

ACERCA DE LA NATURALEZA DE LA INFERENCIA COMUNICATIVA*

Eduardo de BUSTOS**

Introducción

Uno de los fines explícitos de la Inteligencia artificial (IA) es la reproducción de los procesos cognitivos humanos, desde la resolución de problemas a la capacidad para comunicar información. En particular, en lo que atañe a la emulación de la conducta comunicativa humana, la IA tiene uno de sus mayores desafíos en, la elaboración de diálogos interactivos que permitan a los humanos un acceso efectivo y simple a sistemas computatorios a través del uso del lenguaje natural¹.

Un principio heurístico que permite abordar este problema, que se fundamenta en lo que se conoce como **metáfora computatoria**, es el de tratar de comprender la conducta comunicativa humana en términos computatorios, en particular los procesos de producción y comprensión lingüística. Desde este punto de vista, la producción y la comprensión lingüísticas han de ser concebidas como el producto final de **cálculos** que realiza nuestro sistema cognitivo sobre una información tanto externa como interna, esto es, adquirida mediante nuestros sistemas perceptuales o recuperada de la memoria.

En este trabajo, me voy a centrar sobre el segundo aspecto de los procesos comunicativos mencionados, el de la comprensión lingüística, considerándolo de alguna forma previo o prioritario al de la producción, en el sentido de que la dirección causal en la comunicación efectiva entre humanos va de la comprensión a la producción.

Trataré de exponer la naturaleza de los siguientes problemas en un nivel puramente conceptual, si se quiere filosófico, y traducirlos luego a términos más familiares en inteligencia artificial:

1. en primer lugar, el problema de determinar qué es lo que se computa en la comprensión lingüística. Esto es, trataré de caracterizar una noción que refiera a la información que desempeña efectivamente un papel causal en las interacciones comunicativas entre humanos y, por extensión, entre el hombre y el ordenador.

2. en segundo lugar, trataré de caracterizar el proceso, o los procesos, que permiten efectuar ese cálculo. En particular, expondré ciertas limitaciones formales de carácter muy general acerca de las propiedades que han de tener los mecanismos inferenciales implicados en la computación del significado.

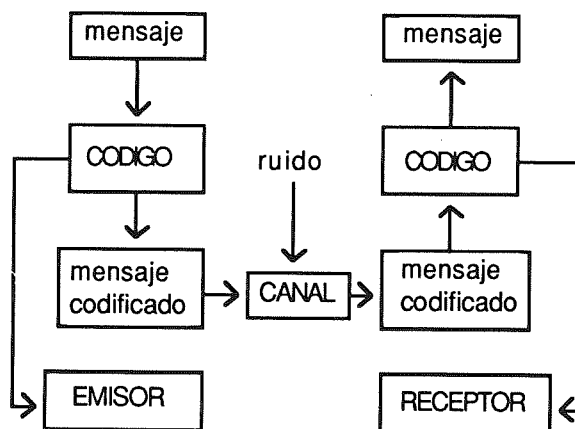
3. en tercer lugar, señalaré algunos de los problemas más importantes en los sistemas de AI que tratan de emular la comprensión lingüística y las orientaciones más prometedoras para su resolución.

* Estas observaciones recogen parcialmente mis intervenciones en **La complejidad, perspectivas interdisciplinares**, UIMP, La Coruña, 1990, y en **La inteligencia artificial hoy**, Universidad de Santiago de Compostela, 1991.

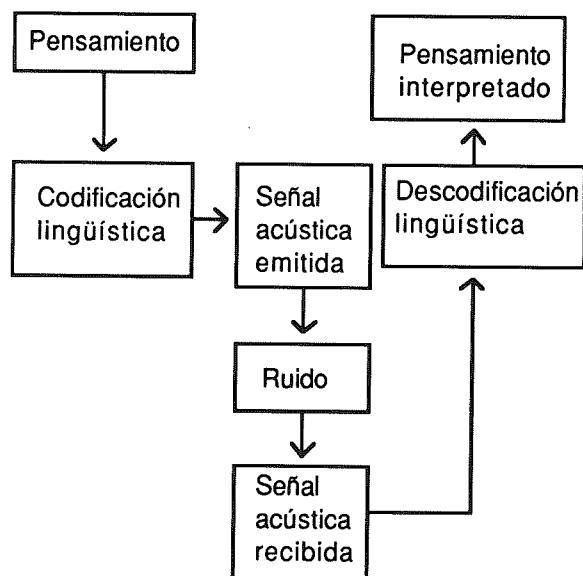
Significado oracional y significado preferencial

La teoría de la comunicación lingüística, al menos desde el siglo XVIII, se ha venido nutriendo de la analogía entre los siguientes esquemas:

I. Diagrama clásico del proceso de comunicación



II. Conducta lingüística verbal



Nótense las similitudes y las diferencias porque **ambas** son importantes para la comprensión de la extensión metafórica de la noción de código a la comunicación. Lo que en el modelo original es en realidad un proceso **cifrado**, de reescritura en una notación diferente, en este modelo es un proceso de **traducción** de pensamiento a expresiones lingüísticas. Incluso si dejamos por un momento la cuestión de si los pensamientos tienen ya una forma lingüística (si forman parte de un **lenguaje interior**, como sostenía Agustín de Hipona y mantiene el psicólogo J. Fodor), es evidente que la codificación lingüística no es **literalmente** comparable al del modelo original. Del mismo modo sucede con el proceso de descodificación, en el que no se produce una simple retranscripción, sino una auténtica traducción, si suponemos que el destino final del pensamiento es el lenguaje mental del destinatario. Entre otras cosas importantes, un elemento esencial que distingue ambos modelos es la existencia de una **gramática**: en el supuesto proceso de codificación lingüística interviene ese conjunto de reglas que nos permite producir una entidad que, de algún modo, contiene la información que contiene la **representación mental** que queremos transmitir al destinatario. La existencia de ese conjunto de reglas, entre otras cosas, separa al lenguaje natural de un código. Para expresarlo de una forma contundente:

una lengua no es un código, y tampoco un código es una lengua. Sólo las lenguas en sentido corriente -lenguas como el inglés, el francés o el latín- tienen gramáticas. Un código no tiene una gramática, ni podría, de alguna forma generar una gramática².

De tal modo que no existe, ni puede existir una gramática Morse o Braille, aunque efectivamente existen **tablas de equivalencias** entre los signos del código y los signos lingüísticos (las letras): lo característico del código es que puedo utilizarlo sin tener ni idea de la lengua en que están escritos los caracteres que cifro. Ello no afectaría a la transmisión, que podría ser descifrada por un receptor igualmente ignorante³.

