, aunque a veces las definiciones contienen indicaciones indirectas a este respecto<sup>33</sup>. Los proyectos de diccionarios de construcción y régimen constituyen una de las más interesantes elaboraciones del diccionario clásico como herramienta para la investigación lingüística<sup>34</sup>. Al construir índices y concordancias con los datos del corpus, se pueden hacer listas de palabras y su frecuencia permite orientarlas como herramienta para la enseñanza de la lengua en cuestión a estudiantes extranjeros. Además, estos diccionarios tienen otras aplicaciones prácticas, relativas a la construcción de correctores ortográficos y demás útiles del procesamiento de textos. Las listas de palabras provistas de su pronunciación son igualmente útiles para la investigación en reconocimiento y síntesis del habla.

---

<sup>33</sup> Por ejemplo, cuando el significado de un verbo establece que su objeto denota «persona», «cosa» o ambas.

<sup>34</sup> En el caso del español, un grupo de investigadores de la Universidad de Santiago de Compostela bajo la dirección de G. Rojo viene elaborando desde 1989 un diccionario de construcción y régimen del español actual. Para su descripción vid. Rojo (1992).

### 3.2.5. *Correctores ortográficos, gramaticales y de estilo*

De entre todas las aplicaciones producidas por la lingüística computacional, las más populares son las que prestan determinadas ayudas al proceso de escritura, ya sea en el ámbito profesional, ya en el uso cotidiano. Aunque las más interesantes desde el punto de vista lingüístico se sitúan en la fase posterior a la escritura, merecen alguna indicación los productos que intervienen en su redacción y composición.

En los últimos años, la fase de planificación en la escritura de un texto se ha visto alentada por programas dialogados que articulan los conocimientos del usuario sobre sus objetivos o su audiencia antes de preparar un texto. En general, se trata de herramientas didácticas que guían el desarrollo del tema o estructuran los párrafos. Además, todos los procesadores de textos son herramientas de edición que permiten almacenar, ver, mover, copiar o dar forma definitiva a un texto de modo más flexible, rápido y eficiente que ningún otro medio. Sólo cuando el texto está redactado pueden intervenir las herramientas de post-escritura, esto es, los programas que corrigen errores ortográficos, gramaticales o de estilo.

Los correctores ortográficos operan básicamente cotejando las palabras del texto con las de un diccionario previamente almacenado<sup>35</sup>. Obviamente, si la palabra en cuestión no aparece en el diccionario se la toma por incorrecta, así que su solvencia varía en función del tamaño del diccionario. Este procedimiento, propio de las primeras versiones de correctores, tiene la ventaja de que el diccionario puede fácilmente ser examinado y adaptado a las necesidades particulares del usuario –por ejemplo, permite añadir vocabulario técnico o especializado–. Pero desde el punto de vista computacional invierte enormes cantidades de tiempo y memoria y, desde el punto de vista lingüístico, la posibilidad de intervención directa del usuario acaba legitimando formas incorrectas. En los últimos años los correctores ortográficos han evolucionado, añadiendo a ese diccionario de formas léxicas una pequeña gramática en la forma de reglas de derivación morfológica y de formación de palabras. Aunque el procesamiento es más complicado, los resultados son bastante satisfactorios. Sin embargo, esta alternativa no está tampoco exenta de problemas. El corrector provisto de una gramática es bastante difícil de construir y las reglas morfológicas no garantizan palabras correctas. Si se instruye al sistema con la información de que [-ción] es en español un sufijo con el que se consiguen sustantivos abstractos a partir de verbos, se explica la formación de «deducción», sobre «deducir» o de «manifestación» sobre manifestar. Pero se permitirían también formas como «\*replicación» a partir de «replicar» y se justificaría «tumefacción»

<sup>35</sup> Vid., por ejemplo, para el español Casajuana y Rodríguez (1985).

sobre un supuesto verbo «\*tumefacer» y no sobre el «entumecer» tan extraño a la regla. Además, ningún corrector usa conocimiento pragmático y los sistemas comercializados ni siquiera emplean informaciones sintácticas o semánticas, de modo que un simple error mecanográfico que dé lugar a una secuencia absurda no puede ser detectado por el sistema.

En los años noventa han empezado a difundirse sistemas que enriquecen el corrector ortográfico con abundante información gramatical. La expansión en esta dirección es inevitable pues sólo un sistema provisto de un analizador sintáctico garantiza una mínima solvencia en este campo. Generalmente aparecen acompañados de abundante material didáctico y orientados a la enseñanza de segundas lenguas. Ahora bien, el corrector gramatical va ligado al desarrollo de sistemas de almacenamiento electrónico del texto a gran escala. Por fin, la posibilidad de introducir programas que incorporen manuales de estilo está siendo explorada también. Sus posibilidades parecen, en principio, menores porque una vez resuelta la corrección gramatical, las cuestiones de estilo resultan bastante subjetivas. En cualquier caso, el futuro anuncia una poderosa expansión de estos productos en el camino del hipertexto y de las herramientas de autoedición<sup>36</sup>.

#### **4. LA INTEGRACIÓN DEL TRATAMIENTO DEL HABLA Y EL TRATAMIENTO DEL TEXTO**

Tal como señalábamos en el apartado 1.1.5, existe una tendencia cada vez mayor hacia la integración entre las tecnologías del habla y el tratamiento del texto. Por una parte, la conversión de texto a habla y el reconocimiento de habla continua y de grandes vocabularios requieren el uso de herramientas propias del procesamiento del lenguaje natural, como categorizadores de partes de la oración o analizadores sintácticos, y de recursos lingüísticos, como los corpus escritos. Pero por otra parte, las tecnologías del habla –muy especialmente el reconocimiento– han venido utilizando en los últimos veinte años métodos de trabajo de naturaleza estocástica después de un periodo dedicado a intentos de construir sistemas basados en reglas (knowledge-based systems); esto ha sido posible gracias a la disponibilidad de grandes corpus de datos que permiten extraer automáticamente modelos de unidades y modelos de lenguaje de tipo probabilístico. En el tratamiento del texto, estas mismas técnicas han sido recientemente adoptadas para el etiquetado en partes de la oración (tagging) y el análisis sintáctico (parsing), en la traducción automática basada en corpus bilingües paralelos convenientemente alineados y en lexicografía.

---

<sup>36</sup> Vid., por ejemplo, para el español Rodríguez et alii. (1992).

Church y Mercer (1993) definen este tipo de investigación como empírica y muestran su aplicación en las áreas anteriormente mencionadas. Asistimos pues a la incorporación de herramientas propias del procesamiento del lenguaje natural en los sistemas desarrollados por las tecnologías del habla, al tiempo que los métodos utilizados en reconocimiento se aplican a diversas áreas del tratamiento del texto. La experiencia en ambos campos ha llevado al desarrollo de sistemas denominados híbridos, que combinan aproximaciones estadísticas con aproximaciones basadas en reglas, intentando aprovechar las ventajas de cada una de ellas para mejorar los resultados obtenidos. A continuación se presentan sucintamente dos campos en los que confluyen el procesamiento del lenguaje natural y las tecnologías del habla: se trata de los *sistemas de diálogo persona-máquina* y de la *traducción automática oral*.

#### 4.1. Los sistemas de diálogo

Hemos indicado al principio del capítulo (Cf. 2.1) que uno de los objetivos principales de las tecnologías del habla es conseguir una comunicación natural entre sistemas informáticos y hablantes humanos. Para ello, es necesario disponer de métodos de reconocimiento y de síntesis que lleven a cabo la entrada y salida vocal en el ordenador, pero también se requieren conocimientos relacionados con el modo cómo se realiza la interacción entre hablantes y, para ciertas tareas, la posibilidad de generar mensajes adecuados a la situación, lo que requiere inevitablemente un cierto grado de comprensión del lenguaje.

