

EFFECTOS DEL MODELO LINGÜÍSTICO EN UN PARSER APLICABLE A LA TRADUCCIÓN



Margarita Goded Rambaud
Universidad Autónoma de Madrid

ALGUNOS ANTECEDENTES TEÓRICOS

El objetivo de esta presentación tiene que ver con la idea de que las diferentes maneras de dotar de un formato a los contenidos gramaticales tiene ciertas implicaciones en los diferentes niveles de organización que dan lugar a las conexiones que se establecen en la Traducción.

Tras un repaso de las relaciones que existen entre las diferentes tradiciones cognitivas, las unidades conceptuales que éstas promueven y el abanico de modelos gramaticales que se desarrollan en relación con las primeras, se pasa a comentar un ejemplo de cómo un modelo gramatical que permite la representación de determinados aspectos significativos o pragmáticos, es un instrumento más preciso y adecuado a los propósitos del etiquetado en la preparación de parsers que otros que, aunque su propia configuración binaria ofrecería mejor alternativa para el procesamiento automático, al no incluir una formalización notacional precisa de los aspectos textuales e ideacionales presentan menos facilidades para las aplicaciones de traducción automática.

Como dice Moreno Cabrera (1991): «No es posible hacer descripciones gramaticales neutrales, que no supongan teoría lingüística alguna».

Si se tiene una orientación generativista el concepto de gramática vendrá determinado por el concepto de gramática universal y se tenderá a un enfoque de la lingüística dicotómico. El estudio de la gramática se centrará en la competencia por una parte, mientras que, por otra, el desarrollo de las habilidades se centrará en la actuación.

Es, precisamente, un mérito indiscutible de la Gramática Generativa, como continúa señalando el profesor Cabrera, el haber hecho consciente a la comunidad científica de la conexión entre teoría lingüística y gramática.

Si, por otra parte, se tiene un concepto de la gramática basado en que: «El lenguaje humano es una encrucijada donde se encuentran dos caminos que el hombre recorre a la vez: la categorización conceptual de la realidad y la expresión de esa realidad» (Moreno Cabrera, 1991) la gramática tendrá que tener una orientación que considere que la representación formal de los elementos gramaticales viene, en parte, determinada por esa necesidad de categorización y en parte por su funcionalidad comunicativa.

Las principales aportaciones actuales al problema del concepto y representación de la gramática tienen que ver con dos paradigmas lingüísticos que están vinculados, por una parte a las dos tradiciones filosóficas básicas que en Occidente se han referido al conocimiento, y, por otra, a los diferentes modos de conceptualizar la realidad.

La explicitación de esta vinculación tiene que ver a su vez con la explicitación general del marco teórico en que nos desenvolvemos. Por una parte tiene que ver con la idea de lenguaje, bien distinta en ambos paradigmas; y por otra con la idea de cómo se adquiere esta facultad, igualmente diferente en los dos.

El racionalismo idealista, como ya se ha mencionado, se remonta a la tradición platónica que sostiene que el conocimiento es siempre la proyección de nuestras ideas innatas

sobre nuestra conciencia. Nosotros no podemos conocer, encadenados como estamos a nuestros sentidos, sino que gracias a las ideas puras (e.g. ideas innatas o en versión «moderna» programa biológico innato), podemos llevar a nuestra conciencia aquellas ideas que desde siempre estaban en el alma.

Las ideas innatas, ya sean los conceptos geométricos del esclavo de Platón o la gramática universal (sea en versión GB o minimalista), son conceptos lógicos cerrados y bien definidos.

Este paradigma explica que el lenguaje se adquiere porque existe esa predisposición genética a adquirirlo. Lo cual es lo mismo que decir que, para poder obtener energía, los humanos procesamos el oxígeno en vez del CO₂ porque tenemos pulmones y plaquetas en vez de tener clorofila y cloroplastos. Es decir que, como viene a afirmar Jackendoff, el órgano crea la función.

No obstante el racionalismo tiene en el desarrollo del método de análisis científico su mejor y mas creativa aportación. Sistematiza la capacidad que tienen las representaciones ideales de la realidad como modelos abstractos con los que se puede operar; y esta contribución es decisiva para el progreso de las matemáticas, la lógica y la física como lo prueba el desarrollo de estas disciplinas a partir de Descartes, Leibnitz, Kant y Newton. Sin embargo, como decimos, su mejor aportación es fundamentalmente metodológica.

Por su parte el empirismo, es otra tradición cognitiva procedente del mejor discípulo de Platón, Aristóteles, que sosteniendo que el conocimiento procede de los sentidos que dotan a la mente de imágenes que se asocian mediante las leyes de contigüidad, similitud y contraste, inicia la tradición del pensamiento empírico. Este tendrá su mejor momento dentro de la filosofía con Hobbes, Locke y Hume y de alguna manera sienta las bases del desarrollo científico posterior. El empirismo como tradición cognitiva se apoya en la idea de la necesidad de conocer los datos observables para elaborar el conocimiento teórico a partir de ellos y exige que cualquier hipótesis teórica sea contrastada de algún modo con los datos de la realidad que pretende describir.

