

# CONTRASTES EN EL USO DEL ARTÍCULO EN INGLÉS Y CASTELLANO. UN ALGORITMO PARA LA TRADUCCIÓN AUTOMÁTICA

TONI BADIA  
FLORA RAMÍREZ  
Eurotra - España

## 1. *Introducción*

La traducción de los sintagmas nominales del inglés al español presenta numerosos problemas, que empiezan desde la constatación de que las estructuras internas de los mismos son diferentes, hasta la distinta distribución de los especificadores que aparecen en él en cada lengua. El artículo ocupa un lugar prominente en estas discrepancias debido a que interviene en la mayoría de los sintagmas nominales y a que su aparición o ausencia viene determinada por factores sintácticos, semánticos y pragmáticos.

Por otra parte, no existe una bibliografía sobre el artículo que proporcione un tratamiento general pero suficientemente rico para su uso por un sistema de traducción automática.

Este artículo es el resultado de un estudio sobre la determinación en inglés y en español e intenta aportar un conjunto de reglas que resuelvan la proyección del artículo del inglés al español con un porcentaje relativamente alto de éxitos.

El enfoque por el que hemos optado es básicamente sintáctico precisamente por la interrelación de factores de diferentes niveles de descrip-

ción lingüística (sintácticos, semánticos y pragmáticos) que intervienen en la aparición (o ausencia) del artículo. Este enfoque nos ha permitido, por un lado, acercarnos al tema sin que la labor pareciera imposible. Por otro, debido a la utilidad que queríamos dar a nuestro sistema (esto es, que pudiera aplicarse a cualquier sistema de traducción automática), esta caracterización nos parece que puede resolver estos problemas en un porcentaje alto de casos sin que exija unos mínimos muy elevados. Así pues, aunque en el estudio se tienen en cuenta todo tipo de factores, la materialización de la proyección es básicamente sintáctica.

El conjunto de reglas que proponemos ha sido confeccionado para su aplicación dentro del sistema de traducción automática EUROTRA. Por ello, después de una breve descripción de los casos problemáticos en la traducción del artículo, haremos una breve exposición del funcionamiento del sistema EUROTRA (secc. 3.). A continuación, expondremos las reglas generales (secc. 4.); en las secciones siguientes se contemplan los factores intrasintagmáticos que intervienen en la traducción (secc. 5.) y los factores externos (intersintagmáticos principalmente, aunque también se contemplan factores semánticos y pragmáticos, secc. 6.).

## 2. Casos problemáticos en la traducción del artículo

Una breve comparación de las descripciones de los artículos inglés y español permite descubrir que casi necesariamente los usos de los artículos en ambas lenguas no pueden coincidir totalmente. En inglés el artículo determinado en su única forma morfológica sirve tanto para sintagmas nominales en singular como en plural, mientras que el artículo indeterminado solamente tiene forma singular. Por otra parte, en español ambos artículos tienen formas singulares y plurales. Aunque en algunos casos, el artículo indeterminado plural corresponde al cuantificador inglés *some*, llamado débil (Quirk et al., 1985) y representado a menudo en las gramáticas inglesas por *sm*, usualmente los sintagmas nominales españoles con el artículo indeterminado plural se traducen en inglés sin artículo. Estas correspondencias quedan reflejadas en la tabla siguiente.

		español	inglés
determinado	singular	el/la	the
	plural	los/las	the
indeterminado	singular	un/una	a/an
	plural	unos/unas	-/sm

Las oraciones de (2)-(5) ejemplifican estos usos del artículo en las dos lenguas <sup>1</sup>.

- (2) a. Esta mañana el niño ha presenciado un accidente  
 b. This morning the boy has witnessed an accident
- (3) a. Los estudiantes de tercero no han asistido a clase  
 b. The third year students didn't attend their lectures
- (4) a. Un profesor ha preguntado por ti  
 B. A professor was looking for you
- (5) a. Hay unos perros corriendo por el jardín  
 b. Dogs are running in the garden.

Así pues, en una primera aproximación ya encontramos discrepancias en el uso del artículo en las dos lenguas, fundamentadas en la no existencia de la forma de artículo indeterminado plural en inglés. No obstante, las diferencias en el uso del artículo no quedan reducidas a estos (relativamente poco numerosos) casos del artículo indeterminado plural español, sino que aparecen varias circunstancias en las cuales los usos de las dos lenguas no coinciden aun cuando las posibilidades formales morfológicas lo permitirían. Veamos algunos de estos casos.

En primer lugar, encontramos que hay ocasiones en las que en inglés aparecen sintagmas nominales sin determinante, que son traducidos al español por sintagmas nominales con el artículo determinado. Este fenó-

<sup>1</sup> En estos ejemplos, como en otros muchos del artículo, las traducciones propuestas no son siempre las únicas posibles; se trata de las más comunes o, por lo menos, de una de las más comunes.

meno es habitual en los sujetos de todos los verbos y en los objetos directos de algunos de ellos. Veamos un par de ejemplos:

- (6) a. Japanese eat raw fish
- b. Los japoneses comen pescado crudo
- (7) a. His brother hates oranges
- b. Su hermano detesta las naranjas.

El primer ejemplo puede generalizarse a todos los verbos, puesto que se trata de sujetos sin referente determinado en las oraciones llamadas genéricas o habituales: en las oraciones de (6) el sintagma nominal sujeto no es usado para referirse a ningún grupo particular de japoneses, sino al conjunto de los japoneses; asimismo, las oraciones enteras no constituyen ninguna afirmación sobre un evento particular (o sobre un conjunto de eventos), sino sobre una particularidad general de los japoneses.

Sin embargo, el ejemplo de las oraciones de (7) no es generalizable a todos los verbos. En unos casos los objetos directos ingleses sin artículo se traducen al español por sintagmas nominales con artículo determinado, mientras que en otros se traducen por sintagmas sin actualizar. Por ejemplo, los objetos directos de las siguientes oraciones se mantienen sin artículo al traducirlas del inglés al español.