Conviene distinguir pues entre procesos distintos, relacionados con la comunicación, que no obstante pueden no tener nada que ver entre sí:

1) en primer lugar, y considerando lo más elemental, el proceso de transmisión de una señal. Este proceso no implica necesariamente una codificación: por ejemplo, la **transmisión** telefónica consiste en la transformación de energía dinámica -las ondas acústicas- en impulsos eléctricos, los impulsos eléctricos recorren un canal -el hilo telefónico- y son retransformados a su vez en ondas acústicas que impresionan nuestros tímpanos. En ningún momento del proceso es correcto hablar **literalmente** de codificación. Lo único que se ha producido es la transformación de una forma de energía en otra de acuerdo con ciertas leyes físicas.

2) en segundo lugar, hay que distinguir los procesos de **codificación** de mensajes, en los que tales mensajes son **cifrados**, esto es, reescritos de acuerdo con un procedimiento, la clave del código, posiblemente transmitidos en esa forma

cifrada, y finalmente descifrados por un receptor, mediante la aplicación de la **clave**. Como he dicho, tales procesos de codificación y decodificación son independientes de que emisor y receptor tengan conocimiento del **significado** (sea esto lo que sea) de los mensajes transmitidos, no requieren ningún conocimiento lingüístico en particular y la competencia codificadora consiste esencialmente en una habilidad mecánica en el sentido que lo es la del perforador de tarjetas de ordenador, por ejemplo.

3) en tercer lugar, los procesos de **traducción**, que implican el conocimiento lingüístico, la **comprensión** de los mensajes que se trasladan, y que persiguen la conservación de la información semántica.

Uno de los inconvenientes de la metáfora semiótica es que hace suponer que lo **comunicado** se encuentra de alguna forma contenido, encerrado, en la representación semántica de la expresión fonéticamente realizada. Sin embargo, si lo que se comunica en la comunicación tiene algo que ver con la información a que se da acceso al destinatario de la comunicación, con la información que éste adquiere en virtud de una relación **causal** con la expresión lingüística utilizada por el emisor, entonces tal supuesto es radicalmente falso. La representación semántica de una oración⁴ no contiene toda la información que es transmitida por medios lingüísticos en la comunicación. A veces ni siquiera es **esa** información la que se transmite, a pesar de ser irreprochable la comunicación. Por decirlo de otro modo, el esquema II omite precisamente lo que es la médula de la comunicación, el hecho de que la información comunicada no está **dentro** del mensaje, sino que, por decirlo así, el mensaje sólo es la **llave** que da acceso a esa información.

Una de las razones de la carencia del esquema II es que no repara en la diferencia existente entre una **oración** y una **preferencia**⁵ de una oración. Las oraciones son entidades teóricas, abstractas de la teoría lingüística, entidades cuya representación semántica (o algo que se le asemeje) y fonética ha de reproducir al gramática propuesta por la teoría. En contraste, las preferencias son acciones concretas realizadas por los hablantes de una lengua cuando la utilizan. Como tales acciones, están localizadas espacio-temporalmente y son literalmente irrepetibles. La diferencia entre uno y otro tipo de entidades lingüísticas se suele expresar diciendo que las oraciones son expresiones-tipo y las preferencias ejemplares o muestras de expresiones⁶.

Pues bien, mientras que una expresión oracional tipo y un ejemplar de la misma comparten las propiedades estructurales que busca describir y explicar la gramática (su configuración interna, su forma lógica, su clase distributiva, etc.), no tienen por qué compartir necesariamente su significado. Habitualmente el significado de una expresión tipo se puede identificar con la representación semántica que a esa expresión permite asignarle la gramática, si es que permite asignarle alguna. En cambio, el significado de una preferencia concreta de esa oración puede no coincidir con la información contenida en la representación

semántica de ésta. Para darse cuenta de ello, basta considerar cualquier expresión que remita a características extralingüísticas (en realidad, en un sentido u otro, **todas** las expresiones remiten a ese tipo de características):

(1) Hoy es lunes

En ella, la expresión *hoy*, considerada en cuanto expresión-tipo, esto es, abstracta entidad lingüística, tiene una representación semántica que contiene la información de que refiere al día en que se profiera la expresión⁷. Pero en cuanto ejemplar i-ésimo de esa expresión tipo puede contener la información de que refiere al 1 de Julio de 1991. Quien profiera (1) puede transmitir la información de que ese día en particular es lunes, lo cual **forma parte** del significado de la preferencia de (1), aunque no de la expresión-tipo.

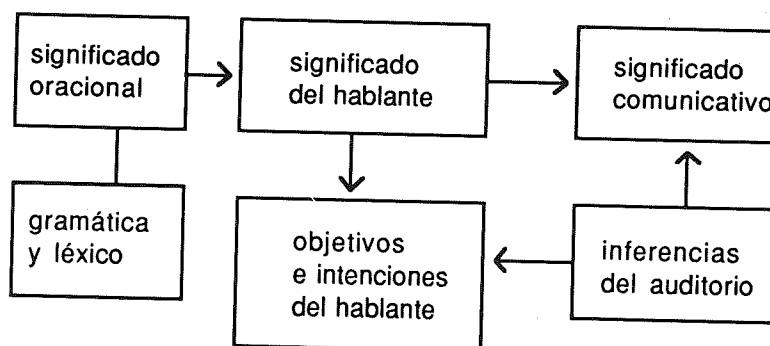
Una distinción estrechamente relacionada con ésta, y que se ha utilizado en lingüística y en filosofía del lenguaje, como criterio para distinguir entre los objetos propios de la semántica y de la pragmática, es la que separa al **significado oracional del significado del hablante**. Veamos cómo explica esta dicotomía un reciente manual de semántica (en la variedad de Massachussetts, MIT)⁸:

Supóngase que Pepita quiere decirte que Juan se ha ido, pero es una hablante del español no muy competente y no ha captado bien los significados de **llegar e irse**. Dice entonces: **Juan ha llegado**, pero lo que realmente quiere decir es que se ha ido. Sin embargo, podríamos desear decir que la oración que profiere, a pesar de sus deseos, no significa que Juan se ha ido, sino que realmente significa que Juan ha llegado. Lo que aquí tenemos es una especie de ambigüedad: en un cierto sentido, lo que Pepita dice es que Juan se ha ido, en otro, que Juan ha llegado /.../ Distingo estas dos clases de significado llamando al primero **significado del hablante** y al segundo **significado oracional**.

El significado oracional es pues equivalente a lo que hemos venido llamando significado de la expresión oracional tipo -que se aplica tanto a las oraciones como a otras clases de expresiones- y puede no coincidir con **lo que el hablante pretende decir**. Ahora bien, ¿cuál es el significado de la **preferencia** de Pepita: **Juan ha llegado**? Por una parte, parece evidente que no es que Juan ha llegado, puesto que Pepita parece querer decir justamente todo lo contrario y es posible que el destinatario capte esa intención, sabiendo que Pepita tiende a trabucar los términos antónimos en español. El oyente entendería entonces que Pepita quiere decir que Juan ha llegado **y así es como interpretaría su preferencia**. Su interpretación de la información transmitida mediante la preferencia de Pepita coincidiría en este caso con el significado que Pepita pretende dar a su uso de la oración, coincidiría con su significado del hablante. Pero, por otro lado, supóngase que el auditorio no tiene conocimiento de la incompetencia lingüística de Pepita: entonces puede interpretar que Pepita quiere decir lo que realmente dice, esto es, que Juan ha llegado. En este caso, la información contenida en la representación semántica de la oración, esto es, su **significado oracional** y, desde el punto de

vista de la descripción de la comunicación entre hablante y auditorio, **ese** es el significado de la preferencia. Por el momento, llamaremos a este significado el **significado comunicativo** de la preferencia, en contraste tanto con el significado oracional como con el significado del hablante.

