

**LINGÜÍSTICA ESPAÑOLA E INTELIGENCIA ARTIFICIAL: APLICACIÓN INFORMÁTICA DE GRAMÁTICAS DE RESTRICCIONES PARA LA CONFECCIÓN DE AGENTES DE DIÁLOGO**

ÓSCAR GARCÍA MARCHENA  
UNIVERSIDAD DENIS DIDEROT, PARIS VII

**Resumen:** La combinación de los últimos desarrollos en lingüística computacional e inteligencia artificial ha permitido confeccionar aplicaciones web para gestionar el diálogo entre hombre y máquina, los *agentes conversacionales*. Si bien estos agentes existen desde los años 70, la inclusión de un modelo lingüístico formal ha permitido simular la comprensión durante el diálogo con humanos. Este artículo presenta la arquitectura de los agentes de diálogo empleados por los laboratorios *VirtuOz*: una estructura de circuitos para la gestión de diálogos, y un analizador sintáctico-semántico, basado en la implementación de una gramática formal: la Teoría Sentido-Texto.

**Palabras clave:** Inteligencia Artificial, Agentes conversacionales, Teoría sentido - texto.

**Resumen:** Combining computational linguistics and artificial intelligence allows us to build web applications for human-machine dialogue: *chatterboxes*. Already existing in the 70s, they have recently been enriched by formal linguistic models in order to simulate understating in dialogue with humans. This article presents the architecture of the chatterboxes created by *VirtuOz labs.*: a circuit structure for dialogue management, and a syntactic-semantic parser, based in the implementation of a formal grammar: the Meaning – Text Theory.

**Keywords:** Artificial Intelligence, Chatterbox, Meaning – Text Theory.

## **1. Introducción: los agentes de diálogo o *chatbots***

El reciente desarrollo de la lingüística formal ha permitido no sólo la creación de modelos teóricos completos y coherentes sobre el funcionamiento de la lengua, sino también su aplicación en las nuevas tecnologías de la comunicación. Una de estas aplicaciones es la confección de *agentes de diálogo* o *chatboxes*: aplicaciones informáticas provistas de una interfaz humanoide (o *avatar*) capaces de mantener un diálogo con un humano a través de un cuadro de diálogo de estilo *chat* o reconocimiento y síntesis vocal, si bien esta segunda opción está aún en fase de desarrollo.

La integración de conocimientos lingüísticos en la construcción de estas aplicaciones ha permitido mejorar exponencialmente sus capacidades comunicativas. Estos agentes conversacionales, o asistentes virtuales, se emplean cada vez más, instalados en un sitio web, para hacer reservas de trenes, dar información sobre los servicios de una empresa, vender sus productos, etc. Internet, las nuevas posibilidades de acceso a la información, y su creciente volumen requieren técnicas de recuperación de información adaptadas. En este momento de automatización progresiva del acceso a la información, la humanización de estos sistemas de comunicación es la clave de su éxito comercial, y esa humanización viene dada por el modelo lingüístico con el que está construido el agente.

La colaboración entre lingüistas e ingenieros ha permitido así dar un salto cualitativo en el área de la inteligencia artificial. La naturaleza exacta de cómo deba ser esta nueva inteligencia artificial es sin embargo difusa. Bajo el objetivo común de realizar máquinas cuyo comportamiento sea un reflejo del comportamiento racional humano, la posición de ingenieros y lingüistas puede ser diferente; la implementación informática permite simular respuestas inteligentes que dan una impresión de inteligencia. Los lingüistas, sin embargo, nos

encontramos con la posibilidad de poner en prácticas modelos cognitivos humanos, es decir, teorías sobre el funcionamiento del lenguaje en el ser humano. Sólo la combinación de ambos intereses (la cointegración de informaciones lingüísticas y estadísticas, por ejemplo) permite la creación de sistemas de una eficacia aceptable. De este modo, la inteligencia artificial, hoy, no aspira a ser sólo fenomenológica, a ser una simulación de un comportamiento inteligente, sino también a materializar una hipótesis psicolingüística, justificando así su denominación.

Este reciente matrimonio de disciplinas está justificado por las más clásicas discusiones en lingüística, como la tesis Sapir-Whorf sobre la relación entre inteligencia y lenguaje: «¿Hasta qué punto el lenguaje interactúa con o es responsable de o da forma a la inteligencia?» Los agentes de diálogo, en esto, ponen a prueba nuestros límites epistemológicos: consideramos el lenguaje un indicio de inteligencia. Encontrar comportamientos impropios al ser humano en el transcurso de una conversación rebela las deficiencias de su «inteligencia». Hoy, las máquinas han pasado ya este examen, el «test de Turing»: En 2003 uno de estos *chatbots*, ((ro)bots de diálogo), *Jabberwock*, gana el Premio Loebner, por su capacidad para impedir que diez jueces distinguan a un humano de un (ro)bot en el transcurso de una conversación. Hoy día, sin embargo, somos conscientes de que la mayoría de las interacciones humanas no son el resultado de un comportamiento «inteligente», sino imitación, repetición y reacción automática a un estímulo, así como actitudes emocionales sin motivación «racional», por lo que este test ha perdido su validez mítica inicial como *examen de inteligencia humana*. Lo que ha quedado patente es la extensa gama de aplicaciones de estos (ro)bots *inteligentes* a la vida diaria.