En el terreno de la psicología el empirismo está relacionado con el asociacionismo y éste, a su vez, con el conductismo. De ahí que la contestación chomskiana al conductismo sea, no sólo una afirmación de la capacidad creativa del lenguaje, sino, a su vez, una profesión de fe racionalista.

La tradición empírica clásica acentúa la distinción entre dos tipos de proposiciones. De una proposición que enuncie una relación entre ideas se dice que es analítica y verdadera «a priori» (si es verdadera). Es decir su verdad es independiente de su verificación empírica.

Por otra parte, de una proposición que se refiera a cuestiones de hecho se dice que es sintética. Su verdad no puede ser conocida sin mas elemento que la proposición misma, sino que necesita verificación empírica, necesita un contraste con la realidad. Y es la verificación empírica la que demuestra si la proposición es verdadera o falsa.

Es decir el empirismo propondría un análisis conceptual en dos planos. Uno, que podríamos llamar sintáctico, en el cual no se cuestiona la verdad o falsedad de sus elementos sino que sólo se enuncia la relación entre sus componentes. Y otro que debe ser verificado y refrendado por la experiencia.

Se podría pues afirmar que la formalización lingüística no puede depender únicamente de un análisis de conceptos y de sus relaciones, sino que necesita el contraste constante con los demás componentes del análisis lingüístico —aspectos semánticos y pragmáticos— para no generar proposiciones lógicamente imposibles. Y, obviamente, necesita un contraste con la realidad: si las formalizaciones propuestas no dan razón del funciona-

miento lingüístico más que por recurso al propio sistema, no son explicaciones científicas sino escolásticas.

Por ello es necesario, en primer lugar, reflexionar acerca de que tipo de conceptos son los conceptos lingüísticos. En algunos casos son conceptos clásicos, bien definidos y delimitados; pero en otros son conceptos más difusos, conceptos probabilísticos cuya adquisición puede ajustarse bien al modelo prototípico o al modelo ejemplar (Langaker, 1987).

Y esto tiene evidentes consecuencias en el tipo de formalización. Pero en cualquier tipo de formalización. Los modelos cognitivos funcionales tampoco escapan a este problema. En realidad formalizar es etiquetar la realidad y ésta es escurridiza y no siempre se deja. En el ámbito de la Física o de la Biología hay algunos ejemplos ilustrativos. La luz, por ejemplo, a veces tiene un comportamiento que se ajusta a descripciones formales que la categorizan como fuente de energía y así es capaz de romper determinadas reacciones químicas mientras que otras veces su comportamiento se describe más fácilmente como partícula pesable y medible. O la capacidad de determinadas cadenas moleculares para desdoblarse que hace difícil su categorización como organismos inertes y más difícil aún el trazado de la línea divisoria entre organismos vivos y materias inertes...

¿Cómo entonces no va a ser un problema la categorización y el etiquetado de unidades lingüísticas? El problema está en considerar las formalizaciones como instrumentos conceptuales que nos permitan una mejor interpretación y descripción de las lenguas no como explicaciones absolutas. Los conceptos los crea el hombre y son necesarios en la medida en que permiten la elaboración de teorías que permiten explicar fenómenos o situaciones pero cuando dejan de ser válidos son tan dispensables como los Kleenex.

Parece pues que, con todas las matizaciones necesarias, los modelos lingüísticos actuales se organizan bien en torno al paradigma funcional o bien entorno al paradigma formal. Aun cuando parece que empiezan a surgir voces a favor de una supuesta integración, ésta no parece fácil. Y no es fácil no porque partan de unos supuestos filosóficos muy distintos, que también, sino porque sus objetivos son también distintos.

Chomsky, desde su posición mentalista racionalista extrema propone un concepto de gramática constituido por la formalización de las ideas innatas que se configuran en los llamados universales lingüísticos. El lenguaje es descrito como una capacidad psicológica / biológica humana o competencia que implica un conocimiento innato de una serie de reglas que constituyen la GU. Incluso en su última formalización minimalista continúa estableciendo una línea divisoria entre el conocimiento del lenguaje y su uso. La gramática sería la representación formal de las reglas que describen esa competencia, ese conocimiento, mientras que la actuación, totalmente independiente de aquella, se refiere a la actividad concreta de los hablantes. Ahora bien, o con Chomsky o contra Chomsky lo que es justo reconocer es que no es posible entender la lingüística actual sin su referencia.

