

Tecnología

Javier Gómez Guinovart

Seminario de Lingüística Informática, Universidade de Vigo

<jgomez@uvigo.es>

Perspectivas de la lingüística computacional

Dave: *Abre las compuertas del muelle de salida, Hal.*
 HAL: *Lo siento, Dave, me temo que no puedo hacerlo.*
 Dave: *¿Cuál es el problema?*
 HAL: *Creo que lo sabes tan bien como yo.*
 Dave: *No sé de qué estás hablando.*
 HAL: *Sé perfectamente que tú y Frank estabais pensando en desconectarme y me temo que esto es algo que no puedo permitir que suceda.*

(2001: una odisea del espacio, de Stanley Kubrick)

1. Introducción

Uno de los objetivos centrales de la lingüística computacional es el de permitir el uso oral de la lengua materna como medio de comunicación entre los ordenadores y los individuos. Se trata de permitir que las personas puedan acceder a todas las facilidades ofrecidas por los ordenadores mediante órdenes vocales expresadas espontáneamente con el vocabulario y la sintaxis de su propia lengua y, al mismo tiempo, que los

ordenadores presenten los resultados de sus aplicaciones en ese mismo idioma de manera natural e inmediatamente comprensible para las personas. Para que este objetivo se dibuje con claridad en el horizonte tecnológico del año 2025 que vislumbramos desde este número especial de *Novática*, conmemorativo de su vigésimo quinto aniversario, será necesario continuar avanzando a un ritmo sostenido en los diversos ámbitos de investigación y desarrollo de la lingüística computacional más implicados en esta empresa.

En este artículo presentaré, en el primer apartado, una panorámica sucinta de algunas de las líneas de trabajo con mayor proyección de futuro en estas áreas, centrándome en los campos específicos de la identificación automática del idioma, del reconocimiento y síntesis del habla, de la comprensión y generación del lenguaje natural, y de la integración de la información visual en el procesamiento del lenguaje y del habla. Dedicaré el segundo apartado al procesamiento del plurilingüismo y al procesamiento documental, dos categorías específicas de aplicaciones de la