En la historia de estos bots podemos distinguir tres etapas evolutivas, con diferentes grados de *inteligencia* según la complejidad de la información lingüística contenida, y la estrategia de conversación empleada. Podemos considerar de primera generación los bots como el mítico *Eliza*, que, con la estrategia de escucha de un psicólogo, tomaba pronombres personales y nombres de las frases del humano y devolvía sus frases hechas cambiando los pronombres de primera a segunda persona y viceversa e incrustando las palabras escritas en frases genéricas adaptables a cualquier situación. Este sistema se reveló útil para lenguas con escasa flexión morfológica, como el inglés, donde el cambio de primera a segunda persona verbal no requería una modificación de la flexión. La programación de una *Eliza* para las lenguas romances era demasiado pesada para sus limitadas posibilidades de aplicación, pero sin embargo se siguen construyendo bots de este tipo con propósitos lúdicos, con estrategias diferentes de escucha, como la confesión religiosa.

La segunda generación de bots conversacionales está representada por *ALICE* y *PANDORA*, y es el resultado de la creación de un lenguaje de marcado específico, el AIML (*Artificial Intelligence Meta Language*), que permite modular su sistema, separando las intervenciones de la máquina de su funcionamiento de selección y gestión de respuestas. Esto permitió la creación de grandes bases de datos con respuestas adaptadas a conversaciones sobre una amplia gama de temas, e incluso las primeras aplicaciones comerciales. La especialización de un robot como informador o vendedor de una actividad determinada acota los tipos de conversaciones posibles, y hace posible una conversación informativamente eficaz.

Los conocimientos lingüísticos de estos robots son sin embargo limitados; si bien aprovechan su potencia computacional para gestionar grandes diccionarios, incluso multilingües, su «comprensión» se reduce al reconocimiento de palabras en una misma frase. A esas palabras hay un paquete de respuestas asociadas en la base de datos, lo suficientemente variadas para no repetirse con frecuencia. No obstante, tienen el mérito de servirse de *expresiones regulares*, que enriquecen la capacidad de comprensión tratando el texto escrito por el usuario como una cadena de caracteres donde las palabras son los segmentos separados por espacios o puntuaciones. Las expresiones regulares, basadas en la sintaxis de Backus-Naur u operadores de lenguajes de script como *Perl*, permiten prever errores ortográficos o incluso simular la reducción gráfica extrema de los lenguajes de sms o chat. Por ejemplo, la frase (1a), grabada en la memoria de reconocimiento del bot como (1b) puede reconocer, entre otras muchas, las posibles intervenciones humanas (1a), (1c), (1d) y (1e) :

(1a) *Hola, ¿qué tal?*

(1b) (h|H)?ola(\s)\* (\i)?((q|Q)u?)|(k|K)(e|é)? ta\*I(\i)\*

- (1c) *hola q tal?*
- (1d) *Hola K tal*
- (1e) *ola ¿qe taal???*

La mayor parte de los agentes de conversación en servicio hoy son de segunda generación; sin embargo, en 2005 la empresa *VirtuOz* comienza a fabricar robots que integran un modelo lingüístico completo, por lo que son considerados de tercera generación. Su sistema se caracteriza por su gran modularidad. La capacidad del lenguaje está gestionada por un sistema diferente de la capacidad de mantener una conversación; en su parte lingüística, sintaxis, semántica, pertenencia de una palabra a una categoría concreta, ortografía, orden de palabras, morfología... todo está gestionado por un módulo independiente que interactúa con los demás. Su mayor logro científico es, sin embargo, el hecho de ser la implementación de una gramática formal teórica completa, la TST o *Teoría Sentido-Texto*.

Su estrategia de conversación es también innovadora, ya que no se limita a gestionar pares de pregunta-respuesta, a reaccionar a cada intervención humana, sino que gestiona modelos de diálogo, en los que se prevén cada respuesta posible del utilizador, pudiendo saltar de uno a otro modelo para comprender cualquier intervención humana, proporcionando una respuesta precisa a cada respuesta posible, e incorporando objetivos a la conversación, como dar una información completa a la primera petición de información. Esto permite que el bot esté entrenado para mantener diálogos propios a oficios especializados, con las estrategias propias de cada actividad. Igualmente, su sistema de « comprensión » permite dar una respuesta pertinente incluso si esa « comprensión » ha sido sólo parcial.