Resumiendo esquemáticamente, tenemos tres nociones, relacionadas con tres aspectos de la comunicación lingüística:



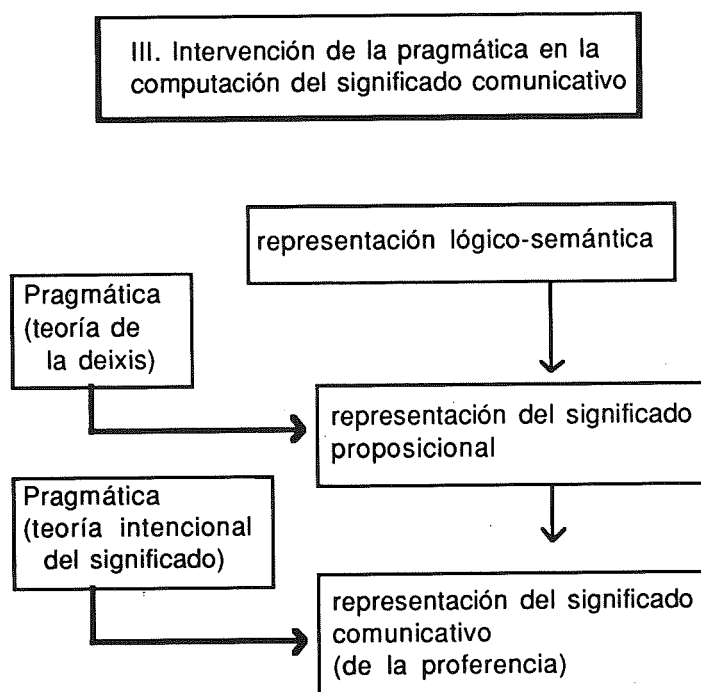
La función de la pragmática en el análisis del significado y el modelo inferencial de la comunicación lingüística.

Resulta pues que, a veces, el significado comunicativo de una preferencia equivale a su significado oracional y a veces resulta que no, que coincide con el significado del hablante, o con otras cosas. Aunque en ocasiones se ha tratado de confinar el análisis pragmático a los casos en que el significado de la preferencia no coincide con el significado oracional, en realidad su ámbito va más allá del que estos casos determinan. Para empezar, ya hemos visto que, incluso en el nivel del significado oracional o semántico, es necesaria la consideración pragmática para dar cuenta de los aspectos deícticos o indéxicos de la oración (los que remiten a características extralingüísticas). Por otro lado, la pragmática es una disciplina cuyos análisis cubren en realidad **ambos** casos:

1) cuando lo que se comunica es la información contenida en la representación semántica -determinada con ayuda de una teoría pragmática de la deixis, y,

2) cuando se comunica algo más, o algo diferente, como el significado del hablante.

Resumiendo de forma esquemática, la teoría pragmática interviene en los siguientes momentos de la computación del significado comunicativo:



El resultado es pues que la teoría pragmática, en su acepción general, es la teoría que determina el significado comunicativo de la preferencia, utilizando como entrada información recogida en la representación semántica de las expresiones-tipo, pero incorporando en el proceso ingredientes no lingüísticos, esenciales no obstante para la determinación de la salida del proceso, el significado comunicativo.

Ahora bien, en la noción de **preferencia** subsiste una indeterminación. Desde el punto de vista de la descripción de la comunicación, las preferencias pueden ser consideradas como acciones o productos de acciones. En otro lugar⁹, hemos propuesto deshacer esa indeterminación denominando a los resultados de las preferencias **prolata**, inscripciones verbales o escritas de entidades lingüísticas, y reservando el término 'preferencia' para el acontecimiento mismo. Las preferencias son pues acciones, acciones verbales, por más señas. Y la comprensión de las acciones verbales, la captación de su significado, sigue un proceso muy diferente al de la descodificación de un mensaje. Por ello, es necesario abandonar el marco teórico de la concepción semiótica de la lengua y utilizar nuevos instrumentos para comprender y reproducir la forma en que algunas de nuestras acciones son comunicativas, esto es, constituyen instrumentos adecuados para transmitir información a nuestros semejantes.

El modelo inferencial de la comunicación lingüística aspira precisamente a la sustitución del modelo semiótico. Pretende dar cuenta de la forma en que los

pertenecientes a una comunidad epistémicamente pertinente- que comparte los conocimientos relevantes para la coordinación de sus acciones, por ejemplo la competencia lingüística, asignan significado a sus acciones verbales, en cuanto agentes, y son capaces de interpretar esas acciones, en cuanto destinatarios o receptores de dichas acciones.

Uno de los supuestos fundamentales en que se basa el modelo inferencial es de carácter negativo o crítico: los procesos de codificación y descodificación no desempeñan ningún papel significativo en la descripción y explicación de la comunidad lingüística. Según D. Sperber y D. Wilson, autores de una de las versiones más completas y conocidas de este modelo, "los humanos no se comunican codificando y descodificando pensamientos"¹⁰. Dicho de otro modo (en términos vagamente cognitivos), los procesos psicológicos que se desarrollan en la mente de los que participan en un intercambio comunicativo, cuando producen o interpretan acciones, no consisten en la aplicación de un código mental subyacente que permita la extensión y comprensión de lo que las acciones significan. La operación básica es otra y el proceso es mucho más sutil y complejo. Lo fundamental es la operación de **inferencia**, que consiste esencialmente en la producción o captación de una información a partir de un conjunto de informaciones antecedentes. Desde el punto de vista cognitivo -no lógico ni lingüístico-, la inferencia es, por una parte, una operación consistente en la manipulación de **representaciones mentales**, que J. Fodor imagina en términos de un lenguaje mental¹¹. Por otra, tiene como resultado el acceso a una información **nueva** a partir de un conjunto de afirmaciones ya poseídas por aquél que practica la inferencia.

Descrito en una forma macroscópica, que más adelante trataremos de detallar, el modelo consiste básicamente en lo siguiente: 1) un conjunto de premisas que expresan el conocimiento que el agente/receptor pone en juego para la producción/comprensión de la acción verbal, y 2) una conclusión, que enuncia el contenido significativo de la acción verbal realizada; la conclusión ha de representar el **significado** de la acción verbal, el menos en la acepción de significado comunicativo, que es la que nos interesa.