Pero no es sólo esto. Chomsky ha tenido el valor y el mérito de desmontar, en cierto modo, aquella parte de su propia teoría que se estaba convirtiendo en un obstáculo. Las reglas transformacionales constituían un aparato formal casi burocrático con una finalidad más de automantenimiento que de descripción del funcionamiento interno del lenguaje. En el modelo GB del 86 se deshace de ellas y propone un solo movimiento transformacional (move- α) de un nivel de abstracción mayor. Además demuestra una mayor flexibilidad y permeabilidad a las aportaciones teóricas críticas que la que muestran sus seguidores e incorpora el significado a un rango de operatividad superior. Como siempre pasa, los creadores de las teorías suelen ser menos dogmáticos que los miembros del aparato que sus ideas generan.

Son muchos los méritos que cabe atribuir a Chomsky y no es el menor el de haber impulsado con sus ricas, polémicas y provocadoras propuestas el debate acerca de cómo se

adquiere una lengua. Su desafío ha valido para impulsar el desarrollo de las teorías de adquisición lingüística.

Además hay aspectos del paradigma formal que hoy se reconocen como aportaciones netas al avance de la lingüística. El hecho de haber acotado el campo de la sintaxis como un terreno de estudio específico con leyes propias supuso una contribución valiosa a la consideración de la gramática como ciencia. Posteriormente con la formulación de la GB, el concepto de mando-c y el de proyección máxima dentro del marco de la reacción son conceptos sólidos e incorporados a la lingüística dentro de la sintaxis.

La primera, que es una especie de regla que define las relaciones de dependencia de los constituyentes sintagmáticos, es el concepto estrella de teoría de la $x /$. Establece la condición en que un elemento manda-c a otro.

Por su parte la idea de proyección máxima establece el grado de alcance máximo al que puede afectar una relación de dependencia entre los constituyentes sintagmáticos. Ambos conceptos tienen que ver con la idea de que el lenguaje como sistema estructurado y organizado que es, es necesariamente jerárquico y vinculante. Estas ideas, como concluía Moreno Cabrera, han pasado a formar parte del patrimonio conceptual mas comúnmente aceptado dentro de la lingüística actual.

Sin embargo el generativismo ha necesitado incorporar a un primer rango de reglas el componente semántico (en la teoría del ligamento: módulo temático o: «cómo las exigencias léxicas de las palabras presentes en la oración inciden en la configuración sintáctica de ésta» (Demonte, 1989), después de haber sostenido durante décadas que el componente sintáctico era primordial y generador de los demás niveles porque era mejor asimilarlo que enfrentarse a él.

El componente semántico se había relegado porque apuntaba a una mayor valoración de los aspectos comunicativos. Si los aspectos significativos son determinantes de los organizativos es que el lenguaje tiene una finalidad, que no es un sistema primordial sino un sistema al servicio de algo. Que la función, la necesidad, crea el órgano; aunque luego el órgano tenga sus propias reglas de organización interna.

Para evitar ese tipo de conclusiones se incorpora el componente semántico como parte de la articulación descriptiva general y eso permite que pueda tener un rango de actuación superior pero sin salirse del marco de organización interna. Es decir salvando el aparato descriptivo.

El programa minimalista (Chomsky, 1992) es toda una pirueta intelectual para seguir sosteniendo postulados intensamente racionalistas como la especie de 'a priori' kantiano que es la GU. Se mantiene como una teoría del lenguaje y de las expresiones que éste genera y como una teoría del estado inicial de los componentes relevantes de la facultad del lenguaje. Incorpora la Pragmática al rango de la competencia y aunque la diferencia del lenguaje —no se entiende muy bien cómo esto es posible— al menos admite que exista un sistema de competencia pragmática. Por primera vez reconoce que tiene que haber una relación entre las propiedades del lenguaje y su uso y a este uso lo denomina *performace systems* y éstos son de dos tipos: perceptivo-articulatorios y conceptuales-intencionales. Esta relación o interface entre la forma fonética y la forma lógica que recuerda al viejo signo lingüístico saussuriano queda articulada de una forma que no establece con claridad la posible relación que apunta entre competencia y actuación.

Y esto es así porque el generativismo insiste en definiciones de la lengua y / o el lenguaje como una capacidad cognitiva sin más en la que se evita cuidadosamente decir *para qué* se tiene esa capacidad. Y lo es porque en el momento en que al concepto de lenguaje incorpore la idea de la finalidad del mismo se desmorona todo el edificio construi-

do a partir de la idea de que el lenguaje es un a priori cognitivo, de que la competencia innata puede quedar formulada en términos de una GU cuya existencia es previa e independiente de las razones por las cuales se ponga en práctica dicha capacidad.

Sin necesidad de perder de vista el indudable valor del lenguaje como instrumento de conceptualización no se puede obviar que el diseño u organización interna de los sistemas tiene que ver con la finalidad a la cual sirven. Sin embargo en toda la literatura generativista se insiste en definiciones del lenguaje que ponen el acento en su evidente característica de ser una capacidad exclusivamente humana, pero se procura evitar mencionar que las capacidades son capacidades para hacer algo; y ese algo en el caso del lenguaje es fundamentalmente transmitir información, es decir comunicarse. Pero la palabra comunicación es poco menos que anatema en el entorno generativista.