Los sistemas de diálogo –spoken language systems<sup>37</sup>– pueden concebirse como un conjunto de módulos, del tipo de reconocedor de habla y conversor de texto a habla, analizador, base de conocimientos y generador de unidades lingüísticas, cada uno de los cuales realiza una tarea concreta guiado por un módulo principal que lleva a cabo el control del diálogo entre la persona y la máquina, proporcionando respuestas adecuadas y coherentes con las preguntas, prediciendo las reacciones del usuario, y tratando aspectos propios de la interacción verbal como, por ejemplo, la elipsis y la anáfora.

Los sistemas de diálogo se definen atendiendo a una serie de parámetros, entre los cuales cabe citar el entorno, la tarea y el usuario. En lo que se refiere al entorno, existen sistemas multimodales en los que la interacción con el usuario se realiza no sólo mediante el habla, sino también recurriendo a información visual o a procedimientos basados en el contacto con una pantalla por ejemplo, pulsando sobre

---

<sup>37</sup> Vid. Pierrel (1988), Luzatti (1993), Peckham (1993) y Boves et alii (1995).

una imagen determinada para obtener más información. Por otra parte, los sistemas basados exclusivamente en el intercambio de información vocal pueden dividirse entre los que llevan a cabo la interacción con el usuario mediante un menú o los que propiamente se denominan sistemas de diálogo, en los que se intenta modelar de la manera más fiel posible la comunicación entre dos personas.

Los diálogos mediante menús en los que el usuario tiene que pronunciar el número correspondiente a la información a la que desea acceder después de haber escuchado la lectura de todas las posibilidades del sistema son útiles, por ejemplo, en servicios de información telefónica que cubran un campo bien delimitado. En el caso de la información municipal que citábamos al referirnos a las aplicaciones de la conversión de texto a habla, la persona que realiza la llamada recibiría mensajes del tipo «diga 1 si desea información sobre farmacias de guardia, 2 si desea información sobre cines», de modo que cada vez se va restringiendo más el ámbito, al igual que en los menús y sub-menús de las aplicaciones informáticas, hasta llegar a los datos que se desea obtener. Puede deducirse fácilmente que si el número de posibilidades que ofrece el sistema es muy elevado, un diálogo mediante menús es casi impracticable, requiriéndose entonces una interacción más compleja basada en un sistema que sea capaz de operar con el principio de pregunta-respuesta.

Un segundo parámetro importante en la caracterización de los sistemas de diálogo es el tipo de tarea que se realiza. En general, el usuario que accede a un sistema de diálogo tiene un objetivo concreto, que puede ser la obtención de información sobre horarios de vuelos, por ejemplo o de un servicio la reserva de un billete. En este último caso, se requiere a menudo una negociación, ya que el cliente buscará, como es natural, los horarios y precios que mejor se adecúen a sus necesidades. Si en el primer caso un sistema basado en menús podría ser adecuado, en el segundo sería muy costoso llevar a cabo la operación deseada por este procedimiento.

Por otra parte, en la interacción entre personas suelen tener lugar intercambios que no están directamente relacionados con el objetivo primario del diálogo, y que han sido descritos ampliamente por los estudiosos de la conversación. Se trata de secuencias con una clara función fática, de comentarios sobre el tiempo y la salud, o de informaciones adicionales relativamente irrelevantes –por ejemplo, una explicación de los motivos personales de un viaje a la hora de reservar un billete–. Estos «sub-diálogos» deben ser claramente diferenciados de la tarea principal por un sistema automático. Una de las técnicas que puede utilizarse para ello es la detección de palabras clave –word spotting– en el módulo de reconocimiento, de modo que aunque el sistema no procese todo el mensaje, identifique aquellas palabras que son relevantes para la tarea.

Finalmente, en el momento de diseñar un sistema de diálogo es importante tener también en cuenta el tipo de usuario al que va dirigido. Además de las características relacionadas con la edad, el nivel cultural o las posibles discapacidades, debe considerarse la experiencia con el uso de sistemas que impliquen la interacción con un ordenador. Por ello, la estrategia de diálogo será distinta en una aplicación pensada para usuarios experimentados que en una a la que pueda acceder cualquier persona. Queda patente pues que un sistema de diálogo integra tanto técnicas de síntesis y de reconocimiento como datos sobre el comportamiento verbal humano en una situación de interacción. Estos abarcan desde la desambiguación de enunciados, la identificación del momento en que se produce un cambio de tema, la detección de los turnos de palabra o la decisión sobre la intencionalidad del hablante. Para ello es necesario llegar, en un sistema realmente eficaz, a la comprensión del lenguaje, utilizando las técnicas propias del procesamiento de la lengua escrita (Cf. 3.2.2).

Es interesante presentar, aunque sea sucintamente, los métodos utilizados en el diseño de los sistemas de diálogo (Fraser y Gilbert 1991). En primer lugar, el investigador puede recurrir a su intuición, intentando prever el conjunto de preguntas y respuestas en una tarea concreta; sin embargo, resulta difícil definir toda la gama de reacciones posibles de un usuario, de modo que suelen introducirse preguntas que delimiten claramente las respuestas. La observación de diálogos naturales es una alternativa que ofrece la posibilidad de definir con mayor realismo tanto el vocabulario como los modelos lingüísticos necesarios para el desarrollo de una aplicación; en cambio, surge el inconveniente de la notable diferencia entre la conversación natural y la interacción con una máquina.

Por este motivo se ha desarrollado una segunda aproximación al diseño de los sistemas de diálogo que consiste en la simulación de una interacción persona-máquina: cuando el usuario cree estar en contacto con un ordenador, en realidad se comunica con un operador humano que escribe las respuestas adecuadas en un conversor de texto a habla, limitándose a un conjunto de posibilidades previamente establecidas. Este procedimiento, conocido como el *protocolo del Mago de Oz*, permite recoger corpora de diálogo simulado en condiciones iguales a las que tendrían lugar con un sistema completamente desarrollado; partiendo del análisis de tales corpora se lleva a cabo de forma más segura el diseño de un prototipo del sistema. Una vez que la primera versión es operativa, puede continuarse el desarrollo mediante la técnica denominada *sistema en bucle* –system-in-the-loop–, en la cual el propio sistema va adaptándose en función de la experiencia progresivamente adquirida.

Los sistemas de diálogo son útiles en todas aquellas situaciones en las que se requiera una interacción. Actualmente existen aplicaciones en el campo de los servicios telefónicos de información (Santos et alii 1994), aunque en muchos casos se trata de procedimientos basados en menús, en los que el usuario indica sus preferencias mediante un número; esto implica un reconocedor de dígitos aislados con vocabulario limitado e independiente del locutor, y un procedimiento de síntesis para proporcionar la información adecuada. Sin embargo, puede pensarse en aplicaciones más complejas como la consulta a grandes bases de datos –por ejemplo un catálogo de biblioteca, los archivos de un periódico o de una empresa– introduciendo, además de un sistema de reconocimiento más sofisticado, módulos que permitan traducir la petición del usuario al lenguaje de consulta de los programas de bases de datos. Estos módulos se conocen como *interfaces en lenguaje natural* y su desarrollo constituye un ámbito propio del tratamiento del texto.

En cambio, existen aplicaciones en las que el objetivo es llevar a cabo una transacción, relacionada, por ejemplo, con la adquisición de billetes o con servicios bancarios. Como hemos indicado anteriormente, para ello es necesario el desarrollo de un módulo capaz de tratar correctamente el diálogo y que esté dotado de un cierto grado de comprensión del lenguaje. Algunas de estas aplicaciones han sido ya desarrolladas: Voyager (Glass et alii 1995), ATIS (Bonneau-Maynard et alii 1993) o PEGASUS (Zue et alii 1994) son sistemas de diálogo en el ámbito de los viajes que incorporan a su vez técnicas de procesamiento de las lenguas (Seneff 1991).