(2) Características de las tres generaciones de *chatbots* :

<i>Ejemplos</i>	Eliza	Pandora, Alice	Equipo VirtuOz
<i>Generación</i>	1	2	3
<i>Modelo lingüístico</i>	Extracción de pronombres y palabras clave ; operadores condicionales ; extracción de listas ;	Bases de datos ; gestión en AIML ; asociación expresión-respuesta ; gran memoria ; incorporación de redes bayesianas, ...	Asociación sentido-textos ; gestión modular en XML ; plataforma privada ; semántica y sintaxis
<i>Estrategia de comunicación</i>	Modelos de escucha (psicólogo, confesor); frases genéricas ;	Especialización de dominio ; interfaces de asistencia ; restricción situacional al objetivo concreto de comunicación ; pares pregunta-repuesta	Conversación dirigida y preprogramada, con objetivos, redes de posibles vías de desarrollo y control del hilo de conversación
<i>Nivel del diálogo</i>	No hay diálogo; sólo escucha.	Diálogo de « sordos » (comprensión de expresiones sueltas)	Diálogo con un aprendiz (posibilidades de interacción limitadas a un oficio)

## 2. Funcionamiento de un agente conversacional de tercera generación

### 2.1. La teoría sentido-texto

El funcionamiento del robot tiene entonces dos partes bien diferenciadas, gestionadas por dos plataformas de desarrollo independientes : un « motor lingüístico » y un « gestor de conversaciones ». La mecánica del motor reposa sobre una teoría formal sobre el lenguaje, *la Teoría Sentido-Texto* (Mel'čuk 1988).

Este modelo parte del concepto de paráfrasis sinonímica: la asociación de un contenido o sentido a diversos *textos* o formas de expresión de una misma idea. La lengua se percibe así como un mecanismo de asociaciones entre un sentido y varios *textos*. La riqueza connotativa es sacrificada por la aplicabilidad computacional, por lo que este modelo ha tenido un éxito particular en su aplicación a los *lenguajes controlados* de la documentación técnica (Nasr : 1996). Esta asociación se lleva a cabo gracias a un modelo formal de transformación (de reducción de) una expresión lingüística a su significado, que pretende ser *universal*; es decir, un *mentalese* que subyace a la expresión de un mensaje en cualquier lengua.

La TST se inscribe así en el marco de las gramáticas formales, que integran datos sobre la semántica, sintaxis de dependencias, morfología y fonética, que interactúan según la estructura detallada en (3a) y (3b) :

- (3a) Semántica <==> sintaxis profunda <==> sintaxis superficial  
 (3b) Morfología profunda <==> morfología superficial <==> fonología

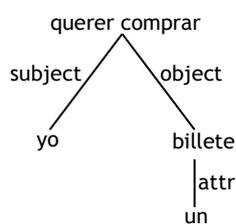
Es decir, que una producción lingüística pasa por varios estados de análisis. Primero, del análisis fonológico se localizan los fonemas que constituyen morfemas, y posteriormente esos morfemas se reducen a su forma « canónica », o flexión menos marcada (lo que equivale a hacer una tripartición del binomio lema – token, correspondiendo el análisis lemativo a la morfología profunda). Igualmente, las relaciones entre palabras se analizan en tres etapas : sintaxis superficial, en la que se perciben las relaciones sintagmáticas de las palabras en el orden en que aparecen, sintaxis profunda, donde las palabras están representadas según sus relaciones de dependencias, siguiendo el modelo de Tesnière (Tesnière : 1959), y finalmente la semántica, donde se incluyen el significado de cada palabra, y el rol semántico de cada relación entre palabras, para así generar una representación completa del significado de una frase.

(4) Niveles de análisis o reescritura en la TST : ejemplos.

<i>ejemplo</i>	<i>nivel de análisis</i>
quiero comprar un billete	Fonología
quiero comprar un billete	Morfología superficial
querer[V : 1ªp, sg, pres, ind] comprar[V : inf] un[Det : num=1] billete[Nc: Gen=com, num=sg]	Morfología profunda
orden lineal (7)	Sintaxis superficial
árbol de dependencias (5)	Sintaxis profunda
grafo semántico (6)	Semántica

El paso entre cada fase de análisis queda asegurado por un mecanismo formal preciso, consistente en sucesivas « reescrituras » en el sentido generativo. Una cadena de caracteres (*morfología superficial*) se reescribe normalizada (con su *morfología profunda*), y posteriormente su *sintaxis superficial* se reescribe en un árbol de dependencias, y éste en un grafo semántico. La TST se inscribe así en la corriente de gramáticas generativas no transformacionales, gramáticas formales, de restricciones o de reescritura, como lo son HPSG (*Head-driven Phrase Structure Grammar*), LFG (*Lexical Functional Grammar*), o TAG (*Tree Adjoint Grammar*). Como ellas, la TST recurre al mecanismo de unificación para asegurar las asociaciones entre las diferentes fases en su implementación. La gramática resultante recibe el nombre de GUST, o *Gramática de Unificación polarizada sentido-texto* (Kahane : 2004).

(5) Sintaxis profunda (árbol de dependencias)



(6) Grafo semántico



Tanto el árbol de dependencias como el árbol semántico se representan mediante grafos, lo que permite poner en común la sintaxis de dependencias con los modos operativos de la teoría de grafos, que es uno de los *modus operandi* de las aplicaciones de inteligencia artificial. Cada palabra está representada por un nodo, y las relaciones entre los nodos por arcos, etiquetados según su función sintáctica o semántica.