El proceso postulado en el caso de la producción de una acción verbal o preferencia significativa consta de lo siguiente -descrito también de una forma muy general- : 1) el agente tiene como objetivo transmitir una cierta información a un receptor o auditorio¹², 2) para ello pone en juego su conocimiento del conjunto de convenciones o procedimientos que, compartidos por la comunidad comunicativa a que pertenecen tanto él como el receptor, permiten expresar ese significado, 3) utiliza esos procedimientos de forma relativa a una representación de la **situación** en que va a realizar la acción; esa representación constituye básicamente lo que se conoce como **contexto** de la acción verbal y, determina, al menos en parte, el significado de la acción llevada a cabo. La forma general que tendría la inferencia llevada a cabo por el agente espues la siguiente:

(i) Si quiero decir (significar, transmitir, hacer saber...) **x**, entonces, dado **C**, he de hacer **z**

donde **x** representa al objeto de intención comunicativa del agente, es decir, lo que anteriormente denominamos el **significado del hablante**, **C** el contexto pertinente para la expresión de esa intención y **z** la acción verbal que constituye el medio apropiado tanto para su expresión como para su comprensión.

Desde el punto de vista de la recepción, el proceso es básicamente inverso, esto es, consiste esencialmente en la reconstrucción de la intención comunicativa del agente:

(ii) Si **A** ha hecho **z**, entonces, dado **C**, ha querido decir **x**.

Esto es, para la comprensión del significado de la acción verbal, el auditorio ha de partir igualmente de una representación del contexto, que puede coincidir o no con la del agente, y de su conocimiento de las convenciones sociales y comunicativas que restringen el ámbito de las posibles interpretaciones de **z**. Utilizando ambos tipos de conocimiento como parte de la información movilizada en sus conjeturas sobre el sentido de la acción **A**, puede llegar a una conclusión sobre el objeto de su intención comunicativa, esto es, acerca del significado de la acción verbal.

La naturaleza de la inferencia del significado comunicativo

Una vez establecido este marco general de lo que es la inferencia comunicativa, vamos a caracterizar algunas de sus propiedades más generales, antes de pasar a analizar más detalladamente su funcionamiento.

Generalmente, se concibe la computación del significado como un proceso determinista, esto es, que alcanza un resultado neto como término del proceso de cálculo. Pero quizás sería más útil considerar la computación del significado lingüístico como una especie de computación de una función de probabilidad, de tal modo que el resultado fuera obtenido con un cierto grado de plausibilidad. Algunos **movimientos** en el discurso así lo sugieren, como la petición de precisiones, informaciones adicionales, aclaraciones de sentido, explicitación de relaciones de relevancia, etc. De tal modo que la computación del significado no sería algo instantáneo, en la mayoría de las ocasiones, sino que se operaría por etapas sucesivas. Incluso así, también resultaría ilusorio pensar que el significado constituye una entidad perfectamente definida o fija, y que la comunicación requiere como condición necesaria su completo análisis. Lo que sucede más bien es que, independientemente de que exista algo así como un significado completo y determinado, la comunicación y el discurso pueden progresar alimentándose únicamente de cálculos parciales de la información que se pretende transmitir.

Este carácter progresivo y parcial de la computación del significado quizás explique algunas de sus propiedades computacionales más sobresalientes, como es la de su rapidez. En este sentido, no se diferencia de otras tareas cognitivas, cuya característica más sobresaliente es,

*cómo puede ser tan fácil en presencia de un número tan amplio de creencias y potencialmente relevantes*¹³

El proceso inferencial ha de disponer en cualquier caso de dispositivos que posibiliten, y expliquen desde el punto de vista cognitivo, esta familiaridad y rapidez con que sucede la comprensión del significado comunicativo. Dicho de otro modo, el mecanismo inferencial ha de ser de tal naturaleza que permita descartar, entre toda la información de que dispone el sujeto, potencialmente utilizable en la inferencia correspondiente, toda la información **no relevante** para el procesamiento del significado comunicativo.

Para que se comprenda mejor la naturaleza **formal** del problema, consideremos un ejemplo lógico, un ejemplo de inferencia demostrativa en que la verdad de la conclusión se sigue de la verdad de las premisas.

Supongamos que $\alpha \models \beta$, esto es, que β es una consecuencia lógica de α . Sabemos que, si esto es así, β se sigue asimismo de $\{\alpha\} \cup \Delta$, siendo Δ cualquier conjunto de fórmulas. Si trasladáramos esto a un plano psicológico o lingüístico, querría decir que el conjunto de las premisas de que se sigue el significado es **potencialmente** infinito, esto es, que la información que el sujeto podría manejar para la resolución de esa inferencia sería ilimitada. Igualmente se sigue que, si la información de entrada de un sistema computacional es ilimitada, ese sistema computacional no puede realizar tareas computacionales de un modo físicamente viable:

Si proponemos un modelo de una actividad cognitiva en términos de una tarea computatoria con un límite exponencial mínimo, y el modelo no puede limitar el tamaño del input, tal modelo no nos dice **nada** acerca de cómo es físicamente posible esa actividad¹⁴ (se entiende que **físicamente** se refiere tanto al cerebro humano como a cualquier otro sistema computatorio).

Además, la situación se reproduce desde el punto de vista de la producción, puesto que dado $\alpha \models \beta$, también es cierto que $\alpha \models \beta \cup \Delta$, Δ es un conjunto ilimitado de fórmulas. Si no existieran mecanismos generales de restricción de la información utilizada tanto en la producción como en la comprensión de significado, tal tarea cognitiva sería inaccesible para el cerebro humano.

La propiedad de que si $\alpha \models \beta$, entonces $\{\alpha\} \cup \Delta \models \beta$, para cualquier conjunto Δ , se denomina **monotonicidad**, y es una propiedad típica de los cálculos lógicos estándar. De lo dicho anteriormente pues se desprenden las siguientes consecuencias:

- 1) el cálculo del significado comunicativo no es un proceso inferencial que se corresponda con la estructura de la inferencia lógica estándar
- 2) la inferencia del significado comunicativo ha de ser no monotónica, esto es, ha de poder ser **revisable** cuando se considere nueva información, como por ejemplo, la contenida en un sistema central-memoria.

Hay que advertir que hemos mencionado hasta ahora solamente un ejemplo de inferencia lógica, este es, demostrativa. Pero nada indica ni nos autoriza a pensar que ésta es la única clase de inferencia implicada en el proceso de comprensión lingüística. Es posible que tal proceso implique otras formas inferenciales no

estrictamente lógicas, como las inductivas o analógicas, con lo que la necesidad de mecanismos que limiten el tamaño del input es, si cabe, aún más perentoria.