#### 4.2. La traducción oral automática

Tal vez el objetivo en el que se concentran actualmente más esfuerzos por integrar el procesamiento del habla y el tratamiento del texto sea la traducción oral automática<sup>38</sup>. Un sistema de traducción oral automática permite a dos personas de diferente lengua materna mantener una conversación hablando cada uno en su lengua respectiva. Para alcanzar tal objetivo, es necesario en primer lugar disponer de un módulo de reconocimiento de habla –a ser posible de habla continua– que convierta en texto el enunciado producido en la lengua-fuente; éste es traducido a la lengua-objeto por un sistema de características similares a los descritos en el apartado 3.2.1., y transformado finalmente en un mensaje oral mediante un sistema de síntesis; un módulo de diálogo controla la interacción entre los dos interlocutores. Naturalmente, este proceso debe realizarse después en sentido inverso, y es necesario que se lleve

---

<sup>38</sup> Vid. de Amores et alii (1994) donde se realiza una exposición general de un sistema de traducción oral automática español-inglés.

a cabo en tiempo real, de modo que no exista una larga espera entre cada una de las intervenciones de los usuarios.

Los sistemas de traducción automática que se están desarrollando en la actualidad son aptos, de momento, para aplicaciones muy restringidas –por ejemplo transacciones bancarias simples como el cambio de divisas– debido a la complejidad lingüística y computacional de la operación. Existen prototipos como AURA que trabaja en inglés y japonés (Morimoto et alii 1993), JANUS para el alemán y el inglés (Woszczyna et alii 1993) o VEST en español e inglés (Roe et alii 1991). Aunque muchas de las aplicaciones se orientan hacia servicios telefónicos, el proyecto alemán Verbmobil (Wahlster 1993) tiene como objetivo la traducción de diálogos cara a cara y constituye hoy en día la iniciativa más importante en este campo.

## **5. LA RELACIÓN CIRCULAR ENTRE TEORÍA LINGÜÍSTICA Y LINGÜÍSTICA COMPUTACIONAL**

### **5.1. Las aportaciones de la teoría lingüística a la Lingüística computacional**

La relativa juventud de la Lingüística aplicada obliga a justificar continuamente la pertinencia de los distintos campos de saber que aglutina. El caso de la Lingüística computacional es, sin embargo, un tanto particular en este sentido: la naturaleza de sus herramientas y sus métodos la individualizan con facilidad pero, sobre todo, son sus productos, visibles y atrayentes, los que garantizan su consideración como un campo de estudio bien delimitado. Como cualquier otra área de conocimiento la Lingüística computacional puede definirse mediante tres parámetros: el objeto que estudia, la metodología con que aborda ese estudio y la finalidad que persigue. De entrada, esta disciplina se ocupa de elaborar teorías y procedimientos para conseguir el tratamiento automático de las lenguas. De ahí que, desde el punto de vista metodológico, posea un carácter híbrido, a medio camino entre informática y lingüística. Su objetivo, por fin, se cifraría en obtener productos tecnológicos relacionados con las industrias de la lengua.

En los primeros años de desarrollo de este campo, el procesamiento de las lenguas parecía un objetivo sencillo, que podría conseguirse trabajando palabra por palabra. La ausencia de fundamentos lingüísticos de esta aproximación originaba, sin embargo, productos deficientes; se estaba trabajando con datos de lengua como si tuviesen la estabilidad de los números y pudiesen desvincularse del texto en que se integran, las condiciones en que tiene lugar el acto de comunicación o las expectativas que el emisor alberga sobre la información manejada por su interlocutor. El desastre de ese acercamiento condujo a una segunda generación de modelos que tomaban, aún de un modo desperdigado y ecléctico, las herramientas



gramaticales disponibles para la formalización de esos datos. Con el tiempo, el desarrollo de la psicolingüística, la semántica cognitiva y la pragmática ha dado una nueva dimensión al diseño de los productos correspondientes, a la vez que apunta una nueva línea de trabajo. Como se ha señalado repetidas veces a lo largo de este capítulo, no cabe esperar la consolidación de la Lingüística computacional fuera del marco general de la Lingüística. Y esto no tiene nada de extraño: tampoco cabría esperar un progreso de los sistemas de computación y gestión informática sin el desarrollo de la matemática o de la teoría de lenguajes. Por supuesto, la consideración de este marco disciplinar como un saber independiente no implica su desvinculación de la Lingüística, entendida ésta como etiqueta general, que abarca proyecciones teóricas y aplicadas. Al fin, sólo se trata de subrayar que la Lingüística computacional lleva a cabo una investigación lingüística y no, simplemente, aplica útiles informáticos a datos procedentes del lenguaje<sup>39</sup>. En términos generales se diría que si un objeto de investigación existe independientemente de la investigación que sobre él se realiza, da lugar a una investigación básica; mientras que, si surge como resultado de una investigación previa, ésta será de tipo aplicado<sup>40</sup>. La Lingüística computacional se encuadra en el grupo de las disciplinas aplicadas porque proyecta determinados conocimientos sobre el funcionamiento de las lenguas a la resolución de problemas concretos.

Desde el punto de vista epistemológico, los productos generados por la Lingüística computacional podrían parecer carentes de interés. Algunos de esos productos, como la síntesis del habla, resuelven problemas que sólo lo son en sociedades con un alto desarrollo tecnológico. Otros, como la traducción automática, realizan tareas que pueden abordarse –incluso con mejores resultados– sin ese recurso a los ordenadores. Por fin, un último grupo, representado por las bases de datos lingüísticos, ni siquiera tienen una utilidad directa, sino que se aplican únicamente a la propia investigación. Todos ellos exigen un enorme despliegue de recursos técnicos y saberes especializados pero, aparte del impacto casi mágico con que sacuden a la sociedad, sus aportaciones podrían tacharse de bastante limitadas. Llevando al extremo este planteamiento, la *Lingüística computacional* sería una disciplina derivada de los avances teóricos en *Lingüística e informática*, que

---

<sup>39</sup> En este sentido, se han planteado continuas precisiones terminológicas. La propia etiqueta *lingüística computacional* puede ser cuestionada. Si en estas páginas se mantiene esta denominación es por su mayor difusión y porque cumple el requisito de destacar el papel lingüístico frente a denominaciones como *informática lingüística*.

<sup>40</sup> Estas consideraciones sobre lingüística teórica y aplicada han sido apuntadas por Slama-Cazacu (1980, 1981) y ampliamente ratificadas por Fernández Pérez (1986). Para más indicaciones al respecto véase la introducción a este volumen a cargo de esta última autora.

diseñaría productos más o menos útiles, pero que nunca generaría nuevas interrogantes. Sin embargo, cualquier valoración de los avances logrados por este campo, incluso una valoración de orden estrictamente epistemológico, exige relacionar el uso de herramientas informáticas con una rapidez y una eficacia en las tareas de bajo nivel que permite ocupar el tiempo humano –bastante más valioso que el de la máquina– en las cuestiones más creativas e interesantes. En pocas palabras, las industrias de la lengua aportan una mayor calidad a la vida humana. Al fin, esta condición práctica es una constante de las disciplinas aplicadas: el diagnóstico en *medicina* no es más que la aplicación del conocimiento obtenido en el estudio de las patologías, y los procesos de control en *ingeniería*, sin los cuales no tendríamos obras públicas ni infraestructuras de comunicación, emanan del saber alcanzado en la matemática aplicada.