- (8) a. That boy washes cars
- b. Aquel chico lava coches
- (9) a. These girls used to read books
- b. Estas chicas solían leer libros.

En ambos grupos de oraciones encontramos predicaciones que tienen también un carácter genérico o habitual; la diferencia en relación con la oración de (6) estriba en que en ésta la predicación es sobre un conjunto de entidades en general, mientras que en las demás se trata de una predicación sobre una entidad concreta (o unas entidades concretas). Así pues, en la oración de (6) se predica del conjunto de los japoneses que suelen comer pescado; mientras que en las de (7)-(9) se predica de su hermano, de aquel chico o de estas chicas que detesta las naranjas, que lava coches y que leen libros, respectivamente. La diferencia, pues, entre los objetos directos que llevan artículo y los que no lo llevan en español no radica en la interpretación (genérica o no) de la predicación. Aquello que diferencia unas oraciones de otras en cuanto a este aspecto es el tipo de relación que determina el verbo con respecto al referente del obje-

to directo. Mientras que la relación de detestar es una relación general que puede afectar a la totalidad de las entidades de una clase, las relaciones de lavar o de leer no pueden tener este carácter general: quien lava coches no puede lavarlos todos, ni quien lee libros puede leerlos todos, de manera que la relación de una entidad con la clase denotada por el objeto directo no puede agotarla.

Tenemos, pues, ejemplos de discrepancias en el uso del artículo en las dos lenguas determinadas por la función sintáctica de los sintagmas nominales en cuestión; en un caso (el del sujeto) se trata de un fenómeno general, que puede afectar a todos los verbos, y en otro queda restringido a un grupo de verbos, con unas características semánticas particulares.

Una discrepancia radicalmente distinta es la que aparece en los sintagmas nominales en función de atributo. En los casos en que en español el sintagma nominal singular en función de atributo aparece sin artículo, el inglés los construye con artículo indeterminado, tal como queda ejemplificado en las traducciones de las oraciones de (10) y (11).

- (10) a. Su padre es maestro  
 b. His father is a teacher
- (11) a. María es investigadora  
 b. María is' a researcher.

De todas maneras, la correspondencia entre los sintagmas nominales en esta posición no son tan sencillas, puesto que en español el uso del artículo indeterminado es obligado cuando el nombre aparece con ciertos modificadores, como puede verse en las oraciones de (12).

- (12) a. Her professor is a good researcher  
 b. Su profesora es una buena investigadora.

Las diferencias en el uso del artículo que acabamos de examinar dependen de factores externos al mismo sintagma nominal, en concreto del papel sintáctico que desempeñan en la oración. No obstante, éste no es el único motivo para que aparezcan sintagmas nominales con diferencias en su determinación: hay diferentes elementos internos al sintagma nominal que determinan también que tengan diferencias en el uso del artículo. Entre estos factores, vamos a destacar tres: el hecho de que el sustantivo núcleo del sintagma sea un nombre propio, que sea un sustantivo temporal y la coaparición del artículo con otros determinantes (naturalmente en el mismo sintagma).

En general, el uso del artículo con los nombres propios es idiosincrático, de manera que podemos encontrar todo tipo de proyecciones: sin artículo en las dos lenguas (13), con artículo determinado en las dos (14), y sin artículo determinado en inglés y con artículo en español (15); de la última combinación posible (es decir, con artículo en inglés y sin artículo determinado en español) no hemos encontrado ningún ejemplo.

- (13) a. England
- b. Inglaterra
- (14) a. the Alpen
- b. los Alpes
- (15) a. Japan
- b. el Japón.

Por otra parte, los sintagmas nominales temporales se comportan de forma distinta en relación con la presencia de artículo, de manera que aparecen sintagmas que se mantienen sin artículo al ser traducidos del inglés al español (como por ejemplo los de (16)-(17)), mientras que los hay que introducen el artículo determinado en español aun cuando no esté presente en inglés (como se muestra en los ejemplos de (18)-(19)).

- (16) a. The president is coming on May
- b. El presidente vendrá en mayo
- (17) a. I last saw him in 1986
- b. Lo vi por última vez en 1986
- (18) a. The president is coming on Tuesday
- b. El presidente vendrá el martes
- (19) a. They met at five
- b. Se reunieron a las cinco.

Finalmente, hay fenómenos de interrelación de varios determinantes en el seno de un sintagma nominal; en algunos de estos casos, las dos lenguas se comportan de manera distinta. Un primer ejemplo lo constituye el cuantificador *todo*, que en español tiene que coaparecer con el artículo determinado, mientras que su equivalente inglés (*all*) suele ir sin ningún artículo <sup>2</sup>.

<sup>2</sup> Nos referimos, claro está, al cuantificador *todo* que aparece sólo en plural con los nombres contables y en singular con los de materia. Naturalmente, existen otros usos de

- (20) a. All children attended the meeting  
 b. Todos los niños participaron en la reunión.

No obstante, hay otros ejemplos de discrepancia relacionados con la presencia de cuantificadores u otros determinantes; compárese *half the cake* con *la mitad del pastel*, *such a surprise* con *qué sorpresa*, *a lot of milk* con *mucha leche*, etcétera.

El último ejemplo de discrepancias que proponemos está relacionado con los pronombres posesivos: a menudo el inglés usa el posesivo cuando en español la información referente al poseedor o bien es sobreentendida (por ser el verbo suficientemente explícito), o bien es aportada por algún pronombre clítico. Considérense las oraciones de (21) y (22).

- (21) a. He has got his paper under his arm  
 b. Llevaba el periódico debajo del brazo  
 (22) a. He has washed his hands  
 b. Se ha lavado las manos.

En (21) podemos observar que los dos posesivos ingleses se traducen naturalmente mediante el artículo determinado, puesto que en español la relación de posesión (cuando no es marcada) viene dada implícitamente por el verbo *llevar*; así, únicamente se necesitaría un pronombre posesivo si fuera necesario indicar que el referente del sujeto del verbo no es el poseedor del periódico (o del brazo), como sería el caso de la oración inglesa *he has got my paper under his arm* (que debería ser traducida por *llevaba mi periódico debajo del brazo*).