El paso de la sintaxis superficial a las sintaxis profunda o de dependencias, la llamada *cartografía*, es el resultado de la aplicación de unas reglas de orden lineal, que serán propios a cada lengua, según el orden de palabras. Los parámetros que se tienen en cuenta son la relación entre dos categorías (*Cat 1* y *Cat 2*), y la función sintáctica y semántica que se establece entre ellos. Dados unos requerimientos respecto a la naturaleza de la categoría dada, se establece un índice de linealidad que marca la posición de un elemento respecto al otro. El orden de palabras se conceptualiza como una línea horizontal en la que la primera categoría (*Cat 1*) ocupa la posición central, y el índice indica la posición de la segunda categoría (*Cat 2*) respecto a la primera. Así, un índice negativo señala que la *Cat 2* se sitúa a la izquierda de *Cat 1* y viceversa.

(7) Orden lineal para « *quiero comprar un billete* »

Función	Cat 1	Cat 2	Lin. Index	Lin 1 Req	Lin 2 Req
sujeto	V	Pron	-3	modetemps = flex	case = nom
OD	V	Nc	+5	---	---
Attr	Nc	Det	-1	---	---

En (7) se indican las reglas de linealidad que rigen el paso de la sintaxis superficial a la sintaxis profunda. Un pronombre que tiene con un verbo una relación de sujeto se sitúa a la izquierda de éste, siempre que el pronombre vaya en caso nominativo y el verbo tenga marcas de flexión (es decir, que no sea un infinitivo). La movilidad de complementos puede indicarse con la presencia de reglas para diferentes posiciones en la frase. Así, para permitir la posibilidad de un sujeto postverbal, basta con añadir una regla en la que el índice de linealidad sea positivo, con una distancia que tenga en cuenta la posible inserción de complementos entre el verbo y el sujeto. Esta computación del orden de palabras, si bien resulta poco visible al lingüista humano, permite describir con precisión un modelo intuitivo en el que el orden de palabras está estrechamente ligado a la estructura de dependencias : en una lengua donde la posición de los complementos goza de cierta movilidad, como es el caso en español, es fácil distinguir cómo los complementos de un núcleo *orbitan* alrededor de él en jerarquías concéntricas. De este modo, resulta sencillo marcar las posiciones posibles de un complemento, y la información añadida que aporta esta posición. Por ejemplo, este sistema es capaz de describir las variaciones de (8a-e), añadiendo a la categoría en una posición marcada un rasgo (p.e. : [*focus = +*]) que indique el aporte potencial de esa posición al significado global de la frase. Sin embargo, en el estado actual de aplicabilidad del modelo, esta información se tiene en cuenta en pocos casos, como la indicación de la alteración de orden en las oraciones interrogativas.

- (8a) *He visto a la chica*
- (8b) *A la chica la he visto*
- (8c) *A la chica la he visto yo*
- (8d) *He visto yo a la chica*
- (8e) *A la chica yo la he visto*

Todo este proceso de reescrituras sucesivas, ejemplificado en (4), requiere ciertos componentes formales para su realización : un diccionario de lemas con su naturaleza categorial y sus flexiones posibles que permita identificar la morfología superficial y pasar a la profunda. Igualmente es necesaria una gramática de construcciones que indique las posibilidades de combinación de las categorías, y sus relaciones de dependencia. Estas descripciones deben ser a la vez sintácticas y semánticas, para permitir el paso de la sintaxis profunda a la semántica. Finalmente, la correspondencia entre la sintaxis profunda y la superficial requiere unas reglas que describan la ordenación lineal de los sintagmas respecto a sus dependencias, es decir, la cartografía proyectada físicamente por el nivel abstracto de dependencias.

## 2.2. Implementación del modelo : *GUST*

La implementación de este modelo va a permitir la creación del motor lingüístico. El objetivo es conseguir un sistema que, una vez integrada toda la información necesaria,

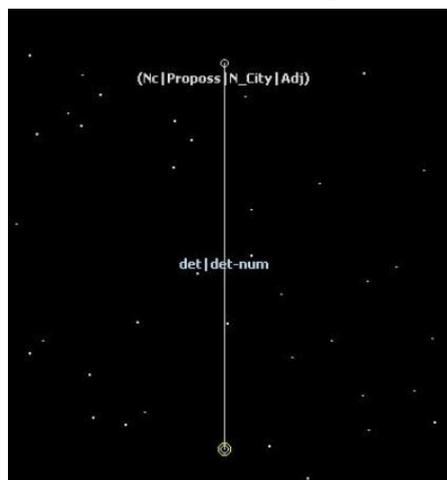
transforme cualquier frase en un grafo semántico. De este modo, obtendremos una representación *universal* del significado de una frase, independientemente del estilo, palabras, sintaxis o incluso lengua en la que esté formulada. Esto permite multiplicar las capacidades de comprensión de la máquina, pues ya no va a comprender una palabra o expresión dada por el conversador humano en función de su forma, sino que captará el sentido de una pluralidad de palabras o frases, independientemente de cómo estén formuladas.

Para las primeras fases de reescritura, el análisis morfológico, se requiere un léxico rico, donde cada entrada contiene información sobre la pertenencia categorial del término, sus rasgos de flexión, y su lema correspondiente. Esto va a permitir separar el lexema de los morfemas flexivos y derivativos, para pasar de una estructura morfológica superficial a una profunda. En cuanto a la estructura sintáctica, la TST, como los modelos de gramáticas formales en los que se inscribe, necesita contar con una gramática de restricciones que indique las construcciones posibles; las combinaciones de las diferentes categorías entre ellas.