También desde el punto de vista de la teoría lingüística, muchas de las soluciones requeridas por la gestión informática de los datos gramaticales producen un esperable rechazo. Son soluciones locales, válidas únicamente en dominios restringidos y tienen mucho de argumentos ad hoc. El procesamiento de las lenguas ha abierto un debate entre la adecuación teórica de las hipótesis lingüísticas y sus posibilidades de solvencia. La gramática ha sido una disciplina fuertemente especulativa y, aún en las visiones más empíricas, no precisa desenvolverse sobre un conjunto de reglas lógicas que den coherencia al desarrollo de unas premisas iniciales. Sin embargo, los desarrollos computacionales precisan una formalización poderosa. Una vez más, la cuestión puede enfocarse desde una perspectiva ligeramente distinta. Las industrias de la lengua constituyen hoy un mercado influido por factores como la naturaleza de sus clientes –normalmente organizaciones internacionales o gubernamentales, industrias o centros de investigación– o el tipo de información que gestionan –conocimiento científico o tecnológico, leyes y comercio–. Sólo su expansión a otros dominios puede hacerles perder gran parte de ese localismo. Además, si la investigación en este terreno no ha avanzado con mayor rapidez no es a causa de limitaciones de tipo teórico, sino porque nuestros conocimientos sobre el lenguaje son todavía más pobres de lo que suponemos. Las aplicaciones computacionales dependen, hoy más que nunca, de una *teoría lingüística* que las avale y les proporcione el apoyo formal imprescindible para la gestión de sus datos.

## **5.2. Las aportaciones de la Lingüística computacional a la teoría lingüística**

En la historia del conocimiento, las disciplinas de corte teórico han permanecido excesivamente separadas de sus aplicaciones. Los investigadores

dedicados a empresas teóricas sospechan de la excesiva solicitud al rendimiento práctico, material o no, de las tecnologías; mientras que quienes se afanan en esta cara de la realidad critican toda aproximación teórica por su desvinculación de los problemas inmediatos. Aunque esta disociación, convertida a menudo en desprecio mutuo, es habitual en todas las ramas del saber, en el caso de la Lingüística se lleva al extremo. Frente a una gruesa tradición gramatical, la mayoría de las disciplinas aplicadas en Lingüística no se han reconocido como tales hasta época muy reciente. En ocasiones porque su cometido no quiso verse como un objeto de estudio relevante –como ocurre con la *Planificación lingüística*– y, en ocasiones, si la práctica correspondiente era una necesidad dictada por la realidad, –como ocurre con la *Traducción* o con la *Didáctica de lenguas*– porque se confiaban a especialistas diferentes del propio lingüista. La *Lingüística* quedaba a salvo de cualquier problema práctico. En el caso que nos ocupa la disociación viene dictada porque su juventud es un imperativo material: hace pocos años que existen soportes técnicos suficientes como para sustentar una *lingüística computacional* de perfil específico.

Ni la *Lingüística computacional* ni ninguna rama de la *Lingüística aplicada* puede considerarse un mero derivado de las propuestas teóricas. Al contrario, teoría y aplicaciones mantienen una relación circular que enmarca su mutua interdependencia. Es cierto que el procesamiento automático de las lenguas debe mucho a la aparición de modelos como la gramática generativa, que daba por primera vez cobertura a un conjunto de reglas con las que condensar la gramática de las lenguas. De modo similar, la lingüística de corpus hunde sus raíces en la tradición filológica de acopio de datos presente ya en los neogramáticos y modelizada de modo más preciso por la gramática de valencias. Pero, del mismo modo, la aplicación a los productos informáticos de presupuestos emanados de estas corrientes sirvió para poner en evidencia algunos defectos de las hipótesis que manejaban.

En el estadio actual de desarrollo cabe plantearse si, realmente, las industrias de la lengua son suficientemente poderosas como para sustentar un disciplina independiente, al modo de la *Sociolingüística*, la *Gramática* o la *Teoría de la traducción*. Por supuesto, el planteamiento es metodológico y tiene que ver con la naturaleza de los problemas implicados o el tipo de conocimiento que se emplea para resolverlos. Lejos de toda valoración sobre su eficacia, se trata de considerar la ubicación de la *Lingüística computacional* en el conjunto de las disciplinas lingüísticas. Quizá, si se alcanzase un grado ideal de desarrollo, cada uno de los productos generados por las industrias de la lengua deberían considerarse campo propio de las diferentes ramas de la *Lingüística teórica* a las que remiten. De hecho, no hay más conexión entre síntesis del habla y *traducción automática*, o entre la construcción de correctores gramaticales y los sistemas de diálogo entre ser humano

y máquina que el factor, secundario, de compartir unas mismas herramientas. También comparten herramientas la *Semántica* y la comprensión automática de textos y no por eso se agrupan juntas. En definitiva, si el progreso de estas industrias fuera óptimo, podrían ser absorbidas por otras disciplinas lingüísticas. El tratamiento del habla, tanto en la fase de síntesis como en la de reconocimiento, podría formar parte de la *Fonética experimental* pues, en realidad, son los fonetistas quienes desempeñan el papel lingüístico en la empresa. La *traducción automática* podría agruparse, con sus características propias, dentro de la *Teoría de la traducción*. La comprensión y la generación de textos podría verse como parte de una teoría gramatical post-estructural, donde cabe el recurso a la *Semántica* y a la *Pragmática*. Los diccionarios automáticos constituirían simplemente un modo particular de hacer *Lexicografía* y los correctores ortográficos o gramaticales serían la aplicación de los conocimientos clásicos de la *Lingüística descriptiva* de tipo normativo.

Por supuesto, esta restricción es peligrosa. Todas las manifestaciones de la *Lingüística aplicada* mantienen dependencias similares con la *Lingüística teórica* –la didáctica de lenguas debe mucho a la teoría gramatical, las patologías lingüísticas a la neurolingüística o la planificación a la sociolingüística– y, sin embargo, se han emancipado de esas visiones teóricas porque aportan un tipo de información ausente de la perspectiva inicial. Pues bien, en el caso de la Lingüística computacional es necesario reconocer una correspondencia similar. Los mecanismos tecnológicos son, a veces, tan llamativos, que se acaba calificando de *Lingüística computacional* cualquier aplicación de los ordenadores a un campo tradicionalmente humanístico. El uso de ordenadores para la investigación en *Lexicografía* no garantiza que se esté haciendo *Lexicografía computacional*, y su aplicación a la *Gramática* no la convierte en *Gramática automática*. Es cierto que con frecuencia el análisis de datos es un proceso lento y laborioso y la posibilidad de automatizarlo resulta especialmente atractiva. Precisamente por ello las nuevas tecnologías son en la actualidad herramientas indispensables para cualquier tipo de investigación, pero las herramientas mismas no transforman la naturaleza de esa investigación.

Parece fructífero, entonces, además de metodológicamente necesario, distinguir entre *Lingüística computacional* y la mera aplicación de útiles informáticos a los datos lingüísticos. Es posible que la barrera sea difícil de establecer pero eso no justifica el desinterés general por trazarla. El terreno específico de la *Lingüística computacional* puede cifrarse en la elaboración de productos derivados del enfoque teórico pero, al tiempo, esa elaboración deber revertir en la calidad de nuestros conocimientos. La *Lingüística computacional* parece un terreno idóneo para la verificación de las teorías lingüísticas (Grishman 1986, Shieber 1987), porque está poniendo en evidencia fenómenos de las lenguas escasamente conocidos y, a medida

que pretenda diseñar gramáticas para autómatas irá desvelando nuevas cuestiones. Baste, en este sentido, con notar cómo la *Teoría de la traducción* empezó a despegarse de la limitada óptica tradicional justo cuando los procesos automáticos mostraron la necesidad de contemplar el texto desde los presupuestos generales de la *Pragmática* y la *Lingüística del texto*. De modo similar, la lingüística de corpus, al proporcionar ingentes cantidades de datos reales sobre las lenguas, ha mostrado que las gramáticas son bastante menos regulares de lo que las teorías suelen considerar; que la aparente armonía de los sistemas lingüísticos, sustentada sobre un haz de oposiciones bien delimitado, se resquebraja ante la reiterada aparición de fenómenos periféricos, que no obedecen a reglas simples. La comprensión y la generación de textos, no sólo consiguen útiles de interés social; también obligan a tomar en consideración procesos cognitivos ignorados por la *teoría lingüística* y abren nuevas vías de progreso a la *Psicolingüística* y a la *Neurolingüística*. La síntesis y el reconocimiento del habla pueden sustentar en el futuro una *teoría fonológica*, probablemente menos brillante, pero más realista. En definitiva, los problemas que estas aplicaciones encuentran para la construcción de sus productos son una fuente empírica de datos lingüísticos. Si la teoría no toma en cuenta esos datos, simplemente estará describiendo una ficción. Tanto para incrementar nuestro conocimiento como para concebir productos más perfeccionados, que sirvan a las necesidades del mundo moderno, parece imprescindible orientar la investigación en lingüística computacional hacia objetivos teóricos y fundamentales.