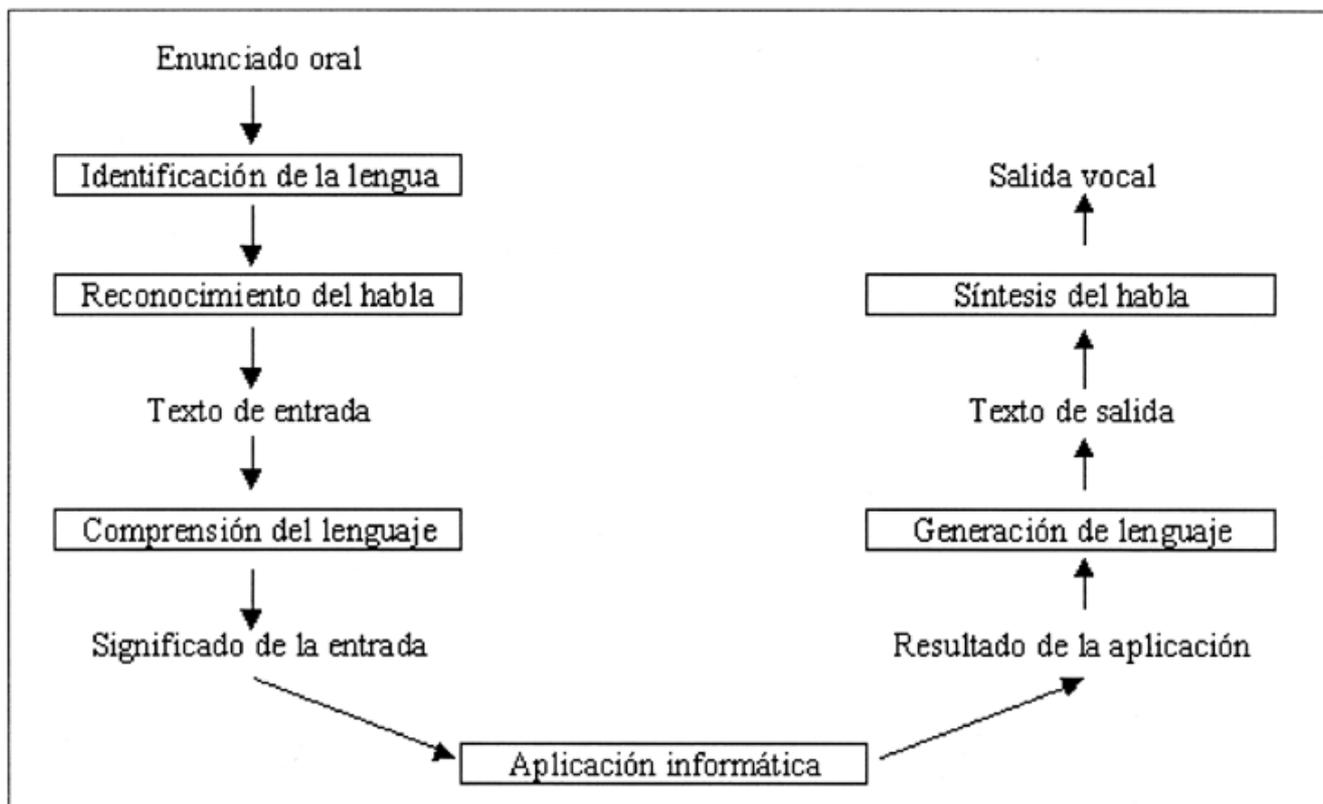


Figura 1: Interacción lingüística oral persona-ordenador

lingüística computacional de enorme incidencia social para las que se auguran importantes desarrollos.

2. Interacción lingüística oral persona-ordenador

Los sistemas de comprensión del lenguaje natural son los programas informáticos que se encargan de deducir el significado de los enunciados lingüísticos de entrada que procesan, mientras que los sistemas de generación del lenguaje natural son los responsables de presentar los resultados de las aplicaciones informáticas en forma de enunciados lingüísticos. La combinación de las técnicas de comprensión y generación de lenguaje permite el establecimiento de una interacción lingüística entre persona y ordenador en situaciones comunicativas delimitadas, como las que se dan en los programas de consulta en lenguaje natural a bases de datos o en los sistemas automáticos de diálogo por línea telefónica. Presentaré a continuación el estado actual y perspectivas de las diferentes disciplinas implicadas en estos sistemas (**figura 1**).

En el campo del **reconocimiento del habla**, implicado en la conversión automática de voz a texto, se está trabajando en tres tecnologías que prometen resultados muy interesantes a corto y medio plazo. Estas tres tecnologías son el procesamiento de la prosodia, las técnicas de resistencia a las interferencias producidas por el ruido del ambiente y la adaptación automática inmediata del sistema a la voz de la persona que realiza la locución.

Con respecto a la primera, conviene recordar que en la actualidad los programas de reconocimiento del habla (cuya encarnación más visible son los programas de dictado para procesamiento de texto en ordenadores personales) se ciñen a la identificación de las unidades fónicas segmentables, es decir, a las vocales y consonantes, descuidando el tratamiento de los elementos prosódicos del enunciado, entre los que se encuentran la curva de entonación y las pausas.

Sin embargo, lo cierto es que no es posible deducir adecuadamente el significado de un enunciado sin identificar determinados aspectos lingüísticos de gran relevancia codificados normalmente en su prosodia, por ejemplo, su modalidad (interrogativa, declarativa o imperativa), su estructura informativa (o sea, qué partes del enunciado contienen información nueva y qué partes se limitan a repetir información conocida) o su estructura sintáctica (cómo se agrupan las palabras en constituyentes sintácticos cuyas fronteras se indican con pausas perceptibles en la enunciación). En consecuencia, el desarrollo de las técnicas de procesamiento de la prosodia resulta un factor crucial en el éxito de la comprensión automática del lenguaje.

También es sumamente importante la investigación encaminada a mejorar la baja fiabilidad del reconocimiento del habla en entornos ruidosos, ya que la superación de este escollo permitirá extender la interacción oral persona-ordenador a la comunicación a través del teléfono y a otros entornos de trabajo críticos, como puede ser la interacción vocal con los ordenadores a bordo de un avión en vuelo o con un robot industrial en una fábrica con un alto nivel de ruido ambiente.