Por otra parte, en las oraciones de (22) aparece la correspondencia entre el posesivo inglés y un pronombre clítico español en dativo, que es el que aporta la información del referente poseedor de las manos. Nótese que el pronombre *se* del ejemplo es la forma dativa clítica del pronombre personal de tercera persona: si el ejemplo hubiera estado en primera o segunda persona del singular se habría usado el pronombre *me*

---

este cuantificador, que podemos ejemplificar con los sintagmas nominales de las oraciones de (i):

- (i) a. Todo ciudadano debe cumplir las normas de convivencia  
 b. Se ha comido todo el pastel.

No obstante, aquí nos referimos únicamente al primer uso reseñado.

o *te*, respectivamente (*I have washed my hands* vs. *me he lavado las manos*, y *you have washed your hands* vs. *te has lavado las manos*).

Esta breve comparación de los usos del artículo en las dos lenguas nos muestra claramente que la traducción del artículo inglés al español no es un asunto que pueda resolverse mediante un par de reglas o tres. Es decir, no es posible partir de una generalización como la que indica que los artículos se mantienen siempre en la traducción del inglés al castellano, y que sólo hay que calcular los casos en que la ausencia de artículo inglés equivale al artículo indeterminado plural español. Hemos visto que la mayoría de las discrepancias están relacionadas con la presencia o ausencia del artículo determinado, por lo que cualquier planteamiento de las reglas de proyección del artículo del inglés al castellano debe ser bastante más complejo de lo que parece en una primera aproximación.

Naturalmente, para poder adentrarnos en los detalles que harán posible la formalización de unas reglas adecuadas para resolver estas discrepancias debemos partir de un conocimiento mínimo de la estrategia general del sistema EUROTRA para la traducción de una lengua a otra. Por esta razón, en la sección siguiente presentamos los aspectos fundamentales de esta estrategia, que van a ser los que nos permitan, en las secciones siguientes, presentar el algoritmo de traducción del artículo inglés al español.

### 3. *Introducción al sistema de traducción automático de EUROTRA*<sup>3</sup>

EUROTRA es un sistema de traducción automática basado en la transferencia. Esto significa que la traducción de una oración de la lengua de origen  $L_o$  a la lengua de destino  $L_d$  se realiza mediante tres estadios consecutivos:

- el análisis de la oración: proceso mediante el cual se consigue una representación abstracta del significado de esta oración, en la que se mantienen los elementos léxicos propios de la lengua de origen, siempre que pertenezcan a clases de palabras abiertas,

---

<sup>3</sup> Una descripción del proyecto EUROTRA puede hallarse en Cerdá et al. (1990), que puede resultar interesante porque adopta una perspectiva diferente de la que tomamos en esta sección. Por otra parte, existen dos recientes presentaciones de EUROTRA, ambas muy completas: Durand et al. (1991) y Allegranza et al. (1991).



- la transferencia de la representación abstracta propia de la  $L_0$  a una representación abstracta propia de la  $L_d$ : en esta etapa se adapta la representación abstracta obtenida en el proceso de análisis a las exigencias de las representaciones abstractas propias de la  $L_d$ ; una parte importante de esta adaptación consiste en sustituir las unidades léxicas de la  $L_0$  por unidades léxicas de la  $L_d$ , y
- la síntesis de la oración en la  $L_d$ : proceso mediante el cual se genera una oración en la  $L_d$  a partir de la representación abstracta correspondiente.

A su vez, los estadios de análisis y síntesis se realizan en cuatro etapas sucesivas, que corresponden a distintos tipos de información lingüística necesaria para el éxito del proceso. En una primera caracterización, podemos definir estas etapas como de información morfológica, de sintaxis de constituyentes, de sintaxis relacional y de sintaxis profunda.

Uno de los ideales presentes en el diseño del sistema es el de preservar lo más posible su modularidad. Es decir, que cada uno de sus componentes pueda ejercer su función independientemente de los demás; esto naturalmente facilita la legibilidad del mismo, así como la posibilidad de ser ampliado, mejorado o corregido. Una primera distribución clara de tareas es la que distingue entre las operaciones lingüísticas y las informáticas. Mientras que las primeras están divididas en etapas según el tipo de información lingüística que tratan, las segundas son más generales. El diseño permite únicamente dos tipos generales de componentes: las gramáticas y los traductores entre gramáticas. Ello significa que las distintas gramáticas motivadas lingüísticamente tienen exactamente el mismo soporte informático para funcionar y que las relaciones entre componentes son idénticas sean cuales sean los componentes relacionados. Con ello se consigue que pueda mejorarse el soporte informático sin afectar a las gramáticas escritas y, lo que es mucho más importante para el lingüista, que las gramáticas puedan ser modificadas, ampliadas, etc., sin que haya que adaptar el soporte informático.

Pero incluso dentro del campo propiamente lingüístico, existe el principio de la modularidad. En general se intenta que cada gramática sea lo más independiente posible tanto de la que le precede como de la que le sigue, de manera que la información que ella trata no dependa de la que trata el componente anterior o el siguiente. De esta manera, se

pretende poder distribuir el trabajo entre los distintos componentes de manera independiente, hasta el punto que, por ejemplo, la modificación del tratamiento morfológico de los verbos afecte relativamente poco a su tratamiento sintáctico.