## REFERENCIAS BIBLIOGRÁFICAS

- Aarts, J. y W. Meijs (1990), *Theory and Practice in Corpus Linguistics*, Amsterdam, Rodopi, 1990.
- Aarts, J. y W. Meijs, eds. (1984), *Corpus Linguistics. Recent Developments in the Use of Corpora in English Language Research*, Amsterdam, Rodopi, 1984.
- Aarts, J. y W. Meijs, eds. (1986), *Corpus Linguistics II. New Studies in the Analysis and Exploitation of Computer Corpora*, Amsterdam, Rodopi, 1986.
- Aarts, J. y T. van den Heuvel (1985), «Computational tools for syntactic analysis of corpora», *Linguistics* 23, 303-35.
- Aarts, J.; P. de Haan y N. Oostdijk, eds. (1993), *English language corpora: design, analysis and exploitation. Papers from the Thirteenth International Conference on English Language Research on Computerized Corpora, Nijmegen 1992*, Amsterdam-Atlanta, Rodopi, 1993.
- Aguilar, L.; J. M. Fernández; J. M. Garrido; J. Llisterrí; A. Macarrón; L. Monzón y M. A. Rodríguez (1994), «Diseño de pruebas para la evaluación de habla sintetizada en español y su aplicación a un sistema de conversión de texto a habla», *Boletín de la Sociedad española para el procesamiento del lenguaje natural*, 10, 1994.
- Aijmer, K. y B. Altenberg, eds. (1991), *English Corpus Linguistics. Studies in Honour of Jan Svartvik*, London, Longman, 1991.
- Ainsworth, W.A., ed. (1990), *Advances in speech, hearing and language processing*, 1, London, JAI Press, 1990.
- Ainsworth, W.A., (1988): *Speech Recognition by Machine*, London, Peter Peregrinus Ltd., 1988.
- Alvar Ezquerro, M. y J. A. Villena Ponsada, eds. (1994), *Estudios para un corpus del español*, Anejo 7 de *Analecta Malacitana*, Univ. de Málaga, 1994.
- Allen, J. (1985): «A perspective on man-machine communication by speech», *Proceedings of the IEEE*, 73/11, 1985, 1541-50.
- Allen, J.; M. S. Hunnicutt y D. H. Klatt (1987), *From Text to Speech: The MITalk System*, Cambridge University Press, 1987.
- Amores Carredano, J.G. de, J.F. Quesada Moreno y D. Tapias Merino (1994), «Traducción automática basada en el formalismo LFG con entrada y salida por voz», *Comunicaciones de Telefónica I+D*, 5/2, 1994, 132-147.

- Austin, J. (1962), *How to do things with words*, Oxford University Press. Trad. esp. de G. Carrió y E. Rabossi, *Palabras y acciones*, Buenos Aires, Paidós, 1971.
- Baker, J. M. (1993a), «Using speech recognition for dictation and other large vocabulary applications», *Applications of Speech Technology. Proceedings of a Joint ESCA-NATO/RSG 10 Tutorial and Workshop. Lautrach Conference Center, Bavaria, Germany, September 1993*, 1993, 103-111.
- Baker, J. M. (1993b), «Dictation, directories and databases. Emerging PC applications for large vocabulary speech recognition», *Eurospeech'93*, 1, 1993, 3-10.
- Baker, M.; G. Francis y E. Tognini-Bonelli, eds. (1993), *Text and Technology: In Honour of John Sinclair*, Amsterdam-Philadelphia, John Benjamins, 1993.
- Bateman, J. A. y E. H. Hovy (1992), «Computers and text generation: principles and uses» en Butler, ed., 1992, 53-74.
- Bloothoof, G.; V. Hazan; D. Huber y J. Llisterri, eds. (1995), *European Studies in Phonetics and Speech Communication*, Utrecht, Publications of the OTS, 1995.
- Boguraev, B. y E. Briscoe, eds. (1989), *Computational lexicography for natural language processing*, London, Longman, 1989.
- Bonneau-Maynard, H.; J. L. Gauvain; D. Goodine; L. F. Lamel; J. Polifroni y S. Seneff (1993), «A French Version of the MIT-ATIS System: Portability Issues», *Eurospeech'93*, 3, 1993, 2059-2062.
- Boves, L.; P. Dalsgaard y L. B. Larsen, eds. (1995), *Proceedings of the ESCA Tutorial and Research Workshop on Spoken Dialogue Systems. Vigs, Denmark, May 30-June 2, 1995*.
- Bristow, G., ed. (1984), *Electronic Speech Synthesis. Techniques, Technology and Applications*, London, Granada, 1984.
- Bristow, G., ed. (1986): *Electronic Speech Recognition. Techniques, Technology and Applications*, London, Collins, 1986.
- Bullon, J. L. y J. C. Pérez (1994), «Conversión de texto a voz en castellano aplicando el algoritmo PSOLA», *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 10, 1994, 217-232.
- Burnage, G. y D. Dunlop (1993), «Encoding the British National Corpus», en AARTS et alii, eds., 1993, 79-95.

- Butler, C. S. (1985), *Computers and Linguistics*, Oxford, Basil Blackwell, 1985.
- Butler, C. S., ed. (1992), *Computers and Written Texts*, Oxford, Basil Blackwell, 1992.
- Calliope (1989), *La parole et son traitement automatique*, J. P. Tubach, ed, Paris, Masson, 1989.
- Carlson, R.; B. Granström y S. Hunnicutt (1990), «Multilingual text-to-speech development and applications» en Ainsworth, ed., 1990, 269-296
- Carré, R.; J.-F. Degremont; M. Gross; J.-M. Pierrel y G. Sabah (1991), *Language Humain et Machine*, Paris, Presses du CNRS, 1991.
- Casacuberta, F. y E. Vidal (1987), *Reconocimiento automático del habla*, Barcelona, Marcombo-Boixareu Editores, 1987.
- Casajuana, R. y C. Rodríguez (1985), «Verificación ortográfica en castellano: la realización de un diccionario en ordenador», *Español Actual* 44, 1985.
- Computational Linguistics* (1980) -. Cambridge, The MIT Press.
- Computer Speech and Language* (1986) -. London, Academic Press.
- Computers and the Humanities* (1966) -. Dordrecht, Kluwer.
- Coseriu, E. (1974), «Los universales del lenguaje (y los otros)», *Gramática, Semántica, Universales*, Madrid, Gredos, 1978, 148-205.
- Chollet, G. (1994), «Automatic Speech and Speaker Recognition: Overview, Current Issues and Perspectives» en E. Keller, ed., 1994, 129-148
- Chomsky, N. (1965), *Aspects of the Theory of Syntax*, Cambridge (Mass.), MIT Press. Trad. esp., introducción, notas y apéndice de C.P. Otero, *Aspectos de la teoría de la sintaxis*, Madrid, Aguilar, 1970.
- Church, K. W. y R. L. Mercer (1993), «Introduction to the special issue on computational linguistics using large corpora», *Computational Linguistics* 19/1, 1993, 1-24.
- Damper, R. I. (1990), «Speech aids for the handicapped» en Ainsworth, ed., 1990, 297-332
- Danlos, L. (1985),  *Génération automatique de textes en langues naturelles*, Paris, Masson, 1985.
- Danlos, L. (1993), «Génération automatique de textes», en Fuchs, ed., 1993, 247-266.



