

# PUENTES ENTRE LA LINGÜÍSTICA COMPUTACIONAL Y LA PSICOLINGÜÍSTICA

Juan Carlos Tordera Yllescas  
Universitat de València

**Resumen:** Las Ciencias cognitivas han considerado que se pueden establecer relaciones entre diversas disciplinas como la Filosofía, la Lingüística, la Antropología, la Inteligencia Artificial o la Psicología. El objetivo del presente trabajo es explicitar las relaciones existentes entre la Psicolingüística y la Lingüística computacional.

**Palabras clave:** Lingüística computacional, Psicolingüística, Ciencias cognitivas

**Abstract:** *Cognitive sciences have assumed that there can be relationships between various disciplines such as Philosophy, Linguistics, Anthropology, Artificial Intelligence, or Psychology. This work aims to make explicit these relations between the Psycholinguistics and Computational Linguistics.*

**Key words:** *Computational Linguistics, Psycholinguistics, Cognitive Sciences.*

## 1. LINGÜÍSTICA COMPUTACIONAL Y PSICOLINGÜÍSTICA

### 1.1. La lingüística computacional...

La Lingüística computacional es un campo de conocimiento interdisciplinar en el que la Lingüística teórica ha jugado un papel destacable, especialmente desde la década de los ochenta. Su objetivo es desarrollar sistemas computacionales que simulen, total o parcialmente, las destrezas y habilidades de un hablante real (v.gr.: decodificar un texto escrito, identificar los sonidos del habla, dotar de sentido a un texto...). Dicha disciplina se caracteriza por ser teórico-aplicada, es decir, no solo se trata de desarrollar una teoría propia (no necesariamente coincidente con los presupuestos de la Lingüística teórica), sino que dicha teoría ha de verse cristalizada en un producto informático. Por ejemplo, en cualquier procesador de textos, como puede ser Word, existe un corrector ortográfico y gramatical que, con mayor o menor fortuna, indica las imprecisiones realizadas por el escritor humano. Estos correctores suelen ser diseñados teniendo en cuenta determinados conocimientos formales de lexicología, morfología y sintaxis.

Dentro de las aplicaciones, existen diferentes campos de investigación práctica como son los identificadores de voz, los sintetizadores de voz, los sistemas de recuperación de información, de recuperación, de consulta, sintetizadores de la información, los traductores informáticos... Como se ha indicado, estos productos (si son de calidad), se desarrollan a partir de una fundamentación teórica propia de la Lingüística computacional. En el presente artículo, nos centraremos en la vertiente teórica de la Lingüística computacional, con el fin de establecer las relaciones existentes entre esta disciplina y la Psicolingüística.

Así pues, siguiendo a Halvorsen (1988: 254), se pueden establecer las siguientes subdisciplinas computacionales teóricas: 1) Tratamiento del habla (síntesis de voz y reconocimiento del habla); 2) Análisis, Generación e Interpretación del lenguaje natural; 3) Traducción automática. Las primeras disciplinas computacionales darían cuenta de la onda sonora del habla, es decir, cómo

se produce (síntesis de voz) o cómo se reconoce (reconocimiento del habla). Las disciplinas siguientes se encargarían de desentrañar los aspectos sintácticos, semánticos y pragmáticos de un mensaje bien producido o bien recibido. Finalmente, el desiderátum (todavía no alcanzado plenamente) de la Traducción automática es que un sistema computacional consiga traducir un texto de una lengua origen a otra lengua meta de forma automática (sin intervención humana).

### 1.2. ... y la Psicolingüística

Si nos iniciamos en los estudios que versen sobre Psicolingüística, el primer hecho que puede llamar la atención para el teórico de la Lingüística computacional es el hecho de que sus subdisciplinas son análogas a las indicadas para la Lingüística computacional (vid. Belinchón, Igoa, y Rivière, 1992 [1998] o López García-Molins, 1991). Los campos teóricos de la Psicolingüística tratan de dar cuenta de fenómenos como la percepción de sonidos y su producción, la comprensión y expresión de textos o discursos, y las relaciones establecidas entre dos (o más) lenguas en la mente de un hablante no monolingüe. Es decir, mientras que la Psicolingüística trata de describir y explicar las habilidades y destrezas de un hablante-oyente real, la Lingüística computacional persigue formalizar dicho conocimiento para un “hablante-oyente digital”. Por ello, en los apartados con los que continuaremos la presente sección, se detallarán los puntos de contacto que existe entre estas disciplinas de la Lingüística computacional y la Psicolingüística.

Por otro lado, se ha de hacer notar el hecho de que no ha pasado desapercibido para los estudiosos e investigadores de ambas disciplinas que existe una importante relación entre la Lingüística computacional y la Psicolingüística/Neurolingüística o, dicho en otros términos, entre el ordenador y la mente humana:

(1)

- a. La similitud entre los ordenadores y la mente humana es obvia y ha sido explotada en ambas direcciones. Los ingenieros electrónicos y sobre todo los técnicos de Inteligencia Artificial tratan de trasladar intuitivamente sus ideas sobre el funcionamiento mental al campo del ordenador. Por su parte, los psicólogos cognitivos toman como modelo el ordenador para plantear hipótesis psicológicas y elaborar interpretaciones teóricas. (De Vega, 1984: 31).
- b. Otro factor relevante que ha acarreado el resurgimiento de numerosos tipos de gramáticas dentro de la LC es la relación interdisciplinaria con la psicología, la lógica y las matemáticas. La LC, a causa de la influencia de reflexiones psicológicas, se interesa más por la semántica que por la sintaxis e intenta dilucidar los procesos cognitivos subyacentes al lenguaje para así poder modelarlos y “computerizarlos” en ordenadores “inteligentes” que simulen una capacidad de comprensión. (Meya y Huber, 1986: 93).