Por último, en relación con los avances de la investigación en el campo del reconocimiento del habla, se espera que la próxima generación de programas de conversión de voz a texto se adapte de forma automática a las características acústicas y articulatorias de la pronunciación de cualquier persona, haciendo innecesaria la fase de personalización del sistema mediante entrenamiento que actualmente se requiere.

Así mismo, pensando en un entorno informático plurilingüe, en el que las personas puedan interactuar con un mismo sistema en las distintas lenguas de su comunidad, habrá que consolidar también las técnicas ya existentes de **identificación automática de la lengua**, una etapa previa al reconocimiento del habla en la que el sistema decide en qué lengua se va a establecer la comunicación con la persona que lo utiliza.

Igualmente, habrá que trabajar aún mucho para conseguir que el proceso de **comprensión del lenguaje** (es decir, la deducción del significado de los enunciados de entrada) funcione con unos niveles elevados de cobertura y precisión capaces de manejar adecuadamente los enunciados que aparecen en las interacciones lingüísticas espontáneas. Un nivel de cobertura alto implicaría que el programa de comprensión no dejara casi ningún enunciado sin analizar, mientras que un grado alto de precisión supondría que a la mayoría de los enunciados analizados se les asignara un análisis correcto. Así, una cobertura apropiada evitaría que la persona usuaria del sistema tuviera que repetir un enunciado de distintas maneras por indicación del programa, mientras que una buena precisión evitaría los errores de interpretación por parte del sistema. Resulta evidente la importancia de este último factor en el caso de sistemas críticos, como los que se pueden encontrar en un avión o en una sala de operaciones.

En cuanto al proceso de **generación del lenguaje**, hasta ahora la investigación se ha centrado sobre todo en la generación de lenguaje escrito y en los entornos comunicativos de interactividad baja. Si en el futuro queremos disponer de sistemas con interacción lingüística oral espontánea persona-ordenador, será necesario orientar los esfuerzos de investigación hacia la generación de lenguaje oral en diálogos interactivos, en los que el programa de generación tenga en cuenta lo dicho en los enunciados previos y adapte su producción lingüística a las intervenciones de la persona interlocutora.

También se deben mejorar las técnicas de **síntesis del habla** para conseguir que la salida vocal del sistema parezca humana. En el estado actual de su desarrollo, la inteligibilidad de la emisión sonora, premisa básica de la voz sintetizada, es un problema prácticamente resuelto. Sin embargo, queda por solucionar la cuestión de la naturalidad de la pronunciación, es decir, conseguir que el habla generada por el ordenador no suene a voz de robot. La clave para alcanzar este objetivo podría ser la curva de entonación adoptada en la generación de los enunciados, uno de los aspectos de la síntesis vocal en los que más se está investigando.

En relación con la comunicación persona-ordenador en lenguaje natural, es de esperar que en el futuro vaya adquiriendo cada vez más importancia la **integración de la**

información visual sobre la situación comunicativa en el procesamiento del lenguaje.

En este sentido, uno de los campos de investigación más relevantes es el del procesamiento de los gestos faciales y corporales. En entornos con ruido, por ejemplo, el reconocimiento del habla puede mejorarse combinando el procesamiento acústico con los datos visuales sobre los movimientos articulatorios que se manifiestan en la cara; y, durante la fase de comprensión del lenguaje, puede resultar imprescindible que el ordenador identifique visualmente el objeto apuntado por el dedo de la persona para deducir el significado de un pronombre del enunciado. Así mismo, la generación en pantalla de un rostro sintético que verbalice adecuadamente el resultado de la síntesis del habla puede aumentar la inteligibilidad y el realismo sensorial de la salida. A más largo plazo, la interacción persona-ordenador en lenguaje natural verbal y gestual acabará integrándose como medio de comunicación entre las personas y los robots inteligentes, en una confluencia de tecnologías tan diversas como la lingüística computacional, el procesado de señal en telecomunicaciones, la microelectrónica, la visión por ordenador y la robótica.