Una muestra de esta actitud puede observarse en algunos autores que analizan el problema de la conexión entre la cognición y la sintaxis desde la perspectiva generativa (Fernández Igoa, 1995) donde se propone un mecanismo que, autoproclamándose «lingüísticamente neutral» (sic), presenta una secuencia que sigue relegando la significación a un nivel muy bajo en la representación del procesamiento lingüístico. En su propuesta dice que:

El etiquetado sintáctico de los constituyentes exige el reconocimiento de la categoría gramatical de las piezas léxicas que actúan como núcleo de cada sintagma, lo que a su vez permite identificar el papel que desempeñan éstos en la estructura de la oración y *con ello la interpretación temática que han de recibir.*

en ese momento a este marcador sintagmático o representación de salida del analizador sintáctico o parser se unen las piezas léxicas.

Es decir, que el desarrollo de la interpretación cognitiva generativa de la incorporación del significado a un nivel de representación mayor continúa a la hora de la verdad en su tradicional situación de sometimiento a la sintaxis. Lo cual no tendría nada de particular y podría admitirse desde otras posturas si el modelo de representación lingüística y el modelo de procesamiento lingüístico no pretendieran un isomorfismo; lo cual no es el caso. Es decir, si se admitiera que la organización lingüística, y como paradigma de ella la sintaxis, es una necesidad que permite automatizar el procesamiento y concentrar la atención en la comprensión del mensaje, en agilizar la comunicación. Pero si se sigue pretendiendo que el modelo gramatical es una suerte de representación de las capacidades psicológicas cognitivas donde los mecanismos de procesamiento mandan sobre finalidad del propio procesamiento entonces el modelo vuelve a ser un núcleo descriptivo de la organización sintáctica —con sus indudables valiosas aportaciones— al que se le van incorporando elementos significativos y pragmáticos, no importa a que nivel, pues a la hora de proponer una aplicación del modelo siempre aparecen en posición de subordinación.

De este modo, desde el punto de vista de lo que sería una propuesta que demostrara que la sintaxis es una forma de representación de la cognición hay que decir que el trabajo de Fernández Lagunilla y otros (1995) casi viene a confirmar lo contrario a lo que se propone.

Las distintas gramáticas englobadas en el llamado paradigma formal comparten estas premisas en mayor o menor medida. Hay posiciones más extremistas como la de Katz (1981) que llega a defender platónicamente que el lenguaje existe con independencia a las mentes de los hablantes. Y también hay posturas menos radicales como la de la Gramática Léxico Funcional de Bresnan y Kaplan que atribuyen un papel central al léxico, lo cual les obliga a tener en cuenta el significado.

Mientras que el enfoque o paradigma formal se centra en las características formales internas del sistema, en el enfoque funcional las formas y relaciones que operan dentro del sistema «se explican en función de cómo se significa por medio de éste» (Mairal, 1995).

Ha sido la influencia de la investigación acerca del procesamiento humano de la información y del aprendizaje lo que ha dado lugar a modelos lingüísticos de orientación cognitivista donde se pretende dar cuenta de los procesos además de representar los productos en sí mismos.

Así a las tendencias funcionalistas se han incorporado recientemente algunos planteamientos cognitivos como en Givón (1984, 1989).

Por otra parte debe mencionarse también cómo la evolución de los modelos cognitivos en psicología ha seguido una evolución similar al desarrollo de los modelos lingüísticos en el sentido de la progresiva incorporación de elementos significativos en la descripción lingüística.

Las teorías cognitivas del aprendizaje ofrecen distintas alternativas aplicables a la caracterización de los conceptos lingüísticos y de su uso, incorporando conceptos de clara aplicación lingüística como la proceduralización, la metacognición, la rutinización de destrezas o su automatización o la funcionalidad o relevancia o interés de la pieza de información o concepto a conocer. Sin embargo, sólo en la medida en que el concepto de lenguaje tenga como rasgo conceptual central el de comunicación y se considere como un instrumento de procesamiento humano de la información, las teorías cognitivas del aprendizaje mencionadas podrán aportar alguna explicación a la caracterización de los conceptos gramaticales, su cambio, uso y aprendizaje. Lakoff (1982, 1987) y Langaker (1987) son los iniciadores más representativos de esta orientación.

Podríamos resumir diciendo que las dos grandes líneas de trabajo en la lingüística actual aceptan como presupuestos básicos que el lenguaje tiene una organización interna sistemática aunque difieren en el sentido de esa organización. Mientras que para los generativistas, esa organización cuyo núcleo generador es la sintaxis, es independiente del sistema cognitivo general del ser humano, además es universal y es el reflejo de una dotación innata, para la línea funcionalista-cognitivista esa organización está en función de la transmisión de la información y forma parte de los sistemas generales de cognición de los seres humanos.