- Dowty, R. D.; L. Karttunen y A. M. Zwicky (1985), *Natural Language Parsing*, Cambridge University Press, 1985.
- Elsnet (1993), *European PhD Studies in Language and Speech. Report by the ELSNET Training and Mobility Task Group, April, 1993. (ELSNET: OTS, Utrecht University, Trans 10, 3512 JK Utrecht, The Netherlands)*, 1993.
- Enríquez, E. (1991), «El problema de las ambigüedades fonéticas y su tratamiento automático», *Boletín de la Real Academia de la Lengua Española* 71-72, 1991, 157-183.
- EUROSPEECH'93, *3rd European Conference on Speech Communication and Technology, Berlin, Germany, 21-23 September 1993*.
- Fant C. G. (1984), «Phonetics and speech technology» en M. P. R. van den Broecke y A. Cohen, eds., *Proceedings of the Tenth International Congress of Phonetic Sciences*, Dordrecht, Foris, 1984, 13-24.
- Fant C. G. (1960), *Acoustic Theory of Speech Production*, The Hague, Mouton, 1960.
- Fernández Pérez, M. (1986), «Las disciplinas lingüísticas», *Verba* 13, 1986, 15-73.
- Fisher, M. (1986), «Voice Control for the Disabled» en G. Bristow, ed., 1986, 309-321.
- Flanagan, J. L. (1972), «The synthesis of Speech», *Scientific American* 226/2, 1972, 45-58.
- Flanagan, J. L. y L. R. Rabiner, eds., (1973), *Speech Synthesis*, Stroudsburg (Penn.), Dowden, Hutchinson & Ross Inc., 1973.
- Fraser, N. M. y G. N. Gilbert (1991), «Simulating speech systems», *Computer Speech and Language* 5, 1991, 81-99.
- Fuchs, C. y B. Victorri (1993), «Compréhension automatique de textes» en C. Fuchs, ed., 1993, 223-246.
- Fuchs, C., ed., (1993), *Linguistique et traitement automatique des langues*, Paris, Hachette, 1993.
- Gabioud, B. (1994), «Articulatory Models in Speech Synthesis» en E. Keller, ed., 1994, 215-230.
- Gagnoulet, C.; D. Jovet y J. Damay (1991), «MARIEVOX: A voice-activated information system», *Speech Communication* 10/1, 1991, 23-32.

- Garrido, J. M. (1991), «Estilización de patrones melódicos del español para sistemas de conversión texto-habla», *Boletín de la Sociedad española para el procesamiento del lenguaje natural*, 7, 1991, 209-220.
- Gazdar, G. y C. Mellish (1987), «Computational linguistics» en J. Lyons, R. Coates, M. Deuchar y G. Gazdar, eds. (1987): *New horizons in linguistics 2*, London, Penguin, 1987, 225-248.
- Gazdar, G. y C. Mellish (1989a), *Natural language processing in LISP*, Reading (Mass.), Addison-Wesley, 1989.
- Gazdar, G. y C. Mellish (1989b), *Natural language processing in PROLOG*, Reading (Mass.), Addison-Wesley, 1989.
- Glass, J.; G. Flammia; D. Goodine; M. Phillips; J. Polifroni; S. Sakai; S. Seneff y V. Zue (1995), «Multilingual spoken-language understanding in the MIT Voyager system», *Speech Communication* 17, 1/2, 1995, 1-18
- Goldman, N. M. (1975), «Conceptual generation» en R. Schank, ed., 1975, 289-371.
- Grice, H. P. (1971), «Meaning» en D. Steinberg y L. Jacobovits, eds., *Semantics: an interdisciplinary reader in Philosophy, Linguistics, and Psychology*, Cambridge Univ. Press, 1971, 53-59.
- Grishman, R. (1986), *Computational linguistics. An introduction*, Cambridge Univ. Press, 1986. Trad. esp. de A. Moreno Sandoval, *Introducción a la lingüística computacional*, Madrid, Visor, 1991.
- Haan, J. P. de y N. Oostdijk, eds. (1993), *English Language Corpora: Design, Analysis and Exploitation*, Amsterdam, Rodopi, 1993.
- Haton, J.-P.; J.-M. Pierrel; G. Perennou; J. Caelen y J.-L. Gauvin (1991), *La reconnaissance automatique de la parole*, Paris, Dunod, 1991.
- Hearn, P. M. y D. F. Button, eds. (1994), *Language Industries Atlas*, Amsterdam-Oxford-Washington, IOS Press, 1994.
- Helbig, G. y W. Schenkel (1969), *Wörterbuch zur Valenz und Distribution deutscher Verben*, Tübingen, Max Niemeyer, 1983.
- Hernández, I.; J. C. Olabe; P. Etxebarria; B. Etxebarria y A. Cuesta (1994), «Ahozka. Un sistema de conversión de texto a voz para el Euskara», *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural* 10, 241-256.
- Hess, W. J. (1995), «Improving the quality of speech synthesis systems at segmental level», en C. Sorin, J. Mariani, H. Meloni y J. Schoentgen, eds., *Levels in*

- Speech Communication. Relations and Interactions*, Amsterdam, Elsevier Science B.V., 1995, 239-248.
- Hiller, S.; E. Rooney; J. Laver y M. Jack (1993), «SPELL; An automated system for computer-aided pronunciation teaching», *Speech Communication* 13, 3/4, 1993, 463-474.
- Hockey, S. (1980), *A Guide to Computer Applications in Humanities*, London, Duckworth, 1980.
- Holmes, J. N. (1988), *Speech Synthesis and Recognition*, Wokingham, Van Nostrand Reinhold, 1988.
- Hornby, A. S., ed. (1974), *Oxford advanced learner's dictionary of current English*, 3ª ed., Oxford Univ. Press, 1974.
- Hovy, E. H. (1988), *Generating natural language under pragmatic constraints*, Hillsdale, L. Erlbaum Associates Pub., 1988.
- Hutchins, W. J. (1986), *Machine translation: past, present and future*, Chichester, Ellis Horwood, 1986.
- International Journal of Corpus Linguistics* 1995 -. Amsterdam, John Benjamins.
- Janot-Giorgetti, M. T. (1983), *Expériences en reconnaissance de la parole. Application à l'apprentissage des langues: le système MicroLEA*, Hamburg, Helmut Buske Verlag, 1983.
- Johansson, S. y A. Stenström, eds. (1991), *English Computer Corpora: Selected Papers and Research Guide*, Berlin, Mouton de Gruyter, 1991.
- Journal of Natural Language Engineering* 1995-. Cambridge University Press.
- Keller, E., ed. (1994), *Fundamentals of Speech Synthesis and Speech Recognition. Basic Concepts, State of the Art and Future Challenges*, Chichester, John Wiley & Sons.
- Kempen, G., ed. (1987), *Natural language generation. New results in artificial intelligence, psychology and linguistics*, Dordrecht-Boston-Lancaster, Martinus Nijhoff Publ., 1987.
- King, M. (1983), *Parsing Natural Language*, New York, Academic Press, 1983.
- Klatt, D. H. (1987), «Review of Tex-to-Speech Conversion for English», *Journal of the Acoustical Society of America*, 82/3, 1987, 737-793.
- Klatt, D. H. (1980), «Software for a Cascade/Parallel Formant Synthesizer», *Journal of the Acoustical Society of America* 67/3, 1980, 971-995. Reed. en R. D.