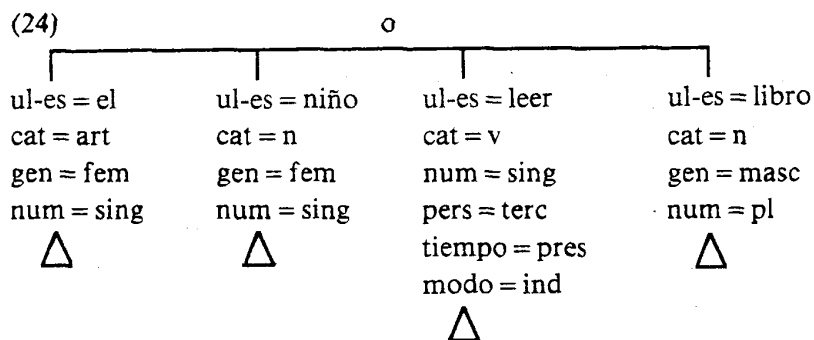
Aunque éste sea el ideal al que tiende toda la implementación lingüística llevada a cabo en EUROTRA, no deja de ser cierto que hay fenómenos que no pueden ser implementados en un único nivel y que, por consiguiente, la información relativa a ellos tiene que ser tratada en más de un componente. Esto se debe en parte al diseño de los componentes informáticos, particularmente a su capacidad expresiva, y en parte a la estructuración de la información lingüística. Puesto que, como veremos, el tratamiento del artículo es uno de estos casos, no merece la pena que nos detengamos más en este aspecto por el momento. Conviene, por otra parte, hacer hincapié en los aspectos regulares de la distribución de la información lingüística, puesto que en parte ellos van a motivar el tratamiento de la traducción del artículo que proponemos.

Como hemos mencionado anteriormente, los estadios de análisis y de síntesis se subdividen en cuatro etapas sucesivas que aportan la información morfológica, de sintaxis de constituyentes, de sintaxis relacional y de sintaxis profunda. Como la diferencia entre uno y otro nivel es fundamentalmente de tipo estructural (puesto que en cada uno de ellos se realizan de manera distinta las relaciones entre los constituyentes), conviene conocerlos para entender el tratamiento dado a fenómenos que no tienen grandes repercusiones estructurales. Para ello, vamos a escoger una oración española muy simple para presentar a continuación los elementos esenciales de su representación en los distintos niveles del sistema. Consideremos, por ejemplo, la oración de (21).

(23) La niña lee libros.

En el primer nivel de análisis, el morfológico, esta oración es representada mediante un árbol en el que cada palabra es una hija directa de la oración; asimismo las palabras pueden tener estructura interna, puesto que algunas (todas en nuestro ejemplo) están formadas por una raíz y unos morfemas. Naturalmente el objetivo de este módulo es conseguir toda la información posible a partir de la estructura interna de las palabras. Así pues, en este caso, el módulo dará para los nombres y el artículo la información correspondiente al género y al número, y, para el ver-

bo, la de número, persona, tiempo y modo. El árbol correspondiente será parecido al siguiente <sup>4</sup>:



En el árbol de (24) se ha omitido la estructura interna de las palabras, puesto que no es relevante para la presente discusión. Hay que notar, sin embargo, que la información extraída de la raíz y de los sufijos flexivos es elevada al modo superior de la palabra, de manera que en los niveles siguientes pueden eliminarse los constituyentes internos de las palabras.

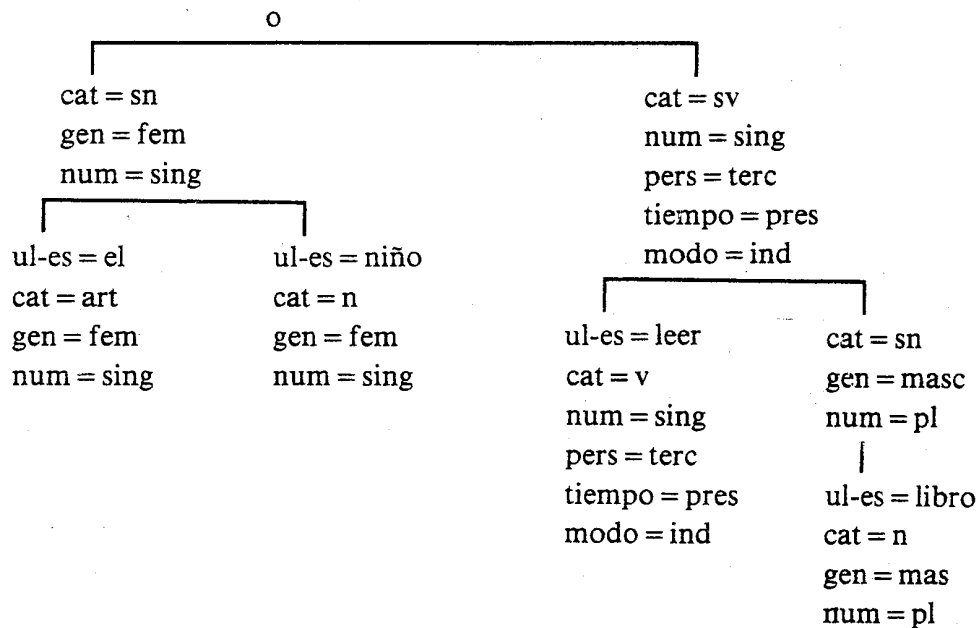
El segundo nivel de análisis elabora una estructura de constituyentes para cada oración. Es decir, a partir de un nombre se crea un sintagma nominal (con todos sus posibles complementos), a partir de un verbo se crea un sintagma verbal, etcétera. En el caso de la oración de (23) el árbol resultante ya tiene una cierta profundidad en su estructura, de manera que los constituyentes inmediatos del nivel morfológico anterior ya no son hijas de la oración sino de sintagmas nominales o verbales.

El árbol de (25) está simplificado en el sentido que sólo contiene la información mostrada en el análisis del módulo morfológico anterior (en (24)). Como es natural, el resultado de un análisis real contendría además otro tipo de información, que va a ser útil en los niveles posteriores, como por ejemplo, la relativa al orden de las palabras o a características del nombre núcleo de un sintagma nominal (si es común, propio, pro-

<sup>4</sup> Las abreviaturas del árbol son claras; la única que merece una indicación es ul-es, que puede traducirse como unidad léxica española; nótese que representa la forma canónica del diccionario, de manera que las variaciones de género, número, persona, modo, tiempo, etc., han sido neutralizadas, puesto que esta información ha sido convertida en rasgos.