Estas afirmaciones recogidas defendidas por otros autores (vid. Belinchón, Igoa y Rivière, 1992 [1998]: 294, o Cuetos, 1999: 564-565) ponen de manifiesto que las investigaciones de los psicolingüistas y de los lingüistas computacionales no han sido tareas aisladas, sino que los avances de una y otra disciplina se han visto reforzados mutuamente. Y esto es así, porque, tal como se recogen en (1a) y (1b), se puede establecer una relación analógica entre el ordenador y la mente humana<sup>1</sup>. Por tanto, es esperable que los problemas abordados sean similares en ambas disciplinas y así esperamos demostrarlo.

No obstante, tal como se desarrollará en las conclusiones, las relaciones de semejanza no son relaciones de identidad, es decir, los problemas ante los que se enfrentan la Psicolingüística y la Lingüística computacional pueden ser parejos, pero no así la forma de resolverlos y ello es así, debido a que el lenguaje natural (el lenguaje del ser humano) no es idéntico al lenguaje

---

<sup>1</sup> Sin embargo, tal como advierte De Vega (1984: 31), dicha relación es solo de semejanza, no de identidad.

artificial (el propio de los sistemas computacionales). Este se reduce a un conjunto de reglas y símbolos; el lenguaje artificial es un puro álgebra que, *per se*, no se produce con intencionalidad. En cambio, el lenguaje natural va más allá de los aspectos digitales (v.gr.: fonología, morfosintaxis...) y está ligado a la intencionalidad, la metarrepresentación (qué cree el hablante que sabe el oyente), los deseos, las emociones..., todos ellos, aspectos analógicos de los que difícilmente puede dar cuenta un “hablante-oyente digital”.

## 2. PRODUCCIÓN Y RECONOCIMIENTO DEL HABLA

En el plano de la producción del habla, cabe indicar que, al igual que ocurre con la síntesis de voz, es decir, la producción automática de la voz humana (Tordera Yllescas, 2011), uno de los mayores retos que plantea la Psicolingüística es describir cómo se articulan los sonidos, puesto que no parece adecuado defender que el sistema nervioso ordene al aparato bucofonatorio para articular sonido a sonido. Es decir, la Psicolingüística/Neurolingüística presta atención a la *coarticulación*, esto es, a la contaminación acústica de los sonidos por los sonidos anteriores y posteriores, al pronunciarse todos conjuntamente (Caplan, 1987 [1992]: 245; Belinchón, Igoa y Rivière, 1992 [1998]: 614-617).

- (2) [...] [la planificación para producir diversos sonidos de forma contigua] afecta, como mínimo, a los músculos implicados en la producción de una sílaba; al mismo tiempo, permite dar cuenta también de la existencia de movimientos de anticipación de los sonidos finales de las palabras e incluso de sonidos que ocupan hasta seis segmentos posteriores a aquel que se está emitiendo. Estos movimientos de anticipación “contaminan” o hacen aún más patente el carácter co-articulado de la emisión de los sonidos del habla. En este sentido, restan credibilidad a la hipótesis de que la planificación articulatoria es un derivado directo de los rasgos fonéticos y/o que el programa motor consta de instrucciones articulatorias invariantes que se aplican una a una con independencia del contexto (i.e., de cuáles sean los fonemas adyacentes). (Belinchón, Igoa y Rivière 1992 [1998]: 615).

Igualmente, un reto que plantea la síntesis de voz es la producción de los aspectos prosódicos. La prosodia (aspecto del lenguaje necesario para una buena comprensión del mensaje) es un aspecto que, en síntesis, está vinculada a la sintaxis (Llisterri et alii, 2003: 8-9; Hualde, 2003; Llisterri et alii, 2004). Igualmente, en Psicolingüística, un tema investigado es cómo la correcta pronunciación de la entonación, el acento, las pausas, la duración... influye en la comprensión del mensaje en casos de ambigüedad o en casos de dirigir la atención hacia un elemento del mensaje (Butterworth, 1990; Belinchón, Igoa y Rivière, 1992 [1998]: 594-597).

En el caso del uso de claves prosódicas para desambiguar los enunciados, Belinchón, Igoa y Rivière (1992 [1998]: 595) exponen el ejemplo de (3) y, en el caso del uso de claves prosódicas con fines enfáticos o contrastivos, Belinchón, Igoa y Rivière (1992 [1998]: 596) aportan el ejemplo de (4).

- (3) a. Si Solti dirige la orquesta / tendrá un gran éxito.  
b. Si Solti dirige / la orquesta tendrá un gran éxito.
- (4) a. TU HERMANO robó el dinero (no mi padre).  
b. Tu hermano robó EL DINERO (no las joyas).

En (3), la entonación y la pausa permiten delimitar de qué verbo depende el sintagma “la orquesta”, mientras que, en (4), la intensidad y la entonación permiten identificar el elemento remático o información nueva. Precisamente, la prosodia para estos casos también es tenida en cuenta por la síntesis de habla, ya que facilitaban la producción de mensajes comprensibles e

inteligibles para el receptor, tal como se muestra a continuación (Llisterri et alii, 2003: 8-9; Hualde, 2003; Llisterri et alii, 2004):

- (5) a. -¿Dónde lo ponemos?  
- *Si te parece bajo el armario. (prep)*  
b. - ¿Qué bajarás ahora?  
- *Si te parece bajo el armario. (verb)*  
c. -¿En qué caso podemos poner una repisa ahí arriba?  
- *(Pues) si te parece bajo el armario. (adj).*
- (6) a. [La vieja]<sub>SN</sub> [lanza [la amenaza]<sub>SN</sub> ]<sub>SV</sub>  
b. [La vieja lanza]<sub>SN</sub> [la amenaza]<sub>SV</sub>
- (7) a. - ¿DÓNDE ha ido Juan? - A LA FIESTA ha ido.  
b. No ha cedido DOS caramelos, sino TRES.

