

Ingeniería del Lenguaje Natural

Fundamentos, logros y retos

José F. Quesada
Director I+D



1. ¿CÓMO FUNCIONAN LOS SISTEMAS O APLICACIONES DE LENGUAJE NATURAL?

Personalmente, prefiero utilizar el término “Ingeniería del Lenguaje Natural” para resaltar el hecho de que se trata de un área de Investigación, Desarrollo e Innovación dirigido a la creación de aplicaciones reales en las que el lenguaje natural constituye el núcleo del sistema.

La introducción del calificativo “natural” para describir un lenguaje es relativamente reciente, ya que todo lenguaje era en sí mismo natural hasta la aparición de los lenguajes formales de la lógica-matemática y los lenguajes de especificación y programación propios de las ciencias de la computación. Actualmente calificamos como “lenguajes naturales” aquellos lenguajes utilizados comúnmente para la comunicación oral y/o escrita entre personas. Por tanto, los sistemas o aplicaciones de la ingeniería del lenguaje natural se podrían caracterizar por el objetivo de lograr la automatización en el uso de dichos lenguajes por parte de los sistemas informáticos.

El funcionamiento de estos sistemas no se puede describir atendiendo a criterios uniformes. Durante las últimas décadas han aparecido distintos enfoques o estrategias para lograr estos objetivos. Incluso, teniendo en cuenta la complejidad del problema se han abordado problemas parciales lo que ha supuesto la aparición de sub-áreas tales como la recuperación de información, la traducción automática o los sistemas de diálogo. Incluso, en cualquiera de estas áreas se pueden diferenciar estrategias estadísticas o basadas en conocimiento, etc.

Desde el punto de vista de la arquitectura de sistemas basados en lenguaje natural, Natural Language constituye una nueva generación tecnológica. La tecnología responde a varios retos planteados de forma simultánea: en primer lugar, obtener sistemas que logren una comprensión del lenguaje natural, en segundo lugar, que la información semántica abstracta obtenida durante la comprensión se integre en el contexto de un sistema de razonamiento capaz de integrarse con fuentes externas de información (bases de datos, sistemas de información, etc.), en tercer lugar, crear un nuevo lenguaje de programación que facilite el diseño de este tipo de sistemas, y en cuarto lugar, conseguir que el sistema se ejecute en tiempo real.



2. ¿QUÉ TECNOLOGÍAS SUBYACEN DEBAJO DE ELLAS?

Lograr un sistema funcionalmente completo desde el punto de vista del lenguaje natural requiere la colaboración de múltiples módulos o componentes.

Entre los principales componentes que integran la tecnología Natural Language podemos destacar:

- + Reconocimiento de voz.
- + Descubrimiento de lenguajes,
- + Análisis léxico-morfológico,
- + Análisis gramatical o sintáctico,
- + Análisis semántico,
- + Estrategias de merging, raising, thought, deviling y phrasing
- + Modelos de representación de la información basados en estructuras Semic,
- + Organización de la memoria de trabajo (áreas de Wernicke, Broca, Hipo-Memory, Assoc-Memory y Focus-Memory),
- + Generación semántica, sintáctica y léxica,
- + Síntesis de voz, ...

No obstante, quizás el gran reto y también la gran aportación de Natural Language ha sido conseguir un entorno donde todas estas tecnologías intervienen en una arquitectura totalmente integrada.

3. LAS VENTAJAS SON EVIDENTES. ¿CUÁL DESTACARÍA POR ENCIMA DE TODAS? ¿POR QUÉ?

De entre las múltiples ventajas derivadas de esta tecnología, me gustaría destacar dos, atendiendo a diferentes aproximaciones a la tecnología.

En primer lugar, desde el punto de vista de los usuarios, la situación actual se puede caracterizar por una gran debilidad en los interfaces de acceso. Es decir, la tecnología informática es bastante reciente, y se ha visto sometida a un espectacular proceso de crecimiento y extensión, y la mayor parte de las actividades productivas actuales usan de una u otra forma sistemas de computación. Ahora bien, los avances en los propios interfaces de acceso en mejoras desde el punto de vista gráfico, mejorando sensiblemente el componente visual de la interacción, pero se han ignorado la parte verbal de la comunicación (reducida habitualmente a menús y botones). El lenguaje natural, tanto escrito como hablado aportará una mayor agilidad y simplicidad desde el punto de vista de los interfaces de usuario.