- Kent, B. S. Atal y J. L. Miller, eds., *Papers in Speech Communication: Speech Production*, New York, Acoustical Society of America, 1991, 765-789.
- Kyts, M.; O. Ihalainen y M. Rissanen, eds. (1988), *Corpus Linguistics Hard and Soft*, Amsterdam, Rodopi, 1988.
- La Tribune des Industries de la Langue*.- Paris, Observatoire Français et International des Industries de la Langue.
- Laface, P., ed. (1990), *Speech Recognition and Understanding: Recent Advances, Trends and Applications*, 1990, Berlin, Springer-Verlag.
- Language Industry Monitor. The World of Language Computing* (1991)-. Amsterdam, Uitverrij LIM.
- Language Technology*.- Amsterdam, Ink International.
- Leech, G. (1991), «The State of the Art in Corpus Linguistics», en K. Aijmer y B. Altenberg, eds., 1991, 8-29.
- Leech, G. y S. Fligelstone (1992), «Computers and corpus analysis», en C. Butler, ed., 1992, 115-140.
- Leech, G., R. Garside y M. Bryant (1994), «The large-scale grammatical tagging of text: experience with the British National Corpus» en N. Oostdijk y P. de Haan, eds., 1994, 47-63.
- Leitner, G., ed. (1992), *New Directions in English Language Corpora. Methodology, Results, Software Development*, Berlin, Mouton de Gruyter, 1992.
- Levinson, S. E. y M. Y. Liberman (1981), «Speech Recognition by Computer», *Scientific American* 244, 1981, 64-76. Trad. esp. de R. Cerdá, «Reconocimiento del habla por medio de ordenadores» en J. Agullo, ed., *Acústica musical*, Barcelona, Prensa Científica, 1989, 106-121.
- Lewis, D. (1992), «Computers and translation» en C. Butler, ed., 1992, 75-113.
- Lieberman, P. y S. E. Blumstein (1988), *Speech Physiology, Speech Perception and Acoustic Phonetics*, Cambridge University Press, 1988.
- Lindblom, B. (1986), «On the origin and purpose of discreteness and invariance in sound patterns», in J. S. Perkell y D. H. Klatt, eds., *Invariance and Variability in Speech Processes*, Hillsdale, Lawrence Erlbaum, 1986, 496-510.
- Literary and Linguistic Computing* (1986)-. Oxford University Press.

- López Gonzalo, E.; J. Álvarez Cercadillo y L. Hernández Gómez (1994), «Metodología para el modelado prosódico de un sistema de conversión de texto a habla en castellano», *Boletín de la Sociedad española para el procesamiento del lenguaje natural* 10, 1994.
- López Gonzalo, E.; E. Rodríguez Banga; C. García Mateo y L. Hernández Gómez (1994), «Modelado lingüístico y acústico para un sistema de conversión de texto a habla», *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 10, 1994, 257-272.
- Luzatti, D. (1993), «Dialogue homme-machine» en C. Fuchs, ed., 1993, 267-290.
- Llisterri, J. (1988), «La síntesis del habla: estado de la cuestión», *Boletín de la Sociedad española para el procesamiento del Lenguaje Natural*, 6, 1988, 17-42.
- Llisterri, J. (1990), «Algunes reflexions sobre el paper de la lingüística en la tecnologia de la veu», *Límits, Revista d'Assaig i d'Informació sobre les Ciències del Llenguatge*, 8, 19-32.
- Marín Gálvez, R. (1994), «Diseño y evaluación de un modelo de duración vocálica del español para la síntesis del habla», *Boletín de la Sociedad española para el procesamiento del Lenguaje Natural*, 10, 1994.
- Mariño, J. B., C. Nadeu y J. Llisterri (1987), «Síntesis automática del habla», en J. Mompín, coord., 1987, 157-165.
- Martí, J. (1987), «Un conversor text-veu en català: Sistema SINCAT (SINtetitzador de CATALÁ)» en C. Martín Vide, ed., *Actas del II Congreso de lenguajes naturales y lenguajes formales*, Barcelona, PPU, 1987, 197-209.
- Martí, J. (1988), «Síntesis del habla (Evolución histórica y situación actual) en C. Martín Vide, ed., *Actas del III Congreso de lenguajes naturales y lenguajes formales*, 1, Barcelona, PPU, 1988, 213-237.
- Martí, M. A.; I. Castellón y M. Taule (1994), «La lexicografía computacional: representació i adquisició del coneixement lèxic», *Anuari de Filologia*, 5 1994.
- Martínez, M.; J. M. Pardo; A. Borrajo; A. Santos; C. L. Barrio; E. Muñoz y A. Quilis (1986), «Conversión automática texto-habla y su relación con el procesamiento del lenguaje natural» en C. Martín Vide, ed., *Actas del I Congreso de Lenguajes naturales y lenguajes formales*, Barcelona, Universitat de Barcelona, 366-375.

- McDonald, D. D. (1981), *Natural language generation as a computational problem: an Introduction*, Univ. of Massachusetts technical report 81/33, 1981.
- Meijs, W. (1992), «Computers and dictionaries» en C. Butler, ed., 1992, 141-165.
- Meijs, W. ed. (1987), *Corpus Linguistics and Beyond. Proceedings of the Seventh International Conference on English Language Research on Computerized Corpora*, Amsterdam, Rodopi, 1987.
- Minsky, M. (1975), «A framework for representing knowledge» en P. Winston, ed., *The Psychology of computer vision*, New York, McGraw-Hill, 1975.
- Mompín, J., coord. (1987), *Inteligencia artificial. Conceptos, técnicas y aplicaciones*, Barcelona, Marcombo-Boixareu Editores, 1987.
- Montague, R. (1974): *Formal philosophy: selected papers of Richard Montague*, R. Thomason, ed., Yale Univ. Press, 1974.
- Monzón, L.; M. A. Rodríguez y G. Escalada (1993), «Módulo de análisis sintáctico para un sistema de conversión texto-voz en castellano», *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 9, 1993, 367-379.
- Moreno, P. J.; M. Martínez; J. M. Pardo y J. A. Vallejo (1989), «Improving naturalness in a text-to-speech system with a new fundamental frequency algorithm» en J. P. Tubach y J. J. Mariani, eds., *Eurospeech 89. European Conference on Speech Communication and Technology. Paris, September 1989*, Edinburgh, CEP Consultants Ltd., 1, 1989, 360-363.
- Morimoto, T.; T. Takezawa; F. Yato; S. Sagayama; T. Tashiro; M. Nagata y A. Kurematsu (1993), «ATR's speech translation system: AURA», *Eurospeech'93*, 2, 1993, 1291-1294.
- Moulines, E. y F. Charpentier (1990), «Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones», *Speech Communication* 9, 5/6, 1990, 453-468.
- Nadeu, C. y J. B. Marino (1985), «Comunicación oral con el computador», *Mundo Electrónico* 149, 1985, 108-116.
- Nirenburg, S., ed. (1987), *Machine Translation. Theoretical and Methodological Issues*, Cambridge University Press, 1987.
- O'Shaughnessy, D. (1987), *Speech Communication. Human and Machine*, Addison Wesley Publishing Co., 1987.