En (5), dependiendo de la categoría que se le asigne a “bajo”, ha de ser pronunciada de forma tónica (verbo o adjetivo) o de forma átona (preposición). En (6), dependiendo del lugar en donde se asigne una pausa potencial se pueden entender dos cosas diferentes: que la señora mayor lanza una amenaza (6a) o que la lanza vieja amenaza a alguien (6b). Finalmente, a partir de los ejemplos de (7), el acento de contraste sirve para indicar cuál puede ser el foco en un enunciado.

Respecto a la percepción del habla, cabe señalar que, al igual que ocurre en el reconocimiento del habla, en Psicolingüística, un desafío al que han prestado atención los expertos es explicar cómo es posible que el ser humano reconozca como unidad invariante una unidad acústica determinada (v.gr.: un fonema, una sílaba...) que sufre no pocas variaciones (Belinchón, Igoa y Rivière, 1992 [1998]: 319-349; Sebastián, Bosch y Costa, 1999, 53-88; Yeni-Komshian, 2000: 113-167).

- 8) Resultaría relativamente sencillo desarrollar modelos para los procesos de percepción del habla, si cada sonido distintivo de una lengua estuviera asociado a un patrón acústico estándar. Desgraciadamente, ni el habla, ni sus características acústicas resultan ser tan sencillas. En vez de presentar patrones *invariantes* (estándares, invariables), los sonidos de la lengua varían considerablemente en lo que respecta a sus características acústicas por varias razones (Yeni-Komshian 2000: 118).

Los psicolingüistas señalan como problema el hecho de que el reconocimiento de los sonidos se ve afectado, tanto por las diferencias de interlocutor como por las de intralocutor (dado que, para un mismo hablante, los sonidos pueden variar por el contexto fónico). Es decir, precisamente, el problema suscitado es el mismo que se plantea en el reconocimiento del habla en Lingüística computacional.

- (9) Una razón muy importante para esta variación reside en que la producción, y en consecuencia la acústica, de un determinado segmento fonético cambia dependiendo del contexto en el que se inserte. [...] Estos efectos del contexto, que resultan del solapamiento de los movimientos del habla, se denominan efectos de coarticulación.

[...] Por otro lado, las propiedades físicas de los sonidos de la lengua, especialmente de las vocales, varían según hayan sido emitidas por hombres, mujeres o niños, cuyos tractos vocales difieren en tamaño y configuración

[...] Otro factor que ocasiona la variación de las propiedades acústicas de los sonidos del habla es que no pronunciamos dos veces la misma expresión de la misma manera. Este tipo de variación es característico del lenguaje natural; si lo suprimiéramos, lo que decimos sonaría como algo artificial. (Yeni-Komshian, 2000: 118-119).

Tal es la influencia que se puede registrar entre los estudios psicolingüísticos del habla y los estudios de Lingüística computacional sobre el habla, que no ha de extrañar que se hayan utilizado aplicaciones del tratamiento del habla para las investigaciones psicolingüísticas. Así por ejemplo, en Yeni-Komshian (2000: 132-133), se describe cómo se utilizó un sintetizador pionero, diseñado a principios de los cincuenta por Franklin Cooper (un ingeniero), Alvin Liberman (un psicólogo) y Pierre Delatre (un lingüista) para detectar las propiedades acústicas perceptivamente relevantes (v.gr.: los dos primeros formantes en el caso de las vocales) frente a las propiedades redundantes es decir, para determinar qué rasgos acústicos son, psicolingüísticamente, prominentes.

### 3. ANÁLISIS, GENERACIÓN DEL LENGUAJE

En la década de los sesenta, el grupo de investigación MITRE fue precursor al tener en consideración las aportaciones de la Lingüística teórica en su labor; en concreto, trató de utilizar en su investigación computacional el modelo estándar la Gramática Generativa Transformacional (GGT), de Noam Chomsky. Desgraciadamente, su aplicación resultó infructuosa, dada la complejidad lingüística que conllevaba dicha teoría (en concreto, determinar si ante una oración se había aplicado o no una determinada regla transformacional).

Pero el Generativismo no supuso solo un hito para la Lingüística computacional, sino también para la Psicolingüística (Owens, 1987 [1990]: 37-48; Belinchón, Igoa y Rivière, 1992 [1998]: 260-274; Wingfield y Titone, 2000: 251-255). Así por ejemplo, una investigación clásica de la Psicolingüística, relacionado con la Gramática Generativa Transformacional, ha sido el estudio de cómo se procesan las oraciones activas y pasivas. De acuerdo con la Gramática Generativa Transformacional, dado que, en la generación de las oraciones pasivas, se ha tenido que aplicar una regla adicional, una regla transformacional (que, en el modelo de Principios y Parámetros, es “muévase-*a*”), lo predecible sería que el hablante-oyente ideal tardara más tiempo en comprender o procesar las oraciones pasivas que no las activas. Y, para corroborar esta predicción, algunos estudios psicolingüísticos calcularon el tiempo de respuesta que tardaban diferentes sujetos al comprender las oraciones activas y pasivas.