Entendiendo por inferencia demostrativa la estrictamente reducible a los sistemas de lógica estándar que conocemos, dos son las fuentes de tales inferencias en las expresiones lingüísticas. En primer lugar, las basadas en la representación lógica de las oraciones. Todas las teorías lingüísticas modernas, y sobre todo la teoría chomskiana del ligamiento y la rección¹⁵ reconocen un nivel de representación lógica, en el que, entre otras cosas, se incluye información acerca del alcance de los cuantificadores, abarque de la ligadura de las variables, etc. Ese nivel de representación es fuente o motor de inferencias de tipo lógico, como las que relaciona las oraciones

2) si los precios suben, la inflación aumenta

3) si la inflación no aumenta, los precios no suben

Tales inferencias se fundamentan en el contenido lógico de ciertas expresiones conectoras en las lenguas naturales, como 'si...entonces', 'y', 'no', 'todos'. Este contenido lógico nos permite efectuar inferencias razonamientos la verdad supuesta de los enunciados tomados como premisas, y seguramente desempeñan algún papel en ciertos contextos de la comunicación lingüística (por ejemplo, en el razonamiento científico).

Sin embargo, creo que su papel en la computación del significado comunicativo es más bien escaso¹⁶. Supóngase que alguien significa algo que no coincide con lo que hemos denominado **significado oracional o proposicional**, sino que más bien se identifica con un cierto **significado del hablante**. Las relaciones entre los significados de uno y otro tipo son de tal naturaleza que **nada** obliga a que el significado oracional de la expresión utilizada y su significado del hablante tengan la misma forma lógica. En consecuencia, la inferencia que permite pasar de la forma lógica del significado proposicional a la del significado comunicativo (del hablante, por ejemplo) no es una inferencia lógica, sino basada en mecanismos diferentes. Dicho de otro modo, la computación de la forma lógica de una expresión lingüística no es una condición suficiente para la averiguación de la forma lógica de aquello que significa comunicativamente dicha expresión.

Existe otro tipo de inferencias que comparten ciertas propiedades con las inferencias lógicas. Entre ellas se encuentran las inferencias léxicas, las inferencias basadas en el conocimiento de la estructura conceptual de los términos empleados en expresiones oracionales. Son inferencias que permiten concluir de (4) la oración (5):

(4) El terrorista asesinó a un inocente

(5) El terrorista dio muerte a un inocente

Del significado de una pieza léxica se pueden extraer conclusiones que incluyen información representada en el significado atribuido a esas piezas léxicas en un diccionario. Con los adecuados instrumentos y procedimientos lógicos, tal tipo de inferencias se pueden reducir a las inferencias lógicas y, por tanto, les son aplicables las mismas consideraciones ya expuestas: son de escasa utilidad para

averiguar el significado comunicativo en los casos en que éste no coincide con el significado oracional.

Con esta observación no quiero dar la impresión de que la representación de información léxica es un asunto resuelto, ni en las teorías semánticas lingüísticas ni en la IA. Aunque no me detendré en detalles, basta indicar los siguientes aspectos problemáticos de esa representación:

1) En primer lugar y sobre todo, la representación léxica que se emplee ha de partir de una decisión metodológicamente esencial: distinguir entre lo que es conocimiento lingüístico y lo que es conocimiento enciclopédico, entre el conocimiento de la estructura semántica y de los hechos acerca de la realidad representada por el elemento léxico. Aparte de consideraciones teóricas, esta distinción tiene trascendencia para la arquitectura del sistema computatorio: supone una distribución de las inferencias que intervienen en esa computación. Las inferencias léxicas o semánticas han de estar conectadas con un diccionario o lexicón, que no es sino un módulo del sistema lingüístico global, mientras que las inferencias enciclopédicas, por así decir, son las que proceden de una base de conocimiento, que es exterior al sistema lingüístico.

2) Además, existen problemas de representación del significado léxico, esto es, problemas que surgen de las limitaciones expresivas de los lenguajes utilizados en esa representación. En mi opinión, las dos propiedades que causan más problemas en este ámbito son la gradualidad y la tipicidad del significado de muchas expresiones lingüísticas.

En general, se puede observar que la riqueza del mecanismo inferencial demostrativo utilizado por los humanos (o por cualquier otro sistema que compute información) depende de la riqueza de la información representada tanto en el análisis lógico como en el léxico. Lo que es evidente, en cualquier caso, es que los sistemas formales de representación de este tipo de información aún no son lo suficientemente ricos como para dar cuenta de muchas inferencias que operamos automática o semi-automáticamente. Para poner este punto de relieve, bastan dos ejemplos sencillos. En primer lugar, en cuanto al nivel lógico, no existe, en mi opinión, un tratamiento adecuado de la subordinación competitiva. A todos nos parece natural el hecho de que la verdad de (6) se siga de la de (7),

(6) Tenía que despertarme a las ocho

(7) Olvidé que tenía que despertarme a las ocho

y, sin embargo, carecemos de una forma relativamente natural de expresar la forma lógica de (7) de modo que implique la de (6). Este tipo de fenómenos se han agrupado bajo el rótulo de **implicaduras convencionales**, para indicar que no son idénticas a las implicaciones lógicas, pero que no obstante están asimiladas al sistema de la lengua. De tal modo que resulta preciso representar de una forma u otra la información lógica o léxica que da pie a tales tipos de inferencias.

En segundo lugar, es evidente la necesidad de representar de alguna forma el núcleo convencionalizado de los usos de ciertas expresiones lingüísticas. Por poner un ejemplo: la conjunción **pero** tiene un contenido lógico en español (y en otras lenguas) idéntico al de la expresión **y**, pero con unas condiciones de uso diferentes. **Pero** expresa una contraposición entre la información conllevada por dos expresiones lingüísticas entre las que se coloca. Su **uso** está indicado cuando se da tal contraposición y, **de su uso**, se puede inferir que, quien la utiliza, establece esa contraposición. El hecho de que entre dos informaciones exista o se conciba una contraposición es extralingüístico, psicológico si se quiere. Pero el hecho de que tal contraposición pueda expresarse mediante la utilización de **pero** es un hecho lingüístico, comunicativo, que exige una representación en uno u otro nivel de la teoría.

De todos modos, volviendo a las consideraciones generales de representación del significado comunicativo, existe una característica general de la forma en que se ha de efectuar esa representación, en la medida en que contribuya a la determinación del significado de manera demostrativa. Me refiero a que, desde la perspectiva inferencialista que mantenemos, la información lógica o léxica no puede tener el carácter estático o codificado de un diccionario estándar. En sus representación, la información ha de estar orientada hacia la generación de inferencias convencionalmente sancionadas por el uso de la lengua. Si se quiere traducir esto a una terminología más familiar en IA, la información lógica o léxica ha de representarse en forma **procedimental, no declarativa**.

He mencionado la opinión de que la inferencia demostrativa desempeña un escaso papel en la computación del significado comunicativo. A la luz de lo dicho habría que matizar esa opinión, pero quizás nos llevaría ello demasiado lejos. Es preferible considerar ahora la naturaleza y función de la inferencia **no demostrativa** en la computación del significado.