- Oostdijk, N. y P. de Haan, eds. (1994), *Corpus-based research into language*, Amsterdam-Atlanta, Rodopi, 1994.
- Pardo, J. M. (1988), «Reconocimiento del habla: una introducción», *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural* 6, 1988, 3-16.
- Pechham, J. (1993), «A new generation of spoken dialogue systems: results and lessons from the SUNDIAL project», *Eurospeech'93*, 1, 1993, 33-42.
- Pierrel, J. M. (1988), *Dialogue oral homme-machine: Connaissances linguistiques, stratégies et architecture des systèmes*, Paris, Editions Hermes, 1988.
- Pols, L. C. W. (1991), «Quality assessment of text-to-speech synthesis by rule» en S. Furui y M. M. Sondhi, eds., *Advances in Speech Signal Processing*, Marcel Dekker Inc, 1991, 387-416.
- Procter, P., ed. (1978), *Longman dictionary of contemporary English*, Harlow, Longman, 1978.
- Puigvi, D.; D. Jiménez y J. M. Fernández (1994), «Parametrización de las pausas ortográficas en castellano. Aplicación a un conversor de texto a habla», *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 10, 1994.
- Rabiner, L. y B. H. Huang (1993), *Fundamentals of Speech Recognition*, New York, Prentice Hall, 1993.
- Ríos, A. (1993), «La información lingüística en la transcripción fonética automática del español», *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 9, 1993 381-387.
- Rodríguez Crespo, M. A.; J. G. Escalada Sardina; A. Macarrón Larumbe y L. Monzón Serrano (1993), «AMIGO, Un conversor texto-voz para el español», *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural* 9, 1993, 389-400.
- Rodríguez, C.; C. Rubio; A. Sánchez y L. Sopena (1992), «Herramientas de ayuda a la redacción de textos: un sistema de verificación léxica, sintáctica y estilística», *Voz y Letra. Revista de Filología*, 3/1, 1992, 155-174
- Rodríguez, H.; M. A. Martí e I. Castellón (1995), *Formalismes lògics per al tractament del llenguatge natural*, Barcelona, Universitat Politècnica de Catalunya, 1995.

- Roe, D.; F. Pereira; R. Sproat; M. Riley; P. Moreno y A. Macarrón (1991), «Toward a spoken language translator for restricted-domain context-free languages», *Eurospeech'91. Second European Conference on Speech Communication and Technology. Genova, Italy, 24-26 September 1991*, 3, 1991, 1063-1066.
- Royo, G. (1992), «El futuro 'Diccionario de construcciones verbales del español actual'» en C. Martín Vide, ed., *Actas del VIII Congreso de lenguajes naturales y lenguajes formales*, Univ. de Barcelona, 1992, 41-50.
- Rouault, J. (1989), «Apports, contraintes et limites du langage dans le traitement automatique des langues» en M-J. Reichler-Beguelin, ed., *Perspectives méthodologiques et épistemologiques dans les sciences du langage. Actes du Colloque de Fribourg, 1988*, Bern, Peter Lang, 1989, 141-157.
- Sager, J. C. (1992), *La industria de la lengua, La lingüística computacional- los trabajos de UMIST. La traducción especializada y su técnica*, Universitat de Barcelona, 1992.
- Saito, S., ed. (1991), *Speech Science and Technology*, Amsterdam, IOS Press, 1991.
- Santos, A.; J. Colas y J. Lestani (1994), «Telephone Information Systems: Dialogue Specification Language» en M. Brouwer-Janse y T. L. Harrington, eds., *Human-Machine Communication for Educational Systems Design*, Berlin, Springer-Verlag, 1994, 315-324.
- Santos, A.; P. Muñoz y M. Martínez (1988), «Diseño y evaluación de reglas de duración en la conversión de texto a voz», *Boletín de la Sociedad española para el procesamiento del Lenguaje Natural* 6, 1988, 69-92.
- Schank, R. (1975), «The structure of episodes in memory» en D. Bobrow y A. Collins, eds., *Representation and Understanding: studies in cognitive science*, New York, Academic Press, 1975.
- Schank, R. y R. Abelson (1977), *Scripts, plans, goals and understanding*, Hillsdale, Erlbaum, 1977.
- Schank, R., ed. (1975), *Conceptual information processing*, Amsterdam, North-Holland, 1975.
- Searle, J. (1969), *Speech acts*, Cambridge Univ. Press, 1969. Trad. esp. de L. Valdés, *Actos de habla*, Madrid, Cátedra, 1980.
- Seneff, S. (1991), «TINA: A natural language system for spoken language applications», *Computational Linguistics* 18/1, 1991, 61-86.



- SEPLN 1994. *Actas del X Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural*, Córdoba, 20-22 de julio.
- Shieber, S. M. (1987), «Separating linguistic analyses from linguistic theories» en P. Whitelock, M. McGee Wood, H. Somers, R. Johnson y P. Bennet, eds., (1987): *Linguistic theory & Computer applications*, London-San Diego, Academic Press, 1987, 1-36.
- Sinclair, J.; P. Hanks; G. Fox; R. Moon y P. Stock, eds. (1987), *Collins COBUILD English language dictionary*, London-Glasgow, Collins, 1987.
- Slama-Cazacu, T. (1980), «The circular relation between fundamental and applied research in Linguistics», *Revue Roumaine de Linguistique* 25, 1980, 3-7.
- Slama-Cazacu, T. (1981), «Sur l'objet de la linguistique appliquée», *Revue Roumaine de Linguistique* 26, 1981, 5-21.
- Slocum, J., ed. (1988), *Machine Translation Systems*, Cambridge University Press, 1988.
- Sopeña, L. de (1990), «Nuevas perspectivas de la traducción», *La lingüística aplicada. Noves perspectives - Noves professions - Noves orientacions*, Barcelona: Fundació Caixa de Pensions / Universitat de Barcelona, 1990, 31-37.
- Speech Communication* (1982) -. Amsterdam, Elsevier Science BV.
- Steinbiss, V.; H. Ney; U. Essen; B.-H. Tran; X. Aubert; C. Dugats; R. Kneser; H.-G. Meier; R. Oerder; R. Haeb-Umbach; D. Geller; W. Hsllerbauer y H. Bartosik (1995), «Continuous speech dictation - From theory to practice», *Speech Communication* 17, 1/2, 1995, 19-38.
- Svartvik, J., ed. (1992), *Directions in Corpus Linguistics, Proceedings of Nobel Symposium 82, Stockholm 4-8 August 1991*, Berlin, Mouton de Gruyter.
- Van Bezooijeno. y L. C. W. Pols (1990), «Evaluating text-to-speech systems: some methodological aspects», *Speech Communication* 9/4, 1990, 263-270.
- Van Santen, J. P. H. et alii, eds. (1995), *Progress in Speech Synthesis*, Berlin, Springer, 1995.
- Verdejo, M. F. (1987), «Comprensión automática: lenguaje natural», en J. Mompín, coord., 1987, 137- 145.
- Vidal Beneyto, J., dir. (1991), *Las industrias de la lengua*. Trad. de M. Alvar Ezquerro et alii. Madrid, Fundación Germán Sánchez Ruipérez y Ediciones Pirámide, 1991.

- Wahlster, W. (1993), «Verbmobil: Translation of Face-To-Face Dialogs», *Eurospeech'93*, 1993, 29-38.
- Waibel, A. (1986), «Suprasegmentals in very large vocabulary word recognition», en E. E. Schwab y H. Nusbaum, eds., *Pattern Recognition by Humans and Machines. Volume 1: Speech Perception*, Orlando, Academic Press, Inc, 1986, 159-186
- Waibel, A. y K. F. Lee, eds. (1990), *Readings in Speech Recognition*, San Mateo (CA), Morgan Kaufmann, 1990.
- Williams, N. (1992), «Computers and writing» en C. Butler, ed., 1992, 247-265
- Winograd, T. (1983), *Language as a cognitive process*, Reading Mass., Addison-Wesley, 1983.
- Witten, I. H. (1982): *Principles of Computer Speech*, London, Academic Press, 1982.
- Woszczyna, M.; N. Coccaro; A. Eisele; D. Lavie; A. McNair; T. Polzin; I. Rogina; C. P. Rose; T. Sloboda; M. Tomita; J. Tsutsumi; A. Aoki-Waibel y W. Ward (1993), «Recent advances in JANUS: A speech translation system», *Eurospeech'93*, 2, 1993, 1295-1298.
- Zue, V.; S. Seneff; J. Polifroni; M. Phillips; C. Pao; D. Goodine; D. Goddeau y J. Glass (1994), «PEGASUS: A spoken dialogue interface for on-line air travel planning», *Speech Communication* 15, 3/4, 1994, 323-340.