Y, al igual que la Gramática Generativa Transformacional cedió paso a diferentes teorías gramaticales aplicables a la Lingüística computacional<sup>2</sup>, del mismo modo, se han llevado a cabo investigaciones psicolingüísticas cuya base gramatical han sido las redes de transición aumentadas (Mitchell, 1987 [1990]: 189-192; Belinchón, Igoa y Rivière, 1992 [1998]: 442-446), la Gramática Léxico-Funcional (Mitchell, 1987 [1990]: 193-194; Belinchón, Igoa y Rivière, 1992 [1998]: 274-275) o la Gramática de casos de Fillmore (Owens, 1987 [1990]: 48-56; Belinchón, Igoa y Rivière, 1992 [1998]: 268-270). E igualmente, se puede registrar el estudio de otras teorías lingüísticas no sintácticas en la Psicolingüística, que han tenido su relevancia en Lingüística computacional. Por citar algunas teorías, se pueden mencionar el modelo de representación por redes semántica de Collins y Quillian (Reeves, Hirsh-Pasek y Golinkoff, 2000: 217-223) o la Teoría de los guiones, planes y metas de Schank y Abelson, teoría que nace dentro de la Inteligencia Artificial, pero que tendrá su relevancia teórica en la Lingüística teórica (Belinchón, Igoa y Rivière, 1992 [1998]: 472-477 y 519-522).

<sup>2</sup> Vid. Grishman (1985 [1990]: 63-67, 122-124 y 144-147), Sparck Jones (1992: 55-57), Moreno Sandoval (1998: 42-43), Martí Antonín y Castellón Masalles (2000: 14-15 y 140-142), Bresnan, J. (1982a), Bresnan, J. (1982b), Bresnan, J. (2001), Dalrymple, M. (2001), Gazdar, G., E. Klein, et alii (1985), Pollard, C. e I. A. Sag (1987), Pollard, C. e I. A. Sag (1994), Sag, I. A., T. Wasow y E. Bender (2003).

Es interesante constatar el hecho de que las gramáticas empleadas en la Lingüística computacional y en la Psicolingüística son las mismas: si el objetivo de la Lingüística computacional es “simular” la competencia lingüística o algunas de sus subcompetencias de un hablante oyente humano, entonces el hecho de que una determinada gramática no tenga únicamente justificación aplicada (como exigiría la Lingüística computacional), sino, además, psicológica, nos puede inducir a pensar que el grado de simulación alcanzado es adecuado. Así por ejemplo, en Lingüística computacional, las gramáticas que han gozado de mayor auge han sido las gramáticas lexicalistas (Tordera Yllescas, 2008), esto es, aquellas que consideran que la sintaxis está codificada en el léxico (v.gr.: la Gramática Léxico-Funcional, la Gramática de Estructura Sintagmática Generalizada o la Gramática de Estructura Sintagmática orientada al Núcleo<sup>3</sup>). Paralelamente, en autores como Hernández Sacristán (2003), se puede apreciar la constatación y defensa de que la rehabilitación de sujetos con agramatismos no puede estar separada de la rehabilitación léxica, pues léxico y sintaxis se interrelacionan. De hecho, como recuerda el autor respecto al desarrollo del lenguaje, es a los dos años de vida del niño cuando se puede hablar de una protosintaxis, que es, precisamente, cuando se produce la explosión de vocabulario. Por tanto, podemos afirmar que el uso de gramáticas lexicalistas en Lingüística computacional es, psicolingüísticamente, más que plausible y acertado.

Y si el conocimiento declarativo de la Lingüística computacional (i.e., las gramáticas) mantiene estrechas relaciones con la Psicolingüística, de igual modo, el conocimiento procedimental (i.e., ¿cómo analizar a partir de las reglas dadas?) ha sido otro punto de contacto. Al igual que la Lingüística computacional ha prestado especial atención a los *parsers* o analizadores sintácticos (que establecen qué reglas se aplican, en qué orden, qué reglas deben ser desechadas...), la Psicolingüística también ha prestado interés por cómo los seres humanos procesan y analizan sintácticamente las oraciones. En concreto, la cuestión que se plantean en las investigaciones psicolingüísticas es si el procesamiento o análisis sintáctico es autónomo (y, consecuentemente, cualquier proceso de desambiguación llevado a cabo por la aportación de información semántica o pragmática es posterior al análisis sintáctico) o, por el contrario, interactivo y, por tanto, en el mismo proceso de análisis, se recibe información de tipo semántico o pragmático que puede ayudar a guiar el análisis. Por ejemplo, en la oración “El policía ve al caballo con los prismáticos”, se interpreta preferentemente, por nuestro conocimiento del mundo, que el sintagma preposicional “con los prismáticos” es un complemento verbal, mientras que en la oración “El policía ve al caballo con la montura desgastada” se interpreta preferentemente, por nuestros conocimientos del mundo, “con la montura desgastada” como modificador de “caballo”. Asimismo, otra cuestión que ha suscitado interés es la forma en que se analizan las oraciones. En concreto, si el analizador lleva a cabo el proceso de análisis de forma determinística o no determinística, es decir, dada una categoría no terminal que se puede reescribir de más de una forma (v.gr.: un SN se puede reescribir como “SN → Det + N + ADJ”, como “SN → Det + N” o bien como “SN → Det + ADJ + N”), el analizador tiene en cuenta todas las posibilidades de análisis o solo una (en este último caso, el parser es determinístico). Y lo significativo es que existen estudios psicolingüísticos como el de Mitchell (1987 [1990]: 192-193) que explican las ventajas y desventajas de los análisis determinísticos no en función de los resultados aportados por pruebas o tests psicolingüísticos, sino por los resultados obtenidos por el parser PARSIFAL de Marcus (parser determinístico), con lo que se vuelve a hacer manifiesta la interrelación Psicolingüística-Lingüística computacional.