En segundo lugar, desde un punto de vista económico o empresarial, esta tecnología pone a disposición de las empresas nuevas posibilidades tanto para mecanismos personalizados de interacción con el cliente o consumidor, como estrategias para la reducción de costes. Es decir, esta tecnología permite abordar funciones que hasta ahora no se podían abordar por la propia dificultad. Podemos por ejemplo conseguir que durante un diálogo con una máquina una persona solicite una beca de estudios interactuando con un sistema basado en lenguaje natural, y el resultado puede ser la misma petición de la beca. O bien, que solicite un préstamo y el sistema le vaya preguntando la información pertinente. Por otro lado, procesos que actualmente están basados en tareas repetitivas que involucran grandes cantidades de personal pueden automatizarse. En este caso, podemos evitar los colapsos que se producen cuando el número de llamadas superan el máximo permitido, podemos dedicar el personal a resolver tareas más productivas para la institución, y podemos contar con sistemas en alta disponibilidad.

4. COMO NADA ES PERFECTO, ¿CUÁLES SON LOS PRINCIPALES DEFECTOS QUE A DÍA DE HOY SIGUE TENIENDO ESTA TECNOLOGÍA?

En primer lugar, hay un grave defecto no imputable a la tecnología en sí, sino quizás al uso (o mejor “abuso”) que de ella se ha hecho durante bastante tiempo. Al igual que otras áreas de la Inteligencia Artificial, los sistemas basados en lenguaje natural han despertado en varias ocasiones bastante interés en los sectores empresariales. No obstante, la inmadurez de la tecnología en sus fases iniciales junto con la irresponsabilidad de algunas empresas al calificar como “lenguaje natural” una tecnología que realmente no lo era, ha provocado un enorme desencanto y rechazo a esta tecnología. Esto supone un reto ya que no sólo hay que demostrar la funcionalidad de la tecnología sino superar los recelos que persisten hacia ella.

A otro nivel, más que de un defecto podríamos hablar de una limitación, derivada de la propia naturaleza del lenguaje natural. Lo ilustraré con un ejemplo. Si hemos de implementar un algoritmo para hacer una operación matemática, partimos de una especificación digamos “estable” y nuestro objetivo se podrá dirigir hacia la eficiencia, generalidad, etc. No obstante, en lenguaje natural no se puede hablar de especificaciones “estables”. El lenguaje natural es algo vivo, sujeto a los cambios que introducen constantemente los hablantes, y esta riqueza dificulta sensiblemente la creación de aplicaciones para un dominio de conocimiento amplio.



5. ¿Y LOS PRINCIPALES RETOS A LOS QUE SE ENFRENTA?

La tecnología de Natural Language ha sido el fruto de un intenso esfuerzo en investigación y desarrollo para lograr una nueva aproximación que superase varios retos clave como son la representación de información semántica, la manipulación eficiente de dicha información, la conexión entre el mundo de la comprensión del lenguaje natural y el razonamiento automático, la obtención de una plataforma de ejecución en tiempo real, multilingüismo y multicanalidad y multimodalidad, etc.

Todos estos retos se han superado en un tiempo record y de una forma brillante. No obstante, el espíritu de Investigación, Desarrollo e Innovación de Natural Language nos lleva a plantearnos nuevos retos, entre los que podemos destacar:

- + Mejora en el reconocimiento de voz: Nuestro enfoque actual para el reconocimiento de voz difiere en cierta medida del seguido por otros grandes fabricantes de tecnología de reconocimiento. Básicamente, nuestra estrategia a este nivel se basa en considerar que el reconocimiento de voz es una parte de la comprensión del lenguaje. Es decir, el reconocimiento de voz no termina con el análisis de la señal acústica y la incorporación de modelos estadísticos y/o gramaticales de lenguajes, sino que el cerebro humano reconoce acústicamente como parte de la comprensión. Siguiendo este esquema, la tecnología actual que hemos obtenido permite mejorar sensiblemente la calidad de los sistemas de diálogo basados en voz. No obstante, esta área permite aún mejoras que estamos abordando en este momento.
- + Formación e integración: Para Natural Language constituye un reto actual el conseguir una base suficientemente amplia de integradores bien formados en esta tecnología. Esto nos ha llevado a trabajar en el área de formación y certificación para conseguir que todos los integradores que ya nos lo están demandando puedan comenzar a realizar desarrollos usando nuestra tecnología.

6. ¿QUÉ INFRAESTRUCTURA TECNOLÓGICA NECESITAMOS PARA INCLUIR ESTAS HERRAMIENTAS?

Una de las principales características que diferencian la tecnología de Natural Language de otras empresas con tecnología más o menos relacionada con el lenguaje natural es la eficiencia.