Según D. Sperber y D. Wilson (1986) el proceso de comprensión inferencial no es demostrativo: esto quiere decir que no existe un procedimiento computatorio determinista en el proceso de comprensión del significado, un conjunto de reglas precisamente definidas y de aplicación unívoca que permita construir una **prueba** de lo que X significa al proferir p, aunque pueda existir una comprobación **a posteriori**, esto es, una confirmación de la corrección de la inferencia realizada por el receptor.

¿Qué implicaciones tiene esto para un enfoque cognitivo de la lingüística computatoria? Parece que no resultaría adecuada ninguna modelización tal que impidiera que la comunicación **podiera fallar**, en la cual el ordenador se comportara como un interlocutor ideal que, dado un mínimo de competencia comunicativa por parte del usuario, siempre entendiera lo que significara, que siempre fuera capaz de captar la intenciones comunicativas del emisor. La comunicación natural **falla** de hecho, seguramente por la naturaleza no determinista del proceso comunicativo.

En cualquier caso, la situación, como se puede suponer, es más bien la contraria: la preocupación del teórico es encontrar un modelo que se desempeñe de forma correcta **alguna vez**, o en un conjunto relevante de situaciones. **Relevante** desde el punto de vista empírico, lo que quiere decir ni más ni menos que tenga una tasa alta de éxitos empíricos cuando la conducta comunicativa del usuario es correcta.

Otra característica destacada por Sperber y Wilson es el carácter global de la inferencia comunicativa. En su sentido, **global** quiere decir que la información utilizada como premisa en la inferencia es recuperada de **cualquier** modelo de la memoria o sistema procesador central en el transcurso del proceso de la comprensión. Se opone a la característica de **localidad**, que tienen los procesos inferenciales basados en información independiente del contexto (¿activable en cualquier situación?) -inferencias lógicas, semánticas, lingüísticas- o contextual, pero acotada (por ejemplo, información co-textual).

Las dos características unidas proporcionan una imagen global acerca de la **arquitectura** del sistema en su conjunto:

Un proceso de inferencia no demostrativa con acceso libre a la memoria conceptual: esto suena en realidad como un proceso cognitivo central normal /.../ Mantenemos que la comprensión inferencial no entraña mecanismos especializados. En particular, argumentaremos que el aspecto inferencial de la comprensión verbal entraña la aplicación de procesos inferenciales centrales no especializados, al resultado (output) de procesos lingüísticos no inferenciales , especializados (p. 65-66)

Esta tesis se basa en una presunta posibilidad de separar lo que es información estrictamente lingüística, almacenada en un sistema periférico modular, y lo que es información extralingüística, impresa en en la memoria de un sistema central. La motivación de Sperber y Wilson es que comparten el modelo Chomsky-Fodor de lo que es la competencia lingüística: un módulo genéticamente controlado que impone estructura en los inputs lingüísticos y es impenetrable a información procedente del sistema central (esto es, lo que los científicos cognitivos denominan el **encapsulamiento** del sistema). Aunque existen argumentos serios en favor de esta forma de concebir la arquitectura cognitiva del cerebro humano, no es preciso suponer que ésta sea la única forma de realizar la tarea computatoria de la comprensión del significado, ya sea en el cerebro humano o en otro sistema computatorio. En particular, resulta problemática la suposición de la impenetrabilidad a la información central del módulo lingüístico. Los que han argumentado en contra de esta concepción, han indicado casos en que la asignación de estructura (puramente lógica o sintáctica) depende directamente de la posibilidad de recuperar información del sistema central. Esto es aún más cierto a medida que se asciende en el nivel de la representación: mientras que la representación morfosintáctica parece menos influida por información ajena al módulo lingüístico, no parece suceder lo mismo en otros niveles. Por ejemplo, considérese el problema de escoger entre dos posibles interpretaciones lógicas de

(8) Todos los profesores han suspendido a un alumno

(8a) $\forall x (Px \rightarrow \exists y Sxy)$

(8b) $\exists y (\forall x \rightarrow \forall x (Px \rightarrow Sxy))$

La elección entre una y otra forma lógica depende del conocimiento de la situación a que se puede referir el enunciado: o bien un alumno tiene una acepción inespecífica, con lo cual la forma lógica más adecuada es (8a), o se refiere a un alumno concreto, con lo cual la forma lógica más representativa es (8b). Este tipo de ejemplos, y otros similares, relativos a identificación anafórica, alcance de los cuantificadores, etc., hacen plausible la idea de que, o bien el módulo lingüístico no está encapsulado, o bien la forma lógica se asigna en forma inferencial central fuera de ese módulo.

En cualquier caso, el problema principal no consiste tanto en **de dónde** procede la información manejada en el proceso inferencial, como **de qué modo** se maneja ésta. Si, como parece cierto, la inferencia sobre la intención comunicativa de un hablante no es demostrativa, ¿cuál es el procedimiento utilizado por el auditorio para computarla? El problema reside en que, mientras que disponemos de un buen modelo de la inferencia demostrativa (aunque no tan bueno como algunos piensan) -la lógica- no sucede lo mismo con la inferencia no demostrativa. Tal y como afirman Sperber y Wilson:

Aunque se supone que la inferencia no demostrativa se ha de basar en reglas inductivas de alguna clase, no existe ningún sistema bien elaborado de lógica inductiva que nos proporcione un modelo plausible de los procesos cognitivos centrales. (op. cit. p. 67).

Y lo que es peor, puede que sean falsos dos supuestos que tratan de paliar esa ausencia.

1) que el sistema de inferencia no demostrativa que utilizamos sea único u homogéneo, esto es, puede suceder que usemos varios procedimientos o técnicas de diferente clase para la realización de esa inferencia

2) que el sistema de inferencia sea computatorio en sentido serial, en el sentido de consistir en un conjunto de reglas que, a semejanza de los sistemas deductivos, son aplicados iteradamente hasta la obtención de un resultado.

No tratándose de una inferencia demostrativa, ¿dónde buscar modelos o situaciones similares en la actividad cognitiva humana? Una sugerencia inmediata es la de la propia actividad investigadora científica. A pesar de cientos de años de filosofía de la ciencia, y de unos cuantos de positivismo más o menos explícito, no se ha llegado a una descripción o explicación lo suficientemente plausible acerca de la formación de hipótesis para dar cuenta de un conjunto de fenómenos. Disponemos en cambio de diferentes explicaciones acerca de cómo funciona la actividad científica a partir de ese punto, esto es, de los procesos de contrastación, falsación y

comparación de teorías. Pero lo que sucede antes del comienzo de ese proceso, en la mente del científico, sigue siendo en buena medida un problema irresuelto. Las hipótesis científicas no se siguen demostrativamente de ningún conjunto de axiomas, ni pueden deducirse de ningún modo mecanizable de información antecedente a la propia hipótesis. El proceso de su formación parece consistir en la eliminación, por una u otra estrategia no deductiva, de la información no relevante para la explicación del conjunto de fenómenos que abarca la hipótesis. Por eso surge inmediatamente el paralelismo con respecto a la computación del significado comunicativo: también ésta consiste en la formación (y posterior comprobación) de una conjetura acerca de la intención comunicativa del hablante. Esta comparación entre las dos actividades cognitivas, la formación de hipótesis científicas y la de hipótesis acerca de las intenciones comunicativas del hablante puede resultar tranquilizadora, pero no es mucho más explicativa. Como ha observado J. Fodor¹⁷, nuestra ignorancia abarca ambos casos:

La razón por la que no existe una psicología seria de los procesos centrales es la misma que la de por qué no existe una filosofía seria de la confirmación científica. Ambos casos ilustran la significación de los factores globales en la determinación de las creencias, y nadie está siquiera comenzando a entender cómo afectan esos factores.