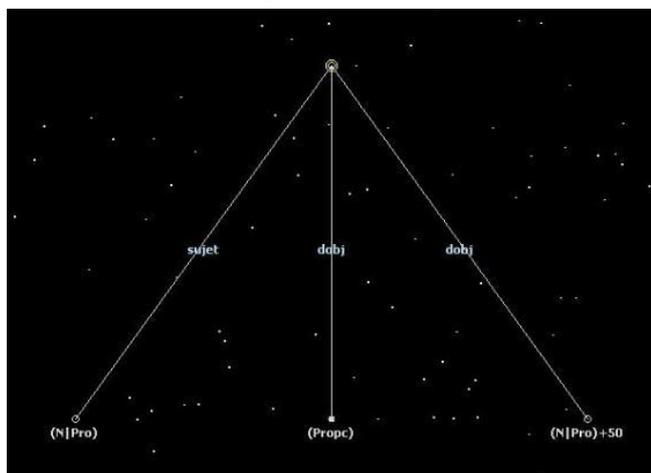
La TST es altamente lexicalista, pues la información sobre las combinaciones gramaticales se sitúa en un grado de abstracción intermedio entre el léxico y la gramática: las categorías. Las categorías, divididas a su vez en diferentes subcategorías, dictan las construcciones en las que se pueden insertar. Así, un determinante es un elemento que se combina con un nombre común mediante una relación de determinación (9); un verbo, por su parte se combina con un nombre o pronombre en una relación de sujeto y, según su estructura actancial, con un número de complementos (10). Según el modelo los sintagmas son entonces proyecciones máximas de una categoría determinada, subcategorizando ciertos elementos de su entorno sintáctico, sus complementos, o pudiendo seleccionar libremente una categoría determinada en el caso de los adjuntos.

Sin embargo, como sucede con otros modelos lexicalistas como HPSG, son numerosos los sintagmas que no son la proyección de una categoría, sino que son categorías sintagmáticas independientes en las que se insertan las categorías precisas. Esto sucede en la TST con los sintagmas subordinados, por ejemplo, en los que una conjunción selecciona el verbo subordinado, pero es seleccionado por otro verbo en una oración matriz, con lo que el sintagma en que se integra no es una proyección suya, sino una combinación de sintagmas.

(9) Estructura de la categoría *Determinante*



(10) Construcción para los verbos monotransitivos



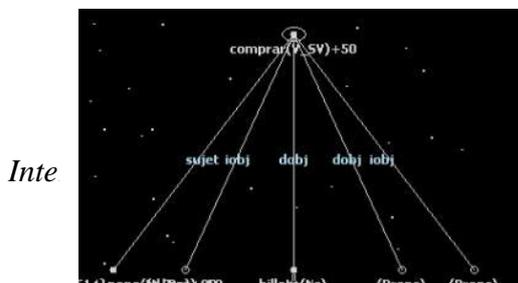
En esta gramática de construcciones (GC) lexicalizada, los verbos tienen un lugar central, al ser ellos el núcleo de la oración al subcategorizar la mayor parte de elementos. La variedad de marcos de subcategorización posibles de un verbo debe estar presente en la gramática, así como los posibles elementos subcategorizados (expresados en (9) y (10) como selecciones disyuntivas «*N / Pro*»). Es necesario asimismo identificar qué complementos tienen una subcategorización *optativa*, es decir, que subcategorizan complementos que pueden estar ausentes sin que varíe el significado del verbo.

En este punto disponemos de un léxico, con indicación de su categoría y morfología, de unas construcciones sintagmáticas de dependencias, que son los contextos sintácticos en los aparecen esas categorías, y de unas reglas de linealidad. Pero necesitamos un mecanismo que inserte las entradas léxicas pertenecientes a una categoría en las construcciones sintagmáticas descritas por estas categorías (léxicas o sintagmáticas). Para esto, la implementación de la TST utiliza una gramática de unificación, que es el mecanismo operacional básico sobre el que se basan las gramáticas formales. Este mecanismo, integrado en la GUST, consiste en añadir una información en los nodos permitiéndoles o prohibiéndoles acoplarse a un nodo preciso de otra construcción. Por ejemplo, en (9) vemos que un determinante selecciona o requiere un nombre, adjetivo o pronombre posesivo. Igualmente, en (10) vemos que un verbo requiere ciertos elementos. En una frase como « *Quiero comprar un billete* », dado que también disponemos de un léxico que nos indica la categoría de cada palabra, y de unas reglas que nos indican qué palabra selecciona cuál en función de su posición lineal (que el determinante es un dependiente de un nombre que se sitúa a su derecha, o que un nombre a la derecha de un verbo mantiene con él una relación de objeto directo), podemos determinar qué categorías se combinan entre sí. El determinante va a seleccionar el nombre, que será a su vez seleccionado por el verbo *comprar* como objeto directo.