Cuando los psicolingüistas estudian cómo se analizan las oraciones, una pregunta capital que tratan de responder es cómo se interpretan las oraciones ambiguas (v.gr.: “El policía ve al niño con los prismáticos”); es decir, si el ser humano tiene en consideración todas las posibles interpretaciones durante el análisis o, por el contrario, solo tiene en cuenta una única interpreta-

<sup>3</sup> Vid. Bresnan, J. (1982a), Bresnan, J. (1982b), Bresnan, J. (2001), Dalrymple, M. (2001), Gazdar, G., E. Klein, et alii (1985), Pollard, C. e I. A. Sag (1987), Pollard, C. e I. A. Sag (1994), Sag, I. A., T. Wasow y E. Bender (2003).

ción que puede ser ratificada al final del análisis<sup>4</sup>. En este segundo caso, con el fin de justificar que el ser humano solo tiene en cuenta una única interpretación, los psicolingüistas partidarios de esta hipótesis suelen ejemplificar sus ideas mediante las oraciones denominadas oraciones de sendero de jardín o de vía muerta (*garden-path sentences*), esto es, oraciones que, al terminar de ser procesadas, el receptor (oyente o lector) puede verse obligado a retroceder hacia atrás ante un análisis erróneo.

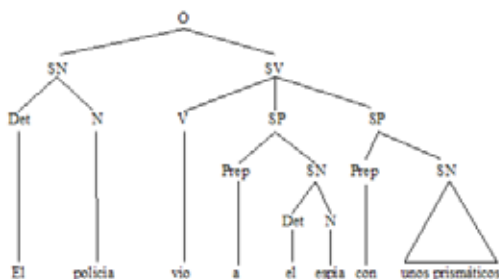
- (10) a. Cuando el bibliotecario subió [los libros] se cayeron al suelo.  
b. Mientras el granjero comía [el pollo] se escapaba del corral.

En las oraciones de (10), el receptor puede verse obligado a volver a escuchar/leer la oración (esto es, dar marcha atrás ante un camino interpretativo erróneo) si ha considerado que los sintagmas “los libros” y “el pollo” es un complemento de los verbos “subió” y “comía” respectivamente, en lugar de interpretar dichos sintagmas como los sintagmas nominales que funcionan como sujeto de la oración principal. Como el ser humano ha de volver a interpretar las oraciones de (10), podría interpretarse que el ser humano no tiene en cuenta todas las posibles interpretaciones de la oración, sino una única interpretación (en este caso errónea) y, por ello, ha tenido que dar una marcha atrás.

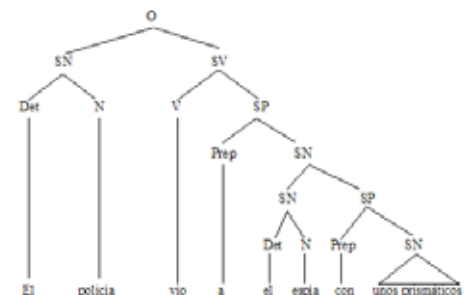
Quienes apoyándose en las oraciones de vía muerta o de sendero de jardín sostienen que el ser humano solo tiene en cuenta una única interpretación en el análisis, defienden que existen dos estrategias que guían al ser humano en la interpretación preferida de las oraciones ambiguas. Según algunos autores, estas estrategias de análisis se denominan adjunción mínima (*minimal attachment*) y cierre tardío (*late closure*).

Según la estrategia de la adjunción mínima, se preferirán aquellos análisis que requieren la generación de un menor número de nudos sintagmáticos. Así, en Belinchón, Igoa y Rivière (1992 [1998]: 431-432), Carreiras y Meseguer (1999: 171-172) y Wingfield y Titone (2000: 260-261), se recoge que el motivo por el que se prefiere interpretar el sintagma preposicional «con unos prismáticos» de la oración “El policía vio a el espía con unos prismáticos” como un modificador verbal y no nominal obedece a esta estrategia. Adaptando los ejemplos aportados por los autores citados, una oración como “El policía vio a el espía con unos prismáticos” se podría analizar de las siguientes maneras:

- (11) a.



- b.



Como se observa a partir de los análisis de (11), la interpretación del sintagma preposicional “con unos prismáticos” como modificador verbal (análisis de (11a)) requiere menos nudos que la

<sup>4</sup> Para quienes defienden que el análisis sintáctico es autónomo y que, además, el análisis es determinista, el ser humano solo podrá tener en consideración una única interpretación al procesar la oración ambigua.

interpretación como modificador nominal (análisis de (11b)), ya que en este último, se ha tenido que reescribir el SN como “SN+SP”, lo que no ocurre en el primer análisis. Así pues, tal como prevé la adjunción mínima, el primer análisis es el preferido.

Respecto al cierre tardío, tal como se recoge en Belinchón, Igoa y Rivière (1992 [1998]: 433), Carreiras y Meseguer (1999: 173) y Wingfield y Titone (2000: 260-261), esta estrategia de análisis sintáctico predice que se preferirán aquellos análisis cuyos nuevos constituyentes se interpretan como parte del constituyente actual en proceso de análisis (esto es, como constituyentes de constituyentes inferiores mejor que constituyentes de constituyentes superiores). Por ejemplo, en (12), por esta estrategia, se prefiere interpretar que el adverbio “ayer” modifica al verbo de la oración subordinada y no al verbo de la oración principal.

- (12) a. Juan dijo que [O María [SV fue a la biblioteca ayer]]  
= Juan dijo que fue ayer cuando María fue a la biblioteca.  
b. [O Juan [SV dijo que [O María fue a la biblioteca] ayer]]  
= Ayer fue cuando Juan dijo que María fue a la biblioteca.