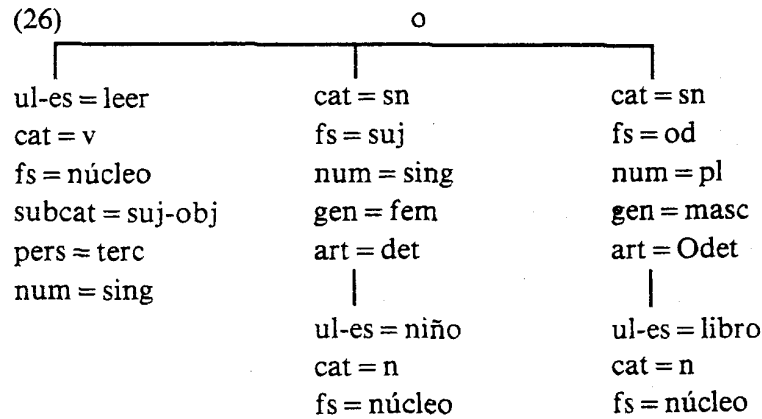
nombre, etc.). Como puede verse en el árbol de (25), la información contenida en las palabras es elevada a los modos sintagmáticos correspondientes; gracias a este proceso van a poder formularse reglas generales en el nivel siguiente, como por ejemplo la de la concordancia del sujeto con el verbo en persona y número, sin tener que adentrarse en los sintagmas correspondientes.

(25)



En el nivel siguiente, el de la sintaxis relacional, se establecen las funciones sintácticas de los constituyentes, de manera que cada uno de ellos recibe un valor para el atributo fs (función sintáctica), fruto de sus propias características y de las de su rector. Para llevar a cabo esta asignación de forma correcta se tienen en cuenta las posibilidades de subcategorización de los rectores; así, por ejemplo, para los verbos se tiene en cuenta si son transitivos, si pueden pasivizar, si tienen un complemento preposicional, o un objeto indirecto, etcétera. Por supuesto se observa si hay constituyentes que cumplen las exigencias de su rector; por ejemplo, si son sintagmas preposicionales con *a*, si concuerdan en persona y número con el verbo, si son otro tipo de complementos preposicionales, etcétera. De todos estos cálculos surge una representación en la que cada constituyente tiene asignada una función sintáctica; téngase en cuenta que los elementos no regidos (por ejemplo, los adjetivos en los sintagmas nominales o los complementos temporales en las oraciones) son conside-

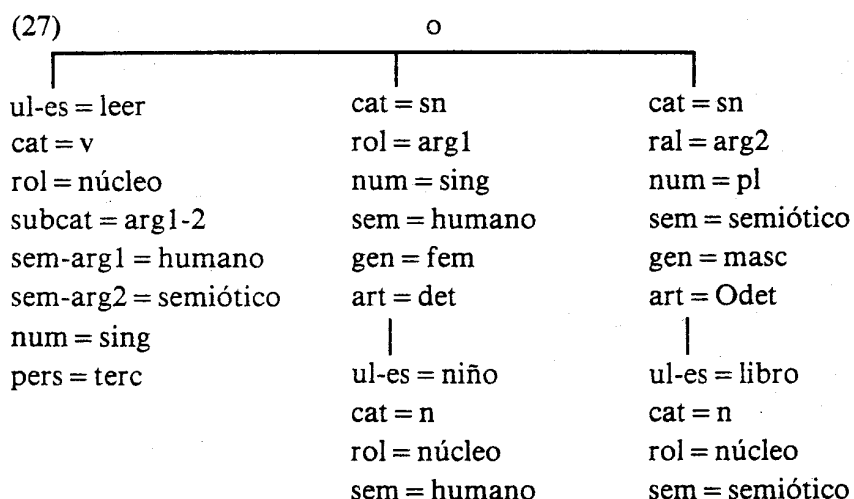
rados modificadores (es decir, tienen el par atributo-valor fs = mod). En consecuencia, la oración que venimos analizando tendría el siguiente árbol de representación:



Además de la asignación de funciones sintácticas, se operan en este nivel una serie de cambios en la representación, dirigidos todos ellos a obtener una representación canónica (independiente del orden superficial de los constituyentes y común en lo posible a todas las lenguas). Por un lado, los constituyentes de las oraciones reciben un orden canónico: el núcleo, el sujeto, el objeto directo, el objeto indirecto, los complementos preposicionales, etc., hasta terminar con los modificadores. Además el sintagma verbal desaparece, de manera que se obtiene una representación plana de los constituyentes de la oración, en la que todos ellos están en una misma relación estructural con el verbo núcleo. Por otra parte, los artículos desaparecen de la representación estructural de los sintagmas nominales y la información sobre su ausencia, presencia y tipo de artículo es llevada por el rasgo art, cuyos valores son det, para el artículo determinado, indet, para el indeterminado, y Odet para la ausencia de artículo. Finalmente, conviene notar que el árbol que presentamos no contiene la información de tipo temporal del verbo o de la oración. Esto es así porque esta información sufre durante este nivel y el siguiente una serie de procesos de cálculo que la convierten en información realmente semántica. Como éste no es un tema al que vayamos a tener que referirnos en lo sucesivo, simplemente lo dejamos de lado <sup>5</sup>.

<sup>5</sup> Para una presentación del tratamiento dado en EUROTRA al tiempo y aspecto verbales, puede consultarse van Eynde (1988).