Asimismo ambos paradigmas difieren en la consideración o definición del lenguaje. Para el funcionalismo en todas sus versiones de las varias finalidades del lenguaje, la más primaria es la comunicación, aunque, obviamente, también reconocen la importancia de las demás funciones.

Por su parte, para las propuestas formales el rasgo diferencial más importante no es que consideren que la finalidad prioritaria sea la conceptualización y sistematización del conocimiento, que son posibles gracias al lenguaje y profundicen desde esa óptica, sino que excluyen las demás funciones; especialmente la comunicativa.

Es muy difícil entonces estudiar el comportamiento de un sistema desgajándolo de una de sus funciones fundamentales. ¿comprendería bien un extraterrestre, de anatomía distinta a la humana, el funcionamiento de un coche o una bicicleta en un mundo desprovisto de bípedos usuarios de medios de transporte más eficaces que sus dos piernas?

CONCEPTO DE GRAMÁTICA. UNIDADES DE ANÁLISIS

Si el lenguaje es una entidad organizada constituida por elementos identificables, la gramática, cualquiera de ellas, tratará de definir sus unidades y describir y explicar su funcionamiento.

El problema está en que cualquier modelo de formalización lingüística necesita operar con conceptos que han tenido que ser previamente definidos. Esa definición implica a su vez una opción o preferencia por un tipo u otro de conceptualización. ¿Se va a operar fundamentalmente con conceptos clásicos, bien definidos y delimitados o se va a operar con conceptos prototípicos y / o ejemplares? El tipo de elección va a forzar un tipo u otro de formalización y por tanto de modelo lingüístico. Aquí de nuevo no sólo el tipo de modelo lingüístico sino más concretamente el tipo de notación está en relación con el tipo de conceptualización.

Cuando se opera con conceptos clásicos de definición clara y precisa, donde la cuestión de pertenencia o exclusión de un conjunto es una cuestión de sí o no, es decir, claramente, una respuesta binaria, la articulación de una representación del tipo de los diagramas arbóreos parece adecuada ya que en cada encrucijada de un diagrama arbóreo de la notación generativa solo es posible una opción o su contraria, y si no se retrocede a la encrucijada anterior para dar salida a otra alternativa. Esto quiere decir que la propia notación generativa favorece un procesamiento binario. De ahí su elegancia.

Sin embargo cuando la pertenencia a un conjunto, nivel de representación o capa es una cuestión de grado, una notación de tipo aglutinante permite simultanear una presencia en distintos ámbitos. Un mismo concepto, un mismo término dentro del modelo de Dik por ejemplo, puede adoptar una posición de argumento o de satélite dependiendo de las funciones semánticas que determinen uno u otro. Así los marcos predicativos del modelo de Dik podrían asimilarse desde el punto de vista cognitivo a los esquemas tal como los define Norman (1986) que «no son estructuras fijas, sino que son configuraciones flexibles, que reflejan regularidades de la experiencia, proporcionando una terminación automática de los componentes ausentes» ... «son estados interpretativos flexibles que reflejan la combinación de la experiencia pasada y las circunstancias presentes».

Por ello es pues conveniente plantearse con qué tipo de conceptos debe operar la Gramática como representación formal de una lengua. Pero esto dependerá, a su vez, del concepto de lenguaje que se tenga. Los conceptos gramaticales no tienen por qué ser únicamente conceptos científicos. Pueden perfectamente coexistir con lo que Vigotskii llama pseudoconceptos, o Rostch conceptos probabilísticos, ya sean ejemplares o prototípicos.

No sólo es necesario plantearse con qué tipo de conceptos debe operar la Gramática sino también qué tipo de explicaciones pueden esperarse de este tipo de formalización. Asimismo cabe plantearse por tanto qué opción dará mejor cuenta de los hechos lingüísticos; si las reglas o las listas de ocurrencias. Langaker (1987) opina que ambas posibilidades pueden coexistir para dar cuenta de los fenómenos lingüísticos.

Cuando finalmente se puede proponer un análisis de las unidades conceptuales que se pueden diferenciar en las diferentes capas que configuran la organización lingüística se pueden identificar algunas de contornos bien definidos como los sonidos, los fonemas y sus alófonos, las distintas agrupaciones en sílabas y finalmente la palabra como unidad de una categoría superior al incorporar el significado. Posteriormente se pueden identificar otro tipo de unidades con una cierta cohesión sintagmática pero con menor contenido semántico independiente como la frase nominal o verbal. A un rango en el que se puede ya codificar ideas aparece como unidad de análisis la cláusula u oración con diferentes niveles de complejidad. Hasta llegar al texto, como objeto de estudio.