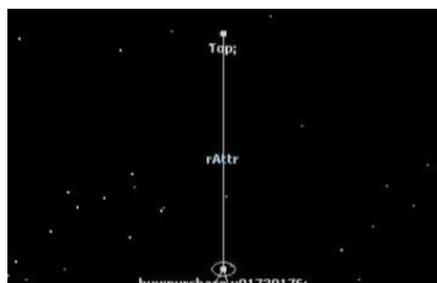
Ahora bien, es necesario indicar qué categorías pueden tomar adjuntos y cuáles no, porque de lo contrario nada impediría que a un nombre ya determinado se le una otro determinante, por ejemplo. La solución es indicar que la categoría que no toma ningún complemento o adjunto está saturada, ya porque ha tomado sus complementos posibles, ya porque no puede tener complementos. Ésta es precisamente la función de la GUST, la de polarizar los nodos, indicando si están saturados o no, permitiendo que un a un nodo no saturado se le una categoría o prohibiendo que lo haga si está saturado. Este mecanismo de unificación asume así una concepción de la capacidad del lenguaje como una capacidad recursiva de combinaciones de sintagmas, lo que ha sido sugerido independientemente por las últimas publicaciones en biolingüística (Balari : 2005). La notación de la saturación de nodos en esta implementación de la TST es visible en (10), donde se describen tres complementos del verbo, un sujeto y un OD realizado por un pronombre (ambos complementos optativos) y un OD realizado por un nombre o pronombre. Esta estructura, que permite asignar grafos sintácticos a estructuras de objeto directo con doblado de clítico acusativo como (8b), (8c) y (8e), y a estructuras sin doblado como (8a) y (8d), contiene un nodo saturado, indicado por un cuadrado blanco : el clítico acusativo, que no permite que se le una ninguna categoría. La polaridad de los nodos de la GUST no es así otra cosa que su saturación o no saturación.

Los pasos de la fonología a la sintaxis profunda han sido ya esbozados; resta indicar cómo se obtiene un grafo semántico (6) a partir de un árbol de dependencias sintácticas (5). Para ello es necesario completar las entradas léxicas con el significado de cada palabra de significado no solamente funcional. Ese significado ha de estar, como el resto de las informaciones lingüísticas que componen el sistema, formalizada. En primer lugar, las relaciones semánticas entre los nodos (en inglés, *agent*, *theme*, *patient* para los complementos seleccionados, y *attribute*, *manner*, *time*, *place* para los adjuntos) están indicadas en el marco de subcategorización del verbo o en la construcción de la categoría. Los arcos que conectan los nodos contienen a la vez una información sintáctica y una semántica, indicando el rol semántico de la relación. Ésta es una de las generalizaciones más problemáticas de la implementación en cuanto a precisión lingüística, pues la asociación entre relación sintáctica y semántica es sistemática : sujeto – *agent*, OD – *theme*, etc., y no se tienen en cuenta otras relaciones posibles, como la de las construcciones inergativas, en las que el sujeto mantiene con el verbo una relación de *patient*. Otros formalismos, como HPSG, no se arriesgan a estas generalizaciones cuya formalización parece aún insuficiente para su implementación, y se limitan a especificar la semántica de una relación-de-sujeto o relación-de-objeto.

(11) Representaciones sintáctica profunda y semántica de « *Quiero comprar un billete de tren* »



X-X.



En cuanto a la formalización de los significados léxicos, es necesario un sistema que permita la homogenización semántica de los sinónimos, y el reconocimiento de hipónimos e hiperónimos, para que a una pregunta del conversador humano como «¿Puedo llevar mi caniche en el tren?» se reconozca «*caniche*» como un hipónimo de perro para dar la respuesta apropiada. Para ello es necesario un léxico formalizado en el que se indiquen las relaciones semánticas jerárquicas se establecen entre las palabras. Para esto se emplea una ontología, concretamente una versión simplificada de *wordnet*, con la que cada significado ocupa un lugar en la red piramidal, indicado por el código que acompaña al significado, que está organizado de más general a más específico.

En resumen, los elementos necesarios para crear un motor de diálogo son, en primer lugar, una **gramática**, descrita como un conjunto de construcciones en las que se integran las diversas categorías gramaticales; en segundo lugar, un **léxico**, con su información morfológica y **categorial**; en tercer lugar, una **ontología**, en la que cada término del léxico ocupa un lugar; finalmente, unas reglas de **linealidad** que aseguran la estructura de dependencias más apropiada.

Esta implementación sin embargo contiene otras herramientas de organización de la información lingüística que, sin ser una parte homogénea del modelo gramatical, son mecanismos imprescindibles que simulan o suplen capacidades intelectivas humanas implicadas en la facultad del lenguaje pero no contempladas en el modelo lingüístico.

En primer lugar nos servimos de un corrector ortográfico, basado en reglas que solucionen errores posibles de taquigrafía del utilizador. El algoritmo permite localizar una letra de más o de menos en una palabra, un carácter que se ha desplazado de su posición en una palabra, o un espacio de más, que corta una palabra, o de menos, que une dos. Este algoritmo aplica esas reglas para verificar si en su base de datos hay alguna palabra que, tras estas modificaciones ortográficas, coincide con la que resulta de alguna de estas operaciones.