En el último nivel de análisis, se convierten las funciones sintácticas en roles sintácticos propios de las gramáticas profundas de dependencias. Se trata de relaciones más abstractas que las funciones sintácticas que, por consiguiente, pueden ser más fácilmente comunes a las distintas lenguas; así, por ejemplo, las oraciones en voz pasiva son deshechas, de manera que el sujeto de una pasiva y el objeto de una activa tienen la misma asignación de rol (la del segundo argumento), mientras que el complemento agente de una pasiva (si lo hay) y el sujeto de una activa tienen asignado el rol de primer argumento. En el caso de la oración que venimos analizando no hay fenómenos que alteren la relación entre los constituyentes de manera que el sujeto del nivel anterior se va a convertir en argumento primero, y el objeto, en argumento segundo. El árbol, pues, va a ser el siguiente:

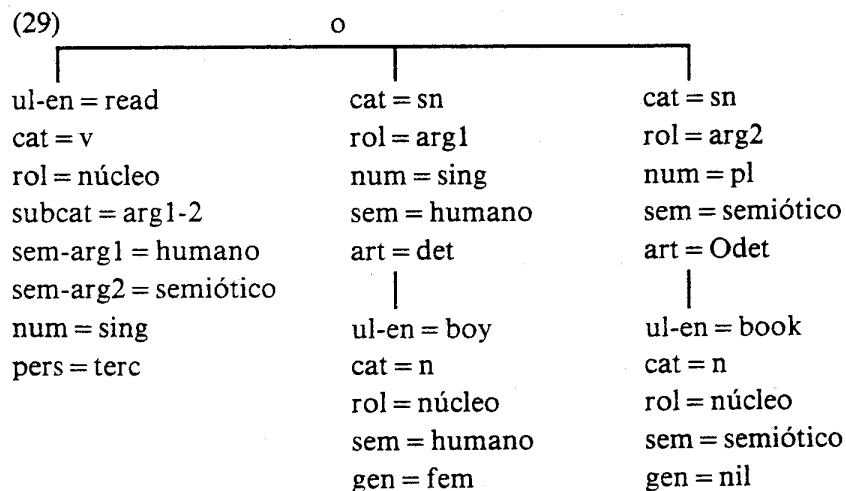


En este nivel, además se llevan a cabo los cálculos semánticos necesarios para la correcta selección de los argumentos. Esto se hace a partir de la codificación de unos valores de semántica léxica que caracterizan la interpretación de los nombres (en nuestro ejemplo, *niña* es caracterizada como humano y *libro*, como semiótico), y de la codificación de las restricciones que imponen los rectores sobre sus argumentos (por ejemplo, el verbo *leer* exige que su argumento primero sea humano y que su argumento segundo sea semiótico) <sup>6</sup>.

<sup>6</sup> En realidad, la relación entre las restricciones impuestas por el rector a sus argumentos y la codificación semántica de éstos no es tan simple como el ejemplo sugiere. De

Esta estructura, llamada «estructura de interficie», es la que es usada como punto de partida para el proceso de transferencia. Supongamos, por ejemplo, que vamos a traducir esta oración del español al inglés. El módulo de transferencia español-inglés contiene una serie de reglas para traducir las unidades léxicas (es decir, los nodos terminales) propias de las representaciones españolas en unidades léxicas adecuadas para las representaciones inglesas. De esta manera, mediante la aplicación de reglas como las de (28), el árbol español de representación abstracta (27) se convierte en un árbol inglés de representación abstracta (el mostrado en (29)).

- (28) a. {cat = v, ul-es = leer} ⇒ {cat = v, ul-en = read}  
 b. {cat = n, ul-es = niño, gen = fem} ⇒ {cat = n, ul-en = girl}  
 c. {cat = n, ul-es = libro} ⇒ {cat = n, ul-en = book}



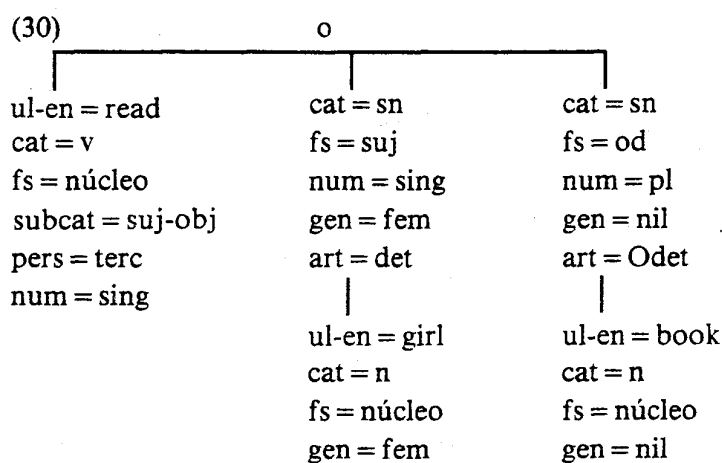
Naturalmente, reglas como las de (28) no son suficientes para llevar a cabo una adaptación completa de una representación abstracta española a una de inglesa. En nuestro caso, que ya de por sí es muy simple, encontramos otro cambio que ha sido necesario; se trata de la acción de borrado del género de los sintagmas nominales. La razón es muy simple, puesto que el género de los nombres no es coincidente en las dos

---

hecho, en la implementación real de EUROTRA los rasgos de semántica léxica conforman una estructura arbórea, de manera que los rasgos con que se tipifican los nombres son rasgos terminales, mientras que los rasgos con que se caracterizan las posiciones argumentales son casi siempre rasgos no terminales en el árbol.