Pues bien, este estudio sistemático del lenguaje se ha ido fragmentando y especializando desde el punto de vista epistemológico en disciplinas y subdisciplinas que se articulan en torno a diversas unidades de análisis. Así, en los niveles fonético y fonológico las unidades de análisis están bien delimitadas y codifican unidades significativas distintivas;

permiten acotar un campo de estudio de manera clara. Igual sucede con el nivel morfológico y léxico donde las palabras y sus subunidades pueden codificar conceptos y relaciones. Y, del mismo modo, el nivel sintáctico puede codificar la idea mediante la cláusula como unidad de análisis. Givón (1993) dice:

Words code concepts that have meaning. Grammar-clad clauses code propositions that convey information.

Sin embargo, esta posibilidad de codificación de unidades más o menos claramente definidas no está tan clara en el caso del texto. Si bien Halliday (1985) define explícitamente el texto como una unidad semántica y pragmática, el texto, al no poder subdividirse a su vez en subunidades significativas menores, no puede representarse de modo que permita su manipulación conceptual.

Bernárdez (1995) ilustra esta idea diciendo que el texto solo puede existir en el uso.

Lo que nos interesa aquí no es tanto que el significado potencial de una oración se realice en contextos determinados ...; nos interesa más que la oración pueda existir en tanto en cuanto objeto abstracto, puramente potencial, y nos interesa precisamente porque en el texto eso no es posible: no existen textos «abstractos».

Así, como también dice Esa Itkonen (1993: 181. cf. Bernárdez), los discursos siempre se enuncian. Desde esta perspectiva entonces, no podemos identificar el texto como unidad de análisis para su posible representación y manipulación teórica aunque podamos identificarlo como entidad en la cual se pueden abordar distintos aspectos de estudio.

Nos encontramos pues con el problema de que las distintas capas organizativas pueden articularse en torno a diferentes tipos de unidades que son susceptibles de definición como tales unidades conceptuales hasta llegar al texto donde si bien efectivamente puede calificarse de unidad significativa no se puede ir mucho más lejos ya que ya que no se pueden identificar diferentes elementos estructurales como sucede en el caso de la oración, la frase, la palabra, la sílaba, el fonema o el alófono. Es como si en realidad el texto fuera algo más parecido a una unidad cognitiva de tipo de los esquemas de Norman o modelos mentales de Johnson-Laird que se construyeran con ocasión de cada uso concreto. Por ello los tipos de representación notacional aglutinante y flexible parecen apropiados a la hora de necesitar representar textos que es en definitiva lo que constituye el objetivo de análisis en la traducción.

A la hora de intentar aplicaciones de traducción automática a problemas concretos puede mencionarse como existe un primer grupo de programas de traducción automática de tipo directo y otro segundo grupo de tipo indirecto basados bien en la interlengua, bien en la transferencia. Pues bien los basados en la interlengua se asientan en la idea de que la predicación subyacente es en sí misma una descripción de la interlengua.

ESTUDIO DE CASOS

Puede comentarse a modo de ejemplo cómo se puede instrumentalizar la aplicación de algunas representaciones para que reflejen con más precisión determinados elementos que deben ser incluidos en el etiquetado previo de cualquier programa de traducción asistida por ordenador.

Como es sabido, en los modelos de representación tipo FG se parte de la configuración semántica para llegar a la configuración sintáctica. Su nivel central de descripción es la predicación que es en sí misma un esquema cognitivo.

Por otra parte, la GB, donde las entradas léxicas dan lugar a las entradas temáticas, es decir, se reducen porque se anula la *Lexical Conceptual Structure*, es decir, la definición, tiene su mejor valor en la precisión y elegancia con la que refleja el comportamiento de los constituyentes sintagmáticos.

Asimismo, mientras los programas que implementan modelos funcionales son de base lógica como el PROLOG y los que implementan parsers de tipo generativo son en general codificados en LISP (*Linguistic String Parser*).

Se trata pues de resolver el problema de la interpretación y etiquetado previo a la traducción de una frase que pertenece a un texto especializado acerca de la restauración de antigüedades. Se refiere a la operación de restituir una placa de material decorativo — madera, carey, metal u otros— a la parte del mueble de la que se había separado previamente para su restauración o limpieza. El contexto es pues de una inmediatez plana y deja poco margen a interpretaciones referenciales. Es por ello que el análisis del nivel sintáctico adquiere mayor relevancia que el pragmático y por lo que, inicialmente, una propuesta de parser apoyado en modelos tradicionales como el GB debería ser suficiente.

La frase que se va a comentar es una frase nominal con dos preposicionales anexas a ella.

Fijación de la superficie decorada al soporte.

Que se puede reescribir también alterando el orden de las frases preposicionales:

- (1) Fijación de la superficie decorada al soporte
- (2) Fijación al soporte de la superficie decorada

Para ella se pueden proponer dos posibles traducciones de las cuales parece mas natural la segunda de ellas.