Como Director de I+D de Natural Language y responsable del desarrollo del núcleo de la tecnología que yo mismo implementé durante aproximadamente dos años, impuse durante el diseño de la arquitectura el obtener los mínimos requerimientos a nivel de infraestructura tecnológica.



Como resultado, el núcleo del sistema permite actualmente ser ejecutado sobre una infraestructura totalmente estándar. A nivel de ejemplo, podemos indicar que un sistema de diálogo para la gestión de la cita previa sanitaria permite ejecutar hasta 15.000 !!! conversaciones simultáneas (concurrentes) sobre un servidor Intel Xeon 3.2Gz con 4 GB RAM, sin requerimientos adicionales a nivel de hardware o software.

7. ¿PUEDE PONERNOS UN EJEMPLO DE UNA INSTALACIÓN TÍPICA? (NÚMERO DE SERVIDORES, SISTEMA OPERATIVO, COSTE ...)

Es prácticamente imposible describir una instalación típica debido a la gran versatilidad que permite la tecnología. No obstante, intentaré describir los principales perfiles en cuanto a aplicación e instalación se refiere.

Actualmente contamos con tres perfiles básicos de aplicaciones:

- + Natural Language Assistant: Se trata de una aplicación o solución dirigida a permitir interacciones basadas en lenguaje natural contra un portal web. El objetivo es lograr una interacción natural con el usuario que desea consultar la información del portal web, hacer preguntas relevantes al área de conocimiento correspondiente, o incluso dialogar con el sistema. Se distinguen tres niveles funcionales:
 - o Topic Level,
 - o Knowledge Level,
 - o Dialogue Level.
- + Natural Language Self-Service: Se trata de sistemas dirigidos a la interacción mediante lenguaje natural con sistemas de información bien a través de la voz (usualmente mediante llamadas telefónicas) o bien mediante texto (en este caso desde un entorno web).
- + Natural Language Transcripator: En este caso, se trata de una solución dirigida a la transcripción, lo más literal posible, de intervenciones habladas, por ejemplo, en sesiones parlamentarias, etc.

Por otro lado, respecto al número de servidores evidentemente dependerá del nivel funcional requerido. Es decir, la instalación más simple puede tener suficiente con un único servidor. Si necesitamos configuraciones en alta disponibilidad evidentemente será necesario incrementar la instalación con configuraciones basadas en componentes stand-by. Por otro lado, la incorporación de módulos de reconocimiento, síntesis, control de la telefonía, implementación de servicios web, etc., puede requerir el uso de servidores adicionales, pero en cualquier caso, toda la arquitectura está orientada a la máxima optimización en el uso de recursos.



Respecto al sistema operativo, nuestro entorno básico de desarrollo y explotación es Linux, aunque al estar implementado el núcleo del sistema en ANSI C nos permite llevar a cabo el plataformado del sistema en distintas arquitecturas y sistemas. Finalmente, respecto al coste, existe también una amplia gama de precios en función de los requisitos funcionales del sistema.

8. ¿ES LA WEB SU PRINCIPAL DESTINATARIO? ¿POR QUÉ?

Actualmente nuestro plan tecnológico está focalizado en dos escenarios prioritarios: la web, mediante el diseño de asistentes virtuales basados en lenguaje natural incluyendo altas prestaciones, y en segundo lugar la automatización de centros de contactación.

Respecto a los asistentes virtuales en lenguaje natural para la web, nuestras principales ventajas competitivas se centran en: Capacidad de comprensión del significado de la frase o consulta realizada por el usuario, establecimiento de un nivel de diálogo, generación de respuestas en lenguaje natural, integración y conexión con sistemas de información, eficiencia, multilingüismo, multicanalidad y disponibilidad de un cuadro de mando de altas prestaciones. Las estrategias de búsqueda de información mediante lo que podemos denominar genéricamente “algoritmos de palabras clave” han constituido una primera y muy brillante solución, como demuestra el caso más famoso de Google. No obstante, en la propia naturaleza de estos sistemas reside su propia limitación. Como alternativa tecnológica aparece la idea de Asistente Virtual. De entrada, para conseguir una optimización de la relación se parte de una reducción del dominio. El asistente virtual se convierte de esta forma en una herramienta especializada en un dominio de conocimiento. Pero de nuevo aparecen varias alternativas. Desde un enfoque trivial, que básicamente reproduce la estrategia de búsqueda de palabras clave en la expresión, aunque con cierta mejora del nivel de acierto, hasta la tecnología Natural Language Assistant, un verdadero motor de comprensión de lenguaje natural que permite interacciones inteligentes.