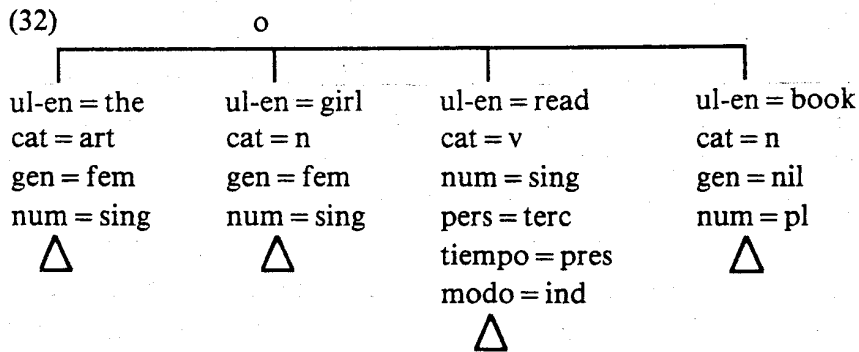
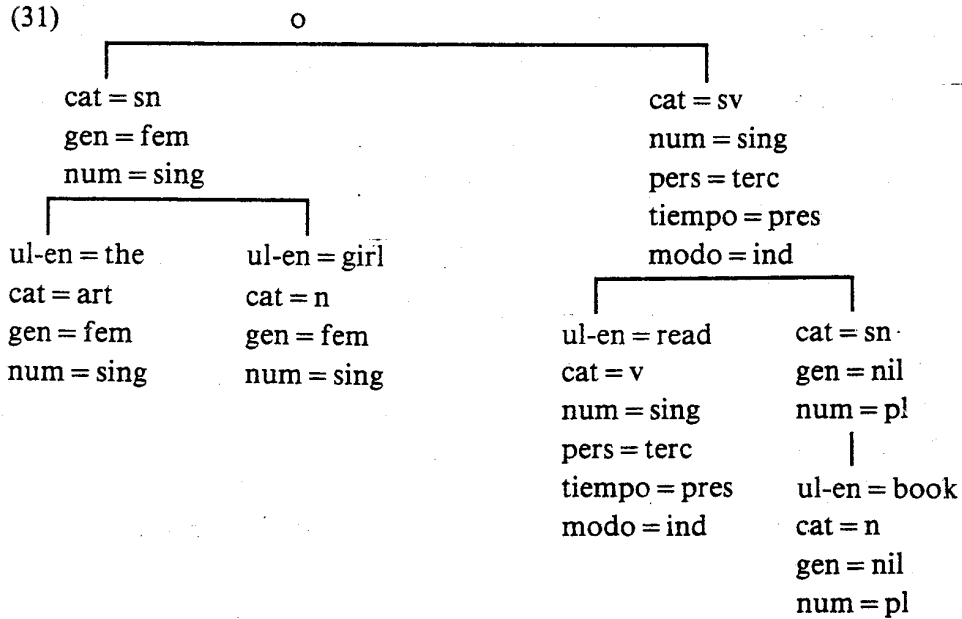
lenguas: en nuestro ejemplo en el sintagma nominal que tiene *book* como núcleo habría aparecido una contradicción, ya que tenía en la representación española *gen = masc* y el nombre inglés *book* no es masculino, sino neutro (que marcamos como *gen = nil*). Esta operación de borrado se aplica únicamente en el objeto de salida. Como puede verse en (28), en la regla de *niño*, la información del género es útil para la lengua origen ya que con ella podremos traducir adecuadamente esta unidad léxica. Debemos, por lo tanto, tener en cuenta que nuestro módulo de transferencia va a contener reglas más generales que las que simplemente traducen unidades léxicas.

Una vez conseguida una representación abstracta inglesa, se puede proceder a ir preparando su conversión en una cadena lingüística superficial; para ello se van a seguir los mismos pasos que para el análisis de la oración española, pero en dirección contraria: de la representación de interficie, pasamos a la de sintaxis relacional, de ésta a la de constituyentes y, finalmente, a la morfológica. Los árboles correspondientes son los de (30), (31) y (32), respectivamente, que al final producen una oración inglesa (la de (33)), que es una traducción de la original española de (23).



Nótese que hasta este nivel la información sobre el artículo de los sintagmas nominales está contenida en el rasgo *art*, igual que en los árboles de representación de la oración española. En el nivel siguiente (el de la estructura de constituyentes), esta información es convertida en la presencia (o ausencia) estructural del artículo, que ya se mantiene hasta la generación de la cadena superficial inglesa.





(33) The girl reads books.

Puesto que la oración que hemos tomado como ejemplo no presentaba ningún problema especial, la correspondencia entre la oración española de (23) y la inglesa de (33) es casi total en todos sus aspectos. No obstante, ya hemos ido indicando que un sistema como el utilizado en Eurotra permite tratar casos mucho más complejos; el ejemplo no ha incluido ninguno de ellos a fin de no complicar las cosas innecesariamente, ya que se trataba de mostrar el marco general en el que va a llevarse a cabo la traducción del artículo que presentamos.

En lo que sigue, vamos a indicar para cada regla que presentamos a qué nivel lingüístico corresponde. Lo haremos mediante el uso de los acrónimos usados realmente en Eurotra, que son los siguientes:

EMSEN	nivel morfológico del inglés
ECSEN	nivel de sintaxis de constituyentes del inglés
ERSEN	nivel de sintaxis relacional del inglés
ISEN	nivel de sintaxis profunda (estructura de interficie) del inglés
ISES	nivel de sintaxis profunda (estructura de interficie) del español
ERSES	nivel de sintaxis relacional del español
ECSES	nivel de sintaxis de constituyentes del español
EMSES	nivel morfológico del español.

Las reglas de los traductores serán caracterizadas mediante la indicación de los dos niveles que relacionan; así, una regla con la indicación ISEN - ISES será del traductor de la representación de interficie inglesa a la española, y una regla con la indicación ISES - ERSSES pertenecerá al traductor que relaciona el nivel de interficie español con el de sintaxis relacional española.

#### 4. *Regla general de la traducción del artículo inglés al español*

Si observamos con un poco de atención los casos reseñados en la sección 2, en los que no puede llevarse a cabo directamente la proyección del artículo inglés al español, veremos que en su mayoría se trata de situaciones en las que la ausencia de artículo inglés se proyecta a uno de los dos artículos españoles<sup>7</sup>. Por ello podemos escribir en el módulo de transferencia inglés-español unas reglas generales que mantengan el artículo cuando el inglés lo tiene, sea cual sea el artículo presente en el texto original, y que, por el contrario, no mantengan la ausencia de artículo cuando en inglés no lo hay. De esta manera, se consigue que en los niveles propios de la generación podamos calcular cuál es el artículo adecuado sin tener que borrar ya el valor original del inglés. Veamos en concreto cómo formulamos estos principios.