Fixing of decorated surfaces into the base / stand / receiving area
Decorated surfaces fixing into the base / stand / receiving area

Obsérvese en primer lugar las representaciones formales de la frase en cuestión tanto en la descripción FG como en la GB.

En el modelo funcional podemos considerar *Fijación* como una entrada léxica con tres argumentos:

$$\begin{aligned} & \text{Fijación } N (X_1) (X_2) (X_3) \\ & \text{df} = (X_1) \text{ Ag } (X_2) (X_3) \end{aligned}$$

A su vez el marco predicativo de (1) podría representarse del siguiente modo:

d1 xi: Fijación N (x i)Φ(go): ((d1 xj: superficie(xj) decorada (xj)) (ag) (rec)): xi (d1 xk: soporte(xk) (xi)loc

que puede ser parafraseada en los siguientes términos:

Entidad definida singular (*Fijación*) tal que es un nombre tal que se aplica a entidad singular definida que es un nombre (*superficie*) tal que tiene la característica de que está decorada tal que se aplica a la entidad definida singular (*soporte*)

Asimismo el marco predicativo de (2)

d1 x: fijación N (x) φ: (d1 x: soporte N (x j)) loc (xi): (d1 X k: superficie N: decorada A (X k)) pos (X i)

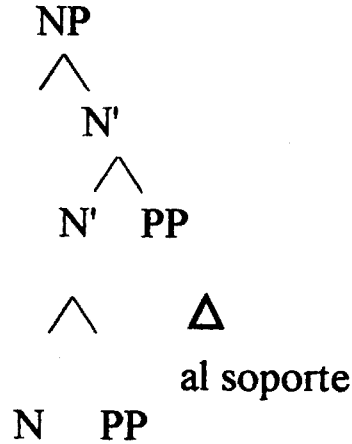
y su traducción:

Fixing of decorated surfaces into receiving area / Decorated surfaces fixing into the receiving area

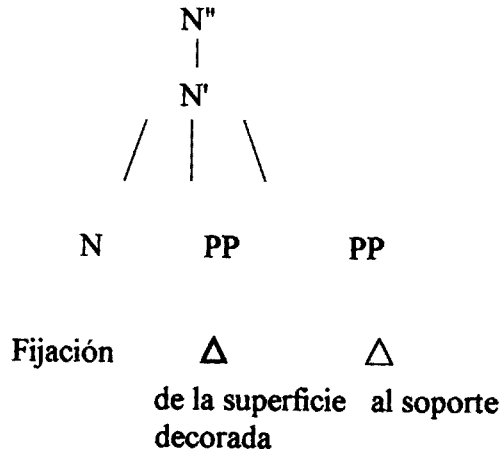
d1 x_i: Fixing N (x_i) Φ (ag) (go): ((d₁ x_j: surface(x_j) decorated (x_i)) (rec)): x_i (d1 x_k receiving area (x_k)) (x_i) loc

El contenido de acción de Fixing lleva anexionada la anotación de (Ag)

Por otra parte la representación GB ofrecería la siguiente configuración en la que ha habido una transformación pasiva previa:



o bien



Hay varios aspectos que se pueden comentar en las representaciones FG y GB respectivamente. En primer lugar las características semánticas de la palabra fijación. En segundo lugar, los efectos del tamaño de frase preposicional y cómo se puede dar mejor

cuenta de ellos y, finalmente, y en esa misma línea, cómo se pueden formalizar los aspectos pragmáticos que pueden tener consecuencias en el orden de palabras o secuenciación de elementos.

El sustantivo fijación (*fixation*) tiene en primer lugar el problema de su categorización como nombre o como nombre verbal. Desde el punto de vista sintáctico es claramente un nombre al admitir pre y postmodificación como cabeza del grupo nominal. Semánticamente se refiere a una acción y sin embargo no aparece ningún participante activo. Son pues las características de esta acción las que interesa analizar aquí (fijación de algo a alguna parte) y son estas características de las que un modelo como la FG de Dik puede dar cuenta. Mientras que la representación en GB simplemente da cuenta de esta carencia de participante activo diciendo que hay una transformación pasiva previa en GF no es preciso representar ese movimiento pues la pasivización queda ya codificada en las características del marco predicativo.

Igualmente el grado de prototipicidad de la palabra *fijación* varía en cada caso. En ambas ocurrencias lleva anexionada una frase preposicional «de la superficie decorada» y «al soporte» pero en fijación al soporte tenemos un nombre menos «nominal», «a less nounish noun» —(Lakoff, 1987), mientras que en fijación de la superficie decorada el mismo nombre *fijación* es más prototípico o más «nounish».