Respecto a la automatización de los centros de contactación, nuestra tecnología supone un gran avance respecto a la generación actual de sistemas basados en reconocimiento mediante “key-word spotting” obteniendo los denominados sistemas IVR. La gran deficiencia de estos sistemas se desprende de la debilidad de la concepción que tienen del reconocimiento de voz, lo que les supone tener que abordar la relación con el usuario como una interacción estructurada en un árbol de posibilidades que dificulta enormemente la conversación. La tecnología de Natural Language supone una nueva generación de tecnología a este nivel, ya que permite una “comprensión del lenguaje hablado”. Por tanto, este es un sector prioritario para Natural Language por cuanto podemos abordar proyectos de automatización de centros de contactación mejorando sensiblemente el modelo de interacción y reduciendo considerablemente los costes de explotación del centro.



9. ¿EN QUÉ SITUACIÓN SE ENCUENTRA EN ESTOS MOMENTOS EL MERCADO ESPAÑOL EN LO QUE A ADOPCIÓN DE ESTAS TECNOLOGÍAS SE REFIERE?

Estas tecnologías constituyen actualmente un mercado emergente. Por otro lado, y a pesar de la presencia reciente de la tecnología de Natural Language en el mercado, podemos resaltar la excepcional acogida que está teniendo.

Desde mi punto de vista, el mercado español actualmente cuenta con un cierto número de sistemas implementados sobre tecnología IVR para dar respuesta a los centros de contactación, y que a nivel de asistentes web son realmente mínimos los organismos que han decidido llevar a cabo su implantación, y de entre los pocos que hay, la mayoría presentan graves deficiencias desde el punto de vista de la capacidad de interacción en lenguaje natural y sobre todo usando diálogo.

Podemos considerar que actualmente la tecnología de Natural Language es la única capaz de asumir simultáneamente la capacidad para interactuar en lenguaje natural, incorporando en el mismo núcleo tanto los modelos de lenguaje hablado como escrito, permitiendo multilingüismo directamente, y siendo capaz de generar diálogos reales, con capacidad de memoria y proyección de expectativas.

10. ¿Y LAS PROYECCIONES PARA EL FUTURO?

Mi sensación es que el mercado tiene una “necesidad” casi imperiosa para incorporar una verdadera tecnología de lenguaje natural.

Los portales web de los grandes organismos (banca, administración, seguros, etc.) están simplemente colapsados por el propio volumen de información disponible, y los usuarios terminan perdiéndose en un mar ingente y habitualmente desorganizado de información. Los centros de atención telefónica se distribuyen entre el extremo de los sobrecostes asociados a una atención totalmente controlada por operadores o el uso de IVR con capacidades mínimas que frustran las expectativas de los usuarios.

Teniendo en cuenta por un lado que la presentación de la tecnología de Natural Language al mercado es realmente muy reciente, y por otro lado, atendiendo a la excelente acogida que han tenido en este corto período de tiempo, nuestras proyecciones para el futuro son muy optimistas.

Nuestros objetivos se centran en la consolidación inmediata de las tres grandes líneas en que hemos estructurado nuestro plan tecnológico, la implantación inmediata de un plan de



formación y certificación de integradores, entornos multilingües para los lenguajes oficiales dentro de España (castellano, catalán, vasco y gallego), así como el acceso durante el mismo año 2007 al mercado internacional comenzando con los lenguajes inglés y francés.

11. ¿POR QUÉ LAS EMPRESAS DEBERÍAN TENER ESTAS HERRAMIENTAS?

Hay varias razones que justifican la adopción de estas tecnologías por parte de las empresas, así como de las administraciones.

- + Acceso a los clientes, ciudadanos o usuarios. Independientemente de los distintos perfiles, los sistemas de lenguaje natural permiten canalizar de una forma coherente las interacciones con los clientes en el caso de una empresa o ciudadanos en el caso de una administración. Estos sistemas permiten mejorar la imagen del propio organismo, permiten agilizar el acceso de los usuarios a la información, y permiten una integración ágil con otras herramientas de la institución tales como entornos CRM, sistemas de información, etc.
- + Reducción de costes, y optimización en el uso de los recursos humanos. Un porcentaje muy alto de las atenciones realizadas por los centros de contactación son actuaciones poco o nada productivas para la institución, pero que no obstante suponen un considerable coste para la misma. La tecnología de Natural Language permite llevar a cabo una reducción considerable en los costes de explotación del centro de contactación.





www.naturalanguage.es