En segundo lugar tomamos un *tagger*, que, aplicando el algoritmo de Viterbi y las cadenas de Markov, permite extraer, previo entrenamiento sobre un corpus (semi)manualmente etiquetado, las probabilidades de aparición de una determinada categoría en el discurso, así como las frecuencias de coaparición lineal de trigramas. Esto nos permite determinar la categoría de una palabra desconocida y asignarle consecuentemente una construcción sintagmática. Otros datos que se integran en el sistema son un marcador de sinonimia, para asignar automáticamente su significado a una palabra a partir de un sinónimo, las *colocaciones*, en forma de lemas compuestos, y ciertas informaciones sobre el nivel de lengua, como un

detector de usos vulgares (como rasgo en forma de par atributo-valor codificado en la información léxica)

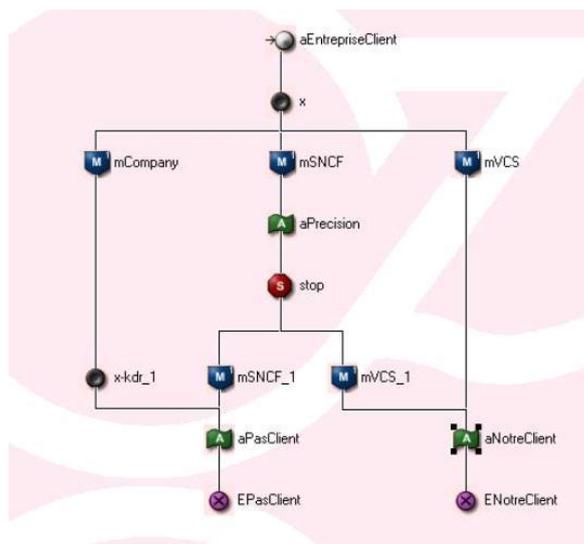
Con todo esto, la confección de un motor lingüístico para el español ha sacado a la luz algunas dificultades para la aplicación de este modelo al español que rebela las deficiencias del sistema de representación arborescente. Este sistema parte del prejuicio que todos los elementos subcategorizados por el verbo aparecen como un nodo seleccionado con el que se establece una relación sintáctico-semántica. Sin embargo, la opcionalidad de expresión del pronombre personal de sujeto plantea un problema: si permitimos su ausencia sintáctica, los grafos semánticos resultantes carecen del elemento que realiza la función semántica del sujeto, aunque éste es siempre recuperable en la morfología verbal. Han sido necesarias modificaciones en el sistema de grafos para resolver este problema. La morfología verbal proporciona el rol semántico, generando siempre un grafo semántico completo, y los pronombres personales carecen de semántica. Sin embargo, de un punto de vista estrictamente lingüístico, esto resulta insuficiente.

Igualmente, otro problema que se plantea es el rol de los clíticos. Si bien se han aportado razones de peso para defender que los clíticos no desempeñan una función sintáctica (Monachesi 2007), el sistema de clíticos del francés (lengua madre de los agentes de tercera generación), donde, a diferencia del español, no es frecuente el doblado de clíticos, permite un análisis de éstos como si fueran complementos subcategorizados. En español, por el contrario, encontramos clíticos que pueden aparecer junto al complemento directo o indirecto. Para resolver esto ha sido necesario recurrir a un sistema de ponderaciones, dando más peso a las estructuras más probables. En (10), por ejemplo, un verbo monotransitivo subcategorizará un sujeto, un pronombre (clítico) acusativo optativo como OD, y un nombre o pronombre (clítico o personal) también como OD, con un peso [+50]. El clítico acusativo está semánticamente vacío, pero no así el segundo. De este modo, de aparecer un nombre y un clítico, el nombre ocupará la tercera posición y el pronombre la segunda, siendo así un clítico reduplicado sin aportación semántica. De aparecer tan sólo un clítico, el peso le inducirá a inertarse en la tercera posición y así tener su aportación semántica.

### 2.3. Gestión del dialogo hombre-máquina

Si bien la implementación que permite la creación de un agente de diálogo de tercera generación consta de una primera parte que puede calificarse de « inteligente », en cuanto a que materializa un modelo cognitivo sobre el funcionamiento del lenguaje (automatizando una representación semántica a cada input lingüístico), esta calificación no están clara para su modulo de gestión de diálogo. El mantenimiento de un diálogo eficiente no es necesariamente indicio de inteligencia (como, por otro lado, hemos visto que tampoco lo es el uso del lenguaje), sino del dominio de una técnica. Un comercial o un informador aprende una técnica conversacional para guiar a su interlocutor, y así se ha procedido para la gestión dialógica: las conversaciones típicas han sido previstas en circuitos de unificación, como (12), en el que se pregunta al interlocutor si es cliente de una compañía de tren, y se prevén respuestas adecuadas a las posibles respuestas.

(12) Circuito de gestión dialógica basado en la unificación



La gran ventaja del sistema consiste en estar basado también en la unificación. Cientos de grafos semánticos están grabados en el sistema (como *M* de *match*, en azul), y asociados a acciones (*A*, en verde), que consisten en respuestas, movimientos eventuales del avatar que representa el agente, navegación a una página web pedida por el usuario, etc. De este modo, la conversación es dirigida mediante unos hilos que toman direcciones determinadas en función de la unificación de la intervención del usuario con el grafo almacenado. La unificación consiste aquí en la correspondencia (*matching*) entre el grafo que está grabado en el circuito (en el bloque *M*), y el grafo generado por la frase del usuario (recordemos que el motor lingüístico ha transformado en grafo semántico la frase del usuario).