Las estrategias de la adjunción mínima y el cierre tardío pueden actuar como estrategias complementarias o bien como estrategias en conflicto. Por ejemplo, en la oración “Mientras el granjero comía [el pollo] se escapaba del corral”, las dos estrategias preferirían aquella interpretación (errónea) por la que “el pollo” sería un complemento verbal respecto a “comía”, ya que, por la adjunción mínima, de este modo no se habría de proponer una nueva oración (pues, en la interpretación correcta, “el pollo” es el SN dominado directamente por O) y, por el cierre tardío, sería preferible la interpretación en la que el constituyente nuevo (en este caso, “el pollo”) puede formar parte del constituyente en proceso de análisis (en este caso, del sintagma verbal “comía”). Sin embargo, en el caso de la oración “El policía vio al espía con unos prismáticos”, aunque la adjunción mínima preferiría la interpretación por la que el sintagma preposicional “con unos prismáticos” es modificador verbal, el cierre tardío preferiría la interpretación en la que dicho sintagma preposicional sería modificador nominal de “espía”, ya que el sintagma nominal sería el constituyente en proceso de análisis al que se puede adjuntar el sintagma preposicional. De hecho, en oraciones como “El policía vio al caballo con una montura”, la interpretación preferida sería aquella en la que el sintagma preposicional “con una montura” es un modificador nominal<sup>5</sup>.

Una vez más, lo interesante de las estrategias de análisis investigadas por la Psicolingüística, como la adjunción mínima y el cierre tardío, es que han sido estrategias estudiadas igualmente por la Lingüística computacional, tal como se recoge en Allen (1995: 160-161), con el fin de ofrecer una guía de análisis a los analizadores sintácticos o *parsers* ante los casos de ambigüedad sintáctica. Obsérvese que este modo de proceder implica que se espera que el sistema computacional analice (e interprete) una oración al igual que lo hace el ser humano; es decir, gracias a la interrelación entre la Psicolingüística y la Lingüística computacional, se puede hablar de una (casi) auténtica simulación computacional de la comprensión del lenguaje natural.

#### 4. TRADUCCIÓN Y OTROS...

En Lingüística computacional, existen, básicamente, dos métodos de llevar a cabo la Traducción automática: el método de transferencia y el método por interlengua (Hutchins y Somers,

<sup>5</sup> Cabe recordar, no obstante, que, como se recoge en Belinchón, Igoa y Rivière (1992 [1998]: 431-433), Carreiras y Meseguer (1999: 170-176) y Wingfield y Titone (2000: 260-261), la preferencia de una interpretación u otra no solo responde a criterios sintácticos (esto es, estrategias de análisis), sino también a criterios léxicos, semánticos y pragmáticos. Si se tiene en consideración que el análisis sintáctico no se realiza de forma autónoma, se puede entender por qué el sintagma preposicional “con una montura” prefiere interpretarse como modificador nominal y no verbal, ya que no tiene mucho sentido considerar que el policía llevaba puesto una montura cuando veía al caballo.



1992 [1995]: 116-121, 160-182 y 192-198; Martí Antonín y Castellón Masselles, 2000: 33-34; Alonso Martín, 2003: 102-107 y 109-126; Hutchins, 2003: 502-505). Los métodos de transferencia y de interlengua se pueden identificar por los niveles lingüísticos implicados; en los de transferencia, intervienen los componentes léxico y sintáctico(-semántico) y, en los segundos, en los métodos de interlengua, se incluye, además, el componente semántico. El sistema de traducción basado en el método de interlengua se caracteriza por generar, a partir del texto origen, una representación (semántica). Dicha representación abstracta (una interlengua) supone un estadio intermedio en el proceso de traducción, a partir del cual se obtendrá el texto meta. Como la representación abstracta es neutra, es decir, no existen propiedades lexicológicas (morfológicas), ni sintácticas, ni semánticas exclusivas de una lengua que se proyecten hacia dicha representación intermedia, los módulos de análisis y generación de cada lengua son independientes entre sí, lo que posibilita que la adición de una lengua al sistema de traducción solo implique un analizador y un generador más de dicha lengua (Hutchins y Somers, 1992 [1995]: 116-118).

En el ámbito de la Traducción automática, también se pueden establecer paralelismos con la Psicolingüística y, en concreto, con las investigaciones referentes al bilingüismo (o plurilingüismo). En Harding-Esch y Riley (1986 [2003]: 51-55) y en López García (1991: 92-100), se recoge un par de conceptos clásicos referentes al estudio de la mente bilingüe, que, aunque ha perdido fuerza teórica en los últimos años, no dejan de ser conceptos a los que merece hacer referencia. En las obras señaladas, se indica que Erving y Osgood, influidos por Weinreich, propusieron en 1954 los conceptos de *bilingüismo coordinado* y de *bilingüismo compuesto*. El *bilingüismo compuesto* sería propio del “individuo que posee un solo contenido cognitivo para dos etiquetas formales diferentes, y el *bilingüismo coordinado*, que es aquel donde cada etiqueta (i.e., cada forma fónica) arrastra su propio contenido, de manera que los términos se acercan sin igualarse” (López García, 1991: 93).

(13)

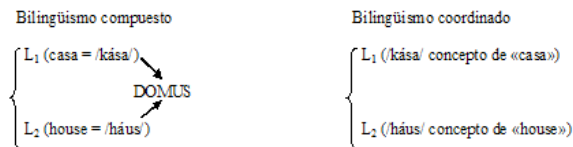


Figura 1. Extraído de López García, 1991: 93.

A nuestro parecer, se puede establecer un paralelismo entre el bilingüismo coordinado y el método de transferencia de la Traducción automática, así como entre el bilingüismo compuesto y el método de interlengua. En el caso del bilingüismo coordinado, para traducir un concepto de la lengua L<sub>1</sub> a la lengua L<sub>2</sub>, el proceso que ha de llevar a cabo el individuo bilingüe es buscar los equivalentes conceptuales de cada lengua implicada, es decir, ha de realizar un proceso de transferencia (v.gr.: “casa” → “house” o, viceversa, “house” → “casa”). En el bilingüismo (o plurilingüismo) compuesto, son indiferentes las lenguas implicadas para acceder al significado de las palabras de cada lengua; las palabras de las diferentes lenguas remiten a un único significado cognitivo que es compartido, es decir, se podría decir que existe una interlengua en la que hallan su significado las palabras del resto de lenguas.