En cualquier caso, la caracterización general del problema nos indica la dirección en la que mirar: hacia el campo de la teoría de la resolución de problemas (de los que la complejidad computatoria forma parte, dicho sea de paso) y no al ámbito de la lógica.

Finalmente, indicaremos algunas características generales de la inferencia no demostrativa, pragmática, implicada en la computación del significado comunicativo:

1) en primer lugar, conviene destacar la propiedad de la **gradualidad**. Mientras que la mayor parte de las categorías lógicas y semánticas tienen un carácter **discreto**, el efecto de los factores pragmáticos de la conducta lingüística real es el de diluir ese carácter discreto y convertirlo en **gradual**. Así, por ejemplo, los contrastes entre definido/no definido, anafórico/no anafórico, genérico/específico, sinonimia/no sinonimia, referencial/atributivo, etc, son categorías analíticas que, aplicadas al habla real, constituyen polos ideales en una escala pragmáticamente determinada, esto es, en función del contexto en que se desarrolla el discurso.

2) Frente al procesamiento deductivo, el pragmático es, por definición, un procesamiento contextualmente dependiente. Esto es, en él desempeña un papel esencial el manejo de información extralingüística.

A lo largo del trabajo he mencionado diferentes problemas con los que se enfrenta cualquier sistema que pretenda simular la comprensión lingüística humana, a partir del punto de la representación lógico-semántica de una expresión lingüística. Para finalizar y cumplir con uno de los tres objetivos propuestos al

inicio, me gustaría indicar, siquiera muy brevemente, cuáles son en mi opinión los enfoques y tratamientos de estos problemas que son más prometedores:

A. Problemas generales acerca de la naturaleza y estructura del sistema

De la perspectiva general sobre la comprensión lingüística que hemos mantenido, se desprende la importancia asignada a la dimensión **inferencial** del sistema en cuestión, tanto en sus aspectos semánticos como pragmáticos. Esto es, es preciso asignar una parte esencial del sistema a la implementación de motores inferenciales que operen tanto sobre el conocimiento estrictamente lingüístico, lógico o semántico, como sobre el conocimiento extralingüístico.

Por poner un ejemplo de sistema concreto que se encuentra próximo a este enfoque, mencionaré el sistema que seguramente es el más importante entre los relacionados con el procesamiento del lenguaje natural, el CYC, dirigido por Douglas B. Lenat, de la Universidad de Stanford. En CYC, el sistema inferencial está ubicado en el nivel heurístico, que se concibe como un nivel complementario al epistemológico, representado por la base de conocimiento. En esta base de conocimiento no existe distinción entre lo que pudiera ser conocimiento lingüístico y conocimiento enciclopédico, de tal modo que no existe una diferenciación entre los diferentes niveles en que se desarrolla la comprensión inferencial. Esa redistribución ha de ser efectuada de acuerdo con los propósitos específicos de cada uno de los sistemas basados en CYC, que en realidad no pretende ser sino una ingente base de conocimiento aplicable a diferentes productos tecnológicos, entre ellos los de procesamiento del lenguaje natural (pero también interrogación de bases de datos, traducción automática, interfaces en lenguaje natural...).

Lo interesante de CYC es que la mayor parte del sistema inferencial en no monotónico, de hecho asignable a lo que se denomina **lógica o razonamiento por defecto** (*default logic*), que se encuentra en la base no sólo de la mayor parte del razonamiento común, sino también en las inferencias sobre el significado comunicativo de las expresiones lingüísticas.

B. Problemas de la representación de conocimiento semántico y pragmático

La representación del significado comunicativo requiere la representación del significado proposicional (en ocasiones) y la representación del significado del hablante. El lenguaje más popular para la representación de ambos tipos de significado ha venido siendo el lenguaje de primer orden con diversas extensiones (lógica modal, epistémica, etc.) o sus trasuntos en IA, como PROLOG. Pero hemos señalado algunos problemas de representación del significado que surgen de la relativa pobreza expresiva de este lenguaje lógico cuando se aplica a la complejidad semántica de una lengua natural. En particular, en lo que se refiere a la

representación del significado lingüístico ligado al motor inferencial semántico, nos encontramos al menos con los siguientes problemas:

A. representación de la gradualidad, tanto en el significado de ciertos elementos categoremáticos (nombres comunes, predicados,...), como sincategoremáticos que son desencadenadores de inferencias. Mi opinión general es que la elaboración y aplicación de nuevos lenguajes formales está contribuyendo al enfoque correcto y a la resolución de este tipo de problemas. En particular quisiera mencionar la utilización de la **lógica borrosa** en la representación de la naturaleza gradual del significado lingüístico y de la teoría de los **topoi** en la representación del potencial inferencial no estrictamente lógico de las partículas sincategoremáticas.

B. representación de la tipicidad. En este campo, en que es particularmente inasible la distinción entre conocimiento lingüístico y conocimiento enciclopédico, la lógica no monotónica, en sus diferentes versiones, parece proveer el lenguaje más adecuado para la representación del conocimiento semántico. Por poner un ejemplo, parte del contenido del **prototipo** del significado de *tigre* es que los tales animales tienen rayas, esto es, que $\forall x(\text{Tigrex} \wedge \neg \text{Anormalx} \rightarrow \text{Rayasx})$. La lógica no monotónica no sólo permite representar esa peculiar naturaleza del significado de *tigre*, sino también dar cuenta de las inferencias basadas en el conocimiento de la tipicidad: concluir que un tigre *a* tendrá rayas, a menos que se disponga de información sobre la anormalidad de *a*. Aquí también entran a formar parte supuestos generales en la interpretación del significado de los elementos lingüísticos en IA, como son el supuesto del **mundo cerrado** (*closed world assumption*) o de la **interpretación minimalista**, que también se conoce como **delimitación** o **circunscripción** de McCarthy, pero que no podemos detallar ahora.