Para enriquecer las capacidades comunicativas se permite una unificación parcial. Se hace un cálculo numérico del grado de similitud entre grafo almacenado (asociado a respuesta) y grafo generado, pues así, aunque sólo una parte de la frase sea “comprendida”, se dará una respuesta pertinente. Asimismo, para evitar el bloqueo del sistema cuando no haya *matching* entre grafos porque la intervención del humano no haya sido comprendida, esta intervención «cae» en un *matching* que contiene un grafo vacío (con el que siempre hay unificación), y conduce a un circuito que imita los bots de primera generación, con lo que cualquier enunciado tendrá una respuesta. Igualmente, el *matching* puede prepararse de una manera más estricta, para asegurar que la intervención máquina sea completamente pertinente, combinando a la vez grafo y expresiones regulares. En el sistema de circuitos todos los módulos de conversaciones tipos están interconectados, de modo que un grafo-intervención humana va a encontrar siempre un grafo con el que conseguir el *matching*.

## 2. Conclusiones

Hemos tratado de esbozar aquí los mecanismos lingüísticos necesarios para la fabricación de un *chatbot* o agente de diálogo «inteligente», al menos con una capacidad de comprensión basada en una hipótesis psicolingüística. Sin embargo este sistema puede ser aún ampliamente mejorado. Las mayores deficiencias de este modelo son, como hemos anotado anteriormente, la asunción de que cada categoría léxica está representada por un nodo, cuando éste ni tiene por qué aparecer, ni tiene que aparecer representado por una sola unidad. Asimismo, numerosas mejoras podrían hacerse en la descripción lingüística, sobre todo en la descripción de la estructura actancial de los verbos. Sabemos que la subcategorización verbal no es unívoca, sino variable, lo que se conoce como *alternancias diatéticas*. Para completar la descripción verbal sería necesario agrupar los verbos en cuanto a su capacidad para sufrir una u otra *alternancia* de su marco de subcategorización, e indicar las posibles subcategorizaciones de cada tipo de verbos.

Finalmente, una mejora que podría convertir estos robots en bots de cuarta generación sería la generación de lenguaje. Si en lugar de tener una respuesta asociada a un grafo para responder al usuario, tuviéramos un grafo de respuesta asociado, entonces podríamos transformar ese grafo en una frase de lenguaje natural. Esta generación de lenguaje a partir de una abstracción semántica es posible hoy día, y utilizada, entre otros, por los sistemas de *machine translation* basados en HPSG (Bender et al.: 2003). La asociación de grafos del humano y de la máquina nos permitiría descubrir las relaciones semánticas latentes en el diálogo, que, una vez formalizadas, podrían permitir reducir los dos grafos a uno solo, sobre el que se efectuarían operaciones de transformación del grafo y de adición de arcos y nodos. De este modo, nos libraríamos de la base de datos de respuestas fijas (que hace que el sistema de respuestas del bot sea un modelo no inteligente de estímulo-respuesta), y nos permitiría separar lenguaje y razonamiento.

#### 4. Referencias

- S. BALARI, “Algunas observaciones sobre el lenguaje desde la perspectiva de las ciencias biológicas », en *Revista Española de Lingüística (RSEL)*, 2005, Vol. 35, 2.
- E. BENDER & al, *Syntactic Theory: A formal Introduction*, Stanford, CSLI Publications, 2003.
- O. BONAMI, & al, “Constituency and word order in French subject inversion” en G. Bouma (ed), *Constraints and Resources in Natural Language Syntax and Semantics*, Stanford, CSLI Publications, 1999.
- J. GINZBURG, *A Semantics for Interactions in Dialogue*, Stanford, CSLI Publications, 2001.
- S. KAHANE, *Grammaire d’Unification Sens-Texte - Vers un modèle mathématique articulé de la langue*, Universidad Paris 7. 2002
- S. KAHANE, “Grammaires d’unification polarisées”, en *11ième Conférence annuelle sur le Traitement Automatique des Langues Naturelles (TALN’04)*, Fès, Maroc, France, 2004.
- I. MEL’CUK, *Dependency Syntax : Theory and Practice*. Albany, N.Y., The SUNY Press, 1988.
- R. MONTAGUE, “Universal grammar” en *Theoria*, 36, 373–398. Reprinted in R. Thomason, editor, *Formal Philosophy*, 188–221, New Haven : Yale University Press, 1970.
- P. MONACHESI, « Les pronoms clitiques dans les langues romanes » en D. Godard (ed), *Langues Romanes, problèmes de la phrase simple*. Paris, CNRS Editions, 2007, pp. 67-123.
- A. NASR, *Un modèle de reformulation automatique fondé sur la Théorie Sens Texte: Application aux langues contrôlées*. Tesis Doctoral en informática, Universidad Paris 7, 1996.
- L. TESNIERE, *Éléments de syntaxe structurale*, Klincksieck, Paris, 1959.
- L. TESNIÈRE, “Comment construire une syntaxe” en *Bulletin de la Faculté des Lettres de Strasbourg*, 1934, 7 - 12, 219–229.