Asimismo debe tenerse en cuenta el tamaño de la frase preposicional donde, en general, la frase más larga queda más naturalmente situada al final. Así podemos comparar las dos variaciones españolas con otro ejemplo similar:

Fijación del clavo a la pared / Fijación a la pared del clavo

donde, en el primer caso, la frase (del clavo) queda más naturalmente vinculada al nombre verbal *fijación* y *del clavo* queda tematizada siguiendo la secuencia natural de fijar algo a alguna parte. Si alteramos esta secuencia podemos observar cómo se sitúa un énfasis en el satélite *del clavo* de modo que la intencionalidad que determinaría un orden de palabras es tenida en cuenta en un tipo de representación con más capacidad de discriminación de este tipo de rasgos conceptuales que en otra en la que siendo ambos elementos de similar longitud y categorización sintáctica ello no sería posible.

Finalmente si comparamos entre sí las dos alternativas españolas podemos observar los efectos sintácticos de algunos aspectos de interés pragmático. En la primera versión (1) de la frase nominal hay una frase preposicional adjuntada al nombre (*Fijación*) jugando un papel temático que permite que la información focalizada quede codificada por *al soporte*. El contenido semántico del nombre que describe una acción está ligada a la frase preposicional *de la superficie decorada* de modo que la acción codificada por el grupo *Fijación* ↔ *al soporte* queda deliberadamente focalizada. Las características del texto sobre restauración de muebles sugieren que este tipo de focalización no es arbitrario o casual. En la segunda versión la acción completada se mantiene junta mientras que la frase preposicional se postpone. Y es el hecho de mantener en un mismo bloque la palabra *fijación* y su argumento más importante (al soporte) lo que da más fuerza al contenido de acción del nombre verbal.

Se pueden pues considerar dos aspectos en relación con ello. En primer lugar que en ambas lenguas la secuenciación natural es «fijar algo en alguna parte» porque «algo» tiene la función semántica de agente, es un argumento, mientras que «en alguna parte» es claramente un satélite y por tanto queda en una posición periférica dentro del marco predicativo.

Aún cuando en español, en comparación con el inglés, se permite una mayor libertad en el orden de palabras, también observamos cómo la segunda ocurrencia parece un es-

pañol más natural si el hablante quiere centrar la atención en la acción. En el original español las autoras eligieron la primera ocurrencia enfatizando la importancia del soporte.

También en la descripción GB se da cuenta del contenido de acción encapsulado en la palabra *fijación*, pero la forma de representarlo es mediante una transformación pasiva previa.

Podríamos concluir diciendo que una traducción al inglés asistida mediante de un tipo de representación que siga el modelo de Dik puede añadir precisión y reducir ambigüedad al permitir una notación más precisa y ajustada de aspectos bastante difusos y que, sin embargo estas dos lenguas codifican. Esta posibilidad de que cada palabra lleve aparejada un conjunto de distintos paquetes de información acerca del comportamiento de diferentes niveles de representación facilitaría el etiquetado de las mismas previo a la aplicación de cualquier parser.

BIBLIOGRAFÍA

- Bernárdez, E. 1995: *Teoría y Epistemología del Texto*. Madrid: Cátedra.
- Chomsky, N. 1992: *A Minimalist Program for Linguistic Theory*. MIT Occasional Papers in Linguistics. Cambridge. MA.
- Demonte, V. 1989: *Teoría sintáctica: De las estructuras a la reacción*. Ed. Síntesis. Madrid.
- Dik, S. C. 1989: *The Theory of Functional Grammar*. Foris Publications. Dordrecht. Holland / Providence RI.
- Fernández Lagunilla, M. & A. Igoa 1995: *Sintaxis y Cognición*. Madrid: Síntesis.
- Givón, T. 1979 a & b: *On Understanding Grammar*. New York: Academic Press.
- Givón, T. 1992: «The grammar of referential coherence as mental processing instructions». *Linguistics* 30.1: 5-55.
- Korst, Bieke van der 1989: *Functional Grammar and machine translation*. > John H. Connolly and S. Dik eds: *Functional Grammar and the computer*. pp. 289-316. Dordrecht: Foris.
- Langaker, R. 1987: *Foundations of Cognitive Grammar*. Stanford. Stanford University Press.
- Lakoff, G. & M. Johnson 1980: *Metaphors we live by*. Chicago: The University of Chicago Press.
- Moreno Cabrera, J. C. 1991: *Curso Universitario de Lingüística General*. Síntesis. Madrid.
- Pozo, J. I. 1989: *Teorías cognitivas del aprendizaje*. Madrid: Morata.
- Rumelhart, D. E. & J. L. McClelland 1992: *Introducción al Procesamiento Distribuido en Paralelo*. Ed. española de J. A. García Madruga. Madrid: Alianza Psicología.
- Schneider, W. & R. Shiffrin 1977: «Controlled and automatic human information processing in detection, search and attention». *Psychological Review* 84: 1-66.
- Shiffrin, R. & Schneider, W. 1977: *Controlled and automatic information processing, II: Perceptual learning, automatic attending and a general theory*. *Psychological Review* 84: 127-90.