El problema que la tipicidad plantea, desde el punto de vista de un sistema de pensamiento del lenguaje natural es de distribución, esto es, de decidir qué parte de esa tipicidad asignar al lexicón, que recoge el conocimiento más o menos compartido por una comunidad lingüística, y qué parte asignar a una base de conocimiento, subconjuntos de la cual pueden formar parte del contexto de un intercambio comunicativo concreto. Creo que el enfoque más apropiado en este caso es pragmático: orientar la distribución dependiendo de los objetivos y fines propuestos en el sistema que se diseñe.

Los dos apartados anteriores hacían referencia a problemas en la representación del significado semántico, pero, de lo anteriormente expuesto, se desprende la importancia del componente pragmático en la representación del significado comunicativo. En este ámbito, me voy a referir a dos problemas que no siempre han recibido la atención que creo se merecen en los sistemas de procesamiento del lenguaje natural:

a. En primer lugar, los problemas que plantea la **representación** del contexto de uso de una expresión lingüística. Con ello, no me refiero tanto a los problemas de las relaciones co-textuales (como la anáfora, catáfora, etc.), sino a la

modelización del interlocutor que creo que es básica en los procesos de comprensión lingüística. Por modelización del interlocutor entiendo primordialmente la atribución de un conjunto de creencias al interlocutor que están relevantemente relacionadas con su conducta lingüística, esto es, entiendo que el contexto comunicativo es esencialmente **doxástico**, frente a enfoques más o menos confesadamente realistas, como pudiera serlo el de la **teoría o lógica de la situación**.

En este ámbito encuentro muy prometedores los recursos expresivos que proporcionan las lógicas epistémicas y autoepistémicas, en la medida en que puedan transformarse en un sistema implementable de reglas de atribución de creencias y de relación de éstas con la formación de intenciones y objetivos comunicativos.

b. En la medida en que parte del significado comunicativo de una preferencia está constituido por las inferencias que no están generadas directamente por el contenido lógico o semántico en sentido amplio, sino por nuestro conocimiento de la situación o del mundo (lo que en teoría del lenguaje se denominan implicaciones conversatorias), es preciso dar adecuada cuenta de tal potencial inferencial en los sistemas de procesamiento del lenguaje natural. Parte de la solución se encuentra por supuesto en la elaboración de bases de conocimiento más amplias de las que hasta ahora han servido de base a esos sistemas, y por ello merecen especial atención los recursos que proporcionan sistemas como CYC. Tales amplias bases, con los mecanismos inferenciales adecuados, monotónicos y no monotónicos, constituyen un instrumento necesario para la computación del significado comunicativo, pero seguramente no será suficiente, será igualmente precisa la incorporación de un sistema de reglas que representen los principios comunicativos de índole pragmática subyacentes a toda interpretación.

En cualquier caso, la potenciación de los recursos formales expresivos y la ampliación de las bases de conocimiento accesibles a los sistemas de procesamiento del lenguaje natural abren una época apasionante en la investigación de este campo y seguramente permiten entrever realizaciones tecnológicas que hoy día se nos antojan aún muy lejanas.

**Dpto. de Lógica y filosofía de la ciencia
UNED-Madrid

Notas

1. R. Weischedel et al. (1990), "Natural language processing", **Annual Review of computer science** 4, 435-452, 437.
2. R. Harris, "The Grammar in your head", en C. Blakemore, **Mindwaves**, 1987, p. 509.
3. Quizás se puede mantener que el único aspecto en que sucede algo estructuralmente similar al proceso de descodificación es en la relación percepción/representación lingüísticas. Si se admite, siguiendo a J. Fodor, que el sistema lingüístico es un sistema **periférico**, al mismo nivel que los sistemas que procesan los inputs perceptivos, entonces se puede mantener que ese sistema lingüístico está sintonizado de tal manera que se efectúa un traslado automático a representación mental de la señal acústica. Dicho de otro modo, la sintonización innata de nuestro sistema perceptivo del lenguaje excluya la

- necesidad de que intervenga en ese proceso de traslado de información procedente del sistema central, a diferencia quizás de lo que sucede en la percepción de señales acústicas no lingüísticas, cuya identificación y categorización requiere la recuperación de información de este tipo. Esa sintonización de nuestro módulo perceptual lingüístico consistiría en una disposición al reconocimiento de estructuras en los estímulos sonoros de índole lingüística (reconocimiento de **universales**). Además, tal disposición siempre se encuentra *enchufada*, por decirlo así. No es posible desconectar el módulo específicamente lingüístico, de tal modo que, literalmente, no podemos dejar de oír expresiones lingüísticas (de nuestra lengua natal) como tales expresiones lingüísticas y no como mero ruido. En este sentido también contrasta con los procesos inferenciales centrales, que suponen un esfuerzo procesador que es optativo, para el individuo. La asignación de significado parece caer del lado de estos procesos inferenciales, escapando a la asignación automática de representación semántica.
4. Para simplificar, hablaré de las oraciones como las entidades lingüísticas que se utilizan en la comunicación, aunque a nadie se le esconde el hecho de que las oraciones, o sus usos concretos, no constituyen la única, o la más adecuada, unidad de análisis de la comunicación.
 5. En inglés, la distinción se establece entre los términos **sentence** y **utterance**.
 6. En inglés, los términos correspondientes son **type** y **token**.
 7. Existen ciertos usos genericos de hoy que no encajan en esta descripción, pero podemos pasarlos por alto.
 8. R.M. Martin (1987), **The meaning fo language**, MIT, p. 19.
 9. J.J. Acero, E. Bustos y D. Quesada, **Introducción a la filosofía del lenguaje**, Madrid: Cátedra, 1982.
 10. D. Sperber y D Wilson, **Relevance**, 1986, pág. 32. Sin embargo, Sperber y Wilson reconocen un cierto papel al modelo semiótico, en combinación con el modelo inferencial, del que sería complementario.
 11. J. Fodor, **The language of thought**, Crowell co., 1975.
 12. Según Sperber y Wilson , es preciso diferenciar entre esta intención puramente informativa y la *intención comunicativa* propiamente dicha, que consiste en la intención de hacer comprensible al auditorio la acción. La distinción es un tanto artificiosa, por lo que la ignoraré.
 13. H. Levesque, (1989), "Logic and the complexity of reasoning", en R.H. Thomason (ed.), **Philosophical logic and artificial intelligence**, Kluwer, p. 83.
 14. H. Levesque, op. cit., pág. 77.
 15. N. Chosky, **Knowledge of language**, para una exposición general. Dicho sea de paso, desconocemos la complejidad computatoria de la teoría chomskiana, puesto que no existe una formalización estricta de sus reglas.
 16. Dejando de lado el diferente contenido semántico de las expresiones de la lengua natural y las contrapartes lógicas. Piénsese que la mayoría de los sistemas lógicos estándar son incapaces de validar la inferencia siguiente:
 - (i) Todos mis alumnos han aprobado
 - (ii) Algunos de mis alumnos han aprobadoLa validación de tal inferencia requiere la introducción del *supuesto* existencial de la expresión "todos".
 17. J. Fodor, **The modularity of mind**, p. 129.