Por último, nos gustaría registrar una última conexión entre la Psicolingüística y la Lingüística computacional; y dicha relación la hemos encontrado en un campo de investigación de la Psicolingüística que, *a priori*, podría parecer alejado de la Lingüística computacional, a saber, las investigaciones y estudios relacionados con el desarrollo del lenguaje. Y el caso es que un concepto clave de estos estudios es el de *maternés*. Según se recoge en Owens (1987 [1990]:

211-217), el maternés es la modalidad lingüística que utilizan los adultos (v.gr.: las madres, de ahí la denominación de maternés) cuando interactúan con los bebés para facilitar la comprensión del mensaje y, por tanto, ayudar implícitamente en el desarrollo del lenguaje. Esta modalidad se caracteriza, prosódicamente, por un tempo o velocidad de habla lento, una entonación exagerada, una intensidad marcada; léxicamente, por un vocabulario restringido temáticamente y no complicado; y, sintácticamente, por oraciones simples (frente a la subordinación) y bien formadas (esto es, gramaticales) sin interrupciones ni reelaboraciones. Pues bien, nosotros hemos observado que, para las distintas aplicaciones disponibles en las que interviene el tratamiento del habla y, especialmente, el reconocimiento del habla, los locutores (quizá al considerar que las máquinas “no son tan maduras e inteligentes como el locutor adulto”) utilizan una modalidad del lenguaje próxima al maternés, solo que, en este caso, el mensaje no va dirigido a un niño sino a un sistema computacional. Por ello, dado el paralelismo apuntado, nos gustaría denominar a dicha modalidad lingüística como el *computernés*.

## 5. CONCLUSIONES

A la luz de los datos aportados, se puede observar que los retos estudiados tanto por lingüistas computacionales como por psicolingüistas son parejos y que, en ocasiones, las soluciones aportadas por uno u otro campo han sido aprovechadas de forma complementaria. Por ello, el futuro de la investigación no puede dejar de ser interdisciplinar en cualquier caso.

No obstante, pese a las semejanzas señaladas, no quisiéramos concluir diciendo que una computadora puede ser tratada igual que la mente humana. “Simular” no implica “equivaler”. Es decir, la mente humana y el ordenador no son lo mismo. Aunque la mente humana y el ordenador se caracterizan por codificar, retener y operar con símbolos y representaciones internas, cabe advertir que el soporte físico o hardware no es el mismo. El sistema nervioso del hombre consta de millones de neuronas que son capaces de establecer miles de complejas conexiones sinápticas y, hasta la fecha, la ingeniería informática ha sido incapaz de simular esta complejidad del sistema nervioso (o no en términos absolutos). Por ello, a nuestro juicio, la heurística (esto es, las técnicas y estrategias de resolver un problema o, dicho en otros términos, la capacidad de plantear un algoritmo) de la mente humana y del ordenador no son idénticos.

Por tanto, hemos de estar de acuerdo con De Vega (1984: 31-32), cuando advierte a este respecto que la analogía solo es funcional y no física, por lo que no parecería adecuado que se estableciera una identificación absoluta entre el modo (específico) de procesar y memorizar la información del ser humano y el modo de procesarla y memorizarla de la máquina. Piénsese, por ejemplo, que una máquina consume menos memoria al memorizar el número 2.987 que 1.000.000 (esto es, un millón); en contraste, posiblemente, para un ser humano, sea más fácil recordar el número 1.000.000. Asimismo, si un ordenador tuviera que retener nombres propios como “Ana” y “María Jesús”, gastaría mayor memoria en retener el nombre propio “María Jesús” que no “Ana” (pues este contiene menos caracteres que aquel). Sin embargo, si un ser humano tiene como hermana o pareja a una persona que se llame “María Jesús” (o si se llama así el propio sujeto), posiblemente le cueste menos retener este nombre que no “Ana”, ya que, al ser humano le es más fácil memorizar determinada información si puede relacionarla con la almacenada previamente. Por último, piénsese cómo suele memorizar el ser humano los números de teléfono. Nos es más fácil memorizar un número telefónico agrupando cifras (96-154-42-63) que no como un todo (961544263). Sin embargo, una máquina no recurre a este tipo de estrategias. Es decir, el ser humano cuenta con una heurística propia de procesar y memorizar la información que están ausentes en las máquinas (al menos, por el momento). Por tanto, el desarrollo de la Lingüística computacional habrá de tener en consideración estos hechos en su aspiración de *simular* la comunicación humana.