3. Procesamiento plurilingüe y procesamiento documental

Otros objetivos ambiciosos de la lingüística computacional están relacionados con el procesamiento plurilingüe y el procesamiento documental, dos aplicaciones de gran incidencia social debido a que, por una parte, la mayoría de los seres humanos habitamos en entornos plurilingües y, por otra parte, la mayoría de las personas usuarias de la informática utilizan los ordenadores para procesar documentos.

Ya mencioné en el apartado anterior las técnicas de identificación de la lengua como un requisito para el funcionamiento plurilingüe de la interacción persona-ordenador en lenguaje natural. Sin embargo, la aplicación más popular del procesamiento plurilingüe es la **traducción automática**. Hasta ahora, los programas de traducción totalmente automática (es decir, los que no requieren intervención humana) sólo ofrecen un nivel de fiabilidad aceptable en la traducción de dominios de conocimiento muy restringidos (por ejemplo, en la traducción de lenguajes sectoriales, como el de los partes meteorológicos) o bien cuando el texto de partida está escrito siguiendo unas normas muy estrictas orientadas a la simplificación de su léxico y sintaxis (como es el caso de los lenguajes controlados). Muchas personas expertas en este campo consideran que, posiblemente, la situación actual muestra el límite de las posibilidades de la traducción automática, aunque tal vez lo que exista sea un cierto desaliento por la enorme complejidad de la tarea. Los futuros desarrollos en este ámbito apuntan a un cambio de paradigma en el modelo perseguido de traducción informatizada. El objetivo ideal de la automatización completa del proceso será substituido por un modelo de traducción interactivo asistido por el ordenador en el que la participación humana será crucial en el proceso. Los nuevos programas de traducción consultarán sus dudas sobre el

texto con una persona experta para obtener así una representación semántica correcta del original, que podrá ser así traducida automáticamente sin error a diversas lenguas.

Por otra parte, una de las líneas de trabajo con más futuro de la lingüística computacional es el procesamiento documental y, en particular, sus aplicaciones de recuperación de la información textual, de resumen automático de documentos y de escritura asistida por ordenador. Las aplicaciones actuales de **recuperación de la información** textual, basadas en su mayoría en técnicas estadísticas, se han extendido y popularizado debido a su presencia como motores de búsqueda en Internet. En el futuro, las aplicaciones de recuperación de información incorporarán un cierto procesamiento lingüístico, tanto de las consultas como de los documentos, manejarán con soltura documentos redactados en distintos idiomas, y permitirán realizar las consultas en la lengua materna de la persona usuaria al margen de las lenguas en que se hallen escritos los documentos. Así mismo, se espera que el desarrollo de las técnicas de **resumen automático** permitan acceder a los documentos recuperados mediante resúmenes de estos documentos generados sobre la marcha, evitando así la lectura completa de los documentos que no sean pertinentes.

En cuanto a las aplicaciones de la lingüística computacional a la **escritura asistida por ordenador**, en los próximos años se espera un avance claro en dos direcciones. Por una parte, los programas de corrección de la sintaxis y el estilo incrementarán su eficacia, incorporando técnicas de análisis lingüístico más refinadas (probablemente, similares a las empleadas en el campo de la comprensión del lenguaje natural) que superarán los resultados proporcionados por las técnicas de análisis superficial y de reconocimiento de patrones empleadas por la mayoría de estos programas en la actualidad. Por otra parte, las herramientas de creación de documentos estructurados se enriquecerán con modelos discursivos mucho más elaborados, surgidos de los campos de la generación del lenguaje natural, del análisis y etiquetado descriptivo de corpus, y del estudio lingüístico cuantitativo de las tipologías textuales, lo que permitirá que los programas de procesamiento de textos ofrezcan una guía mucho más detallada del formato y los contenidos del documento elaborado.

4. Conclusiones

La lingüística computacional pretende incorporar en los ordenadores diversas habilidades lingüísticas, como las de hablar, redactar, resumir o traducir. Aunque la historia de la disciplina es corta (unos 50 años), su desarrollo científico teórico y aplicado ha sido muy importante y su incidencia social, muy elevada. Aun siendo muy probable que nunca se llegue a incorporar en los ordenadores el dominio del lenguaje y la inteligencia que Kubrick vaticinó para HAL en 1968, la vitalidad de la investigación en lingüística computacional no deja lugar a dudas de su imparable avance.