Recordamos que los sintagmas nominales ingleses (al igual que los españoles en el ejemplo anterior de la sección 3) no contienen ningún nodo estructural correspondiente al artículo, sino que la información sobre si tenían o no artículo (y sobre cuál de ellos era, si es que lo había) está representada por un rasgo en el nodo del sintagma nominal:

$$(34) \{art = det\}, \{art = indet\} \text{ o } \{art = Odet\}$$

<sup>7</sup> Hay una excepción importante a esta afirmación: la de los atributos, que será tratada específicamente en la sección 6.5.

Según se indica en (34), art = det representa la presencia del artículo determinado (*the* en el original inglés), art = indet indica la presencia del artículo indeterminado (*a* o *an* en el sintagma nominal inglés), y art = Odet expresa que el sintagma nominal no llevaba ningún artículo (aunque, naturalmente, podía llevar algún otro determinante). Así, en (35) se indica qué codificación corresponde a algunos sintagmas nominales ingleses:

- (35) a. the girl           → art = det  
       b. a girl            → art = indet  
       c. girls             → art = Odet  
       d. many girls      → art = Odet  
       e. the three girls → art = det  
       f. most girls       → art = Odet

En general la información positiva (es decir, art = det o art = indet se mantiene en español), mientras que la negativa tiene que ser calculada de nuevo (puesto que hay muchos casos en los que la ausencia de artículo inglés se traduce por un artículo u otro en español). Tenemos que formular, pues, tres reglas para hacer estas operaciones, una para cada uno de los valores originales del rasgo art.

(36) ISEN - ISES

FART 1 = {cat = sn, art = det} ⇒ {art = det}

FART 2 = {cat = sn, art = indet, rol = argAS; argAO} ⇒ {art = indet}

FART 3 = {cat = sn, art = Odet} ⇒ {art = ?}

La primera regla de (36), FART 1, simplemente indica que un sintagma nominal con artículo determinado en inglés mantiene este artículo en español.

La segunda regla, FART 2, es un poco más compleja, puesto que tiene en cuenta el rol del sintagma nominal. La regla dice que todo sintagma nominal inglés con artículo indeterminado que no sea ni atributo del sujeto (argAS) ni atributo del objeto (argAO) se traduce por un sintagma nominal español que mantiene el artículo indeterminado. Nótese que la negación de un valor se expresa mediante el operador =, y la alternancia de valores mediante el punto y coma (;).

La tercera regla, FART 3, en cambio, no permite copiar la ausencia de artículo inglés al correspondiente sintagma nominal español. Un rasgo con el valor abierto en el lado derecho de una regla de este tipo (expresa-

do por el interrogante del par art = ?) indica que este atributo puede tener cualquier valor; de manera que de momento se deja abierto, sin especificar ninguno.

Esto es lo que permitirá que en los niveles de la generación española puedan llevarse a cabo los cálculos adecuados. Como los sintagmas nominales que en el original inglés ya tenían artículo determinado o indeterminado (excepto para los atributos) siguen teniendo estos artículos, los cálculos que hagamos en los niveles siguientes no les afectarán, puesto que las reglas no pueden sobrescribir un valor por encima de otro. Así, una regla cualquiera de las que vamos a ver en las secciones siguientes que indique que en unas condiciones concretas un sintagma nominal tiene que llevar un artículo determinado no se aplicará a aquellos sintagmas nominales que ya tengan especificado un valor para su rasgo art. De este hecho se deduce que se da preferencia a las reglas que están en un nivel anterior en el proceso de traducción. Por ello, no es indiferente el nivel en el que se plantean unas reglas y otras.

Nuestra implementación tiene en cuenta este aspecto de manera que pertenecen a niveles anteriores aquellas reglas que tienen preferencia de aplicación; de ahí que las últimas reglas en aplicarse tengan un campo muy restringido de aplicación, ya que sólo lo harán en aquellos sintagmas nominales que no tengan especificado un valor para art, y éstos van a ser lógicamente muy pocos.

##### 5. *Condiciones internas del sintagma nominal para la traducción del artículo*

Las primeras condiciones que se van a tener en cuenta para la correcta resolución de los casos en que el módulo de transferencia no ha copiado el artículo son las que resultan de la estructura interna del sintagma nominal. Son, pues, reglas que se aplicarán en el mismo momento de construirse los sintagmas nominales (es decir, sin esperar a que tengan asignado un rol argumental en la oración o sintagma al que pertenecen). De entre estas condiciones vamos a tratar los sintagmas cuyo núcleo es un nombre propio, la coocurrencia con otros determinantes y los sintagmas con sustantivos temporales como núcleo.

En general, la decisión de si los sintagmas nominales con nombres propios llevan o no artículo depende estrictamente de características par-

ticulares de cada nombre en concreto, de manera que es prácticamente imposible proceder a una generalización que permita formular reglas correctamente. Por ello, distinguiremos entre los nombres propios simples y los nombres propios complejos (formados por un nombre propio simple y un nombre común, con o sin preposición).

En cuanto a los nombres simples, aparecerán marcados en el diccionario con un rasgo que indique si normalmente aparecen con o sin artículo. Llamamos artic a este rasgo, que tiene como posibles valores sí y no. En el diccionario español, pues, aquellos nombres que sean propios (o sea, los que están caracterizados como clas-n = propio, es decir, clase de nombre = propio) deberán llevar una de las caracterizaciones siguientes: artic = sí o artic = no. Así, las entradas léxicas correspondientes a los nombres propios *Francia* y *La India* serán respectivamente las siguientes:

- (37) a. {ul-es = Francia, cat = n, clas-n = propio, gen = fem, artic = no}  
 b. {ul-es = India, cat = n, clas-n = propio, gen = fem, artic = sí}

A partir de entradas léxicas como éstas, las reglas siguientes permiten decidir el artículo que deben llevar los sintagmas nominales que tienen nombres propios como núcleo.

(38) ISES

FART PROP1 = {art = det/cat = sn}  
 [ {/cat = n, clas-n = propio, artic = sí} ].

FART PROP2 = {art = Odet/cat = sn}  
 [ {/cat = n, clas-n = propio