## REFERENCIAS

- Allen, J. (1995<sup>2</sup>). *Natural Language Understanding*. Redwood: Benjamin/Cummings.
- Alonso Martín, J.A., (2003). "La traducción automática", en M<sup>a</sup>. A. Martí Antonín (eds.) *Tecnologías del lenguaje*. Barcelona: UOC, 94-129.
- Belinchón, M., J.M. Igoa, y Á. Rivière (1992). *Psicología del lenguaje. Investigación y teoría*. Valladolid: Trotta.
- Bresnan, J. (1982a). "The passive in Lexical Theory", en J. Bresnan, ed., *The Mental Representation of Grammatical Relation*. Cambridge: MIT press, 3-86.
- Bresnan, J. (1982b). "Control and Complementation", en J. Bresnan, ed., *The Mental Representation of Grammatical Relation*. Cambridge: MIT press, 282-390.
- Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Basil Blackwell.
- Butterworth, B. (1990). "Aportaciones del estudio de las pausas en el habla", en F. Valle et alii, eds., *Lecturas de Psicolingüística I. Comprensión y producción del lenguaje*. Madrid: Alianza, 289-309.
- Caplan, D. (1987). *Introducción a la Neurolingüística y al estudio de los trastornos del lenguaje*. Madrid: Visor.
- Carreiras, M. y E. Meseguer (1999). "Procesamiento de ambigüedades sintácticas", en M. De Vega y F. Cuetos Vega, (eds.), *Psicolingüística del español*. Madrid: Trotta, 163-203.
- Chomsky, N. (1957). *Estructuras sintácticas*. Méjico: Siglo XXI editores.
- Chomsky, N. (1965). *Aspectos de la teoría de la sintaxis*. Barcelona: Gedisa.
- Cuetos Vega, F. (1999). "Neuropsicología cognitiva del lenguaje", en DE VEGA, M. y F. Cuetos Vega (eds.). *Psicolingüística del español*. Madrid: Trotta, 571-596.
- Dalrymple, M. (2001). *Syntax and Semantics. Lexical Functional Grammar*. Londres y San Diego: Academic Press.
- De Vega, M. (1984) *Introducción a la psicología cognitiva*. Madrid: Alianza.
- Gazdar, G., E. Klein, et alii (1985). *Generalised Phrase Structure Grammar*. Oxford: Blackwell y Cambridge University press.
- Grishman, R. (1986). *Introducción a la lingüística computacional*. Madrid: Visor.
- Halvorsen, P.-K. (1988). "Las aplicaciones informáticas de la teoría lingüística", en F. J. Newmeyer, (ed.) *Panorama de la Lingüística moderna de la Universidad de Cambridge. II. Teoría lingüística: extensiones e implicaciones*. Madrid: Visor, 247-271.
- Harding-Esch, E. y P. Riley (1986) (2003). *La familia bilingüe*. Madrid: Cambridge.
- Hernández Sacristán, C. (2002). "Patología verbal: sistémica para una orientación holista en el procesamiento del lenguaje", en Hernández y Serra (eds.). *Estudios de Lingüística Clínica*, vol. 0. Valencia: Universitat de València, 130-175.
- Hualde, J.I. (2003) "El modelo métrico y autosegmental", en P. Prieto (ed.) *Teorías de la entonación*. Barcelona: Ariel. 155-184.
- Hutchins, J. (2003): "Machine Translation: General Overview" en Ruslan Mitkov, (ed.) *The Oxford Handbook of Computational Linguistics*. Nueva York y Oxford: Oxford University Press: 501-511.
- Hutchins, W.J. y H.L. Somers (1992). *Introducción a la traducción automática*. Madrid: Visor.
- Llisterri, J. et alii (2003). "El papel de la lingüística en el desarrollo de las tecnologías del habla" en M. Casas Gómez (dir.) - C. Varo Varo (ed.) *VII Jornadas de Lingüística*. Cádiz. Servicio de Publicaciones de la Universidad de Cádiz, 137-191.
- Llisterri, J. et alii (2004). "La conversión de texto en habla: aspectos lingüísticos" disponible en [http://liceu.uab.cat/~joaquim/publicacions/Llisterri\\_Carbo\\_Machuca\\_Mota\\_Riera\\_Rios\\_04\\_Conversion\\_Texto\\_Habla.pdf](http://liceu.uab.cat/~joaquim/publicacions/Llisterri_Carbo_Machuca_Mota_Riera_Rios_04_Conversion_Texto_Habla.pdf) [fecha acceso 2-7-2008].
- López García-Molins Á. (1991). *Psicolingüística*. Madrid: Síntesis.
- Martí Antonín, M<sup>a</sup>.A. e I. Castellón Masalles (2000). *Lingüística computacional*. Barcelona: Edicions Universitat de Barcelona.
- Meya, M. y W. Huber (1986). *Lingüística computacional*. Barcelona: Teide.

- Mitchell, D.C. (1987). "Lectura y análisis sintáctico" en Francisco Valle et alii, (eds.) *Lecturas de Psicolingüística I. Comprensión y producción del lenguaje*. Madrid: Alianza, 187-210.
- Moreno Sandoval, A. (1998). *Lingüística computacional*. Madrid: Síntesis.
- Owens, R.E. (2003). *Desarrollo del lenguaje*. Madrid: Pearson Educación/Prentice Hall.
- Pollard, C. e I.A. Sag (1987). *Information-based Syntax and Semantics. Volume 1: Fundamentals*. Stanford: CSLI.
- Pollard, C. e I.A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: Chicago University press.
- Reeves, L.M., K. Hirsh-Pasek, y R. Golinkoff (2000). "Palabras y significado: de los elementos simples a la organización compleja", en J. Berko Gleason y N. B. Ratner (eds.) *Psicolingüística*. Madrid: McGraw Hill, 169-244.
- Sag, I.A., T. Wasow y E. Bender (2003). *Syntactic Theory*. Stanford: CSLI.
- Sebastian, N., L. Bosch, y A. Costa (1999). "La percepción del habla" en M. De Vega y F. Cuetos Vega (eds.) *Psicolingüística del español*. Madrid: Trotta, 53-88.
- Sparck Jones, K. (1992). "Natural Language Processing", en W. Bright (ed.) *International Encyclopedia of Linguistics*. Nueva York y Oxford: Oxford University Press: 3: 53-59.
- Tordera Yllescas, J.C. (2008). *Introducción a la Gramática Léxico-Funcional*. Valencia: Universitat de València.
- Tordera Yllescas, J.C. (2011). *Lingüística computacional. Tecnologías del habla*. Valencia: Universitat de València.
- Wingfield, A. y D. Titone (2000). "Procesamiento de oraciones", en J. Berko Gleason y N.B. Ratner (eds.) *Psicolingüística*. Madrid: McGraw Hill, 245-295.
- Yeni-Komshian, G.H. (2000) "Percepción del habla", en J. Berko Gleason y N.B. Ratner (eds.) *Psicolingüística*. Madrid: McGraw Hill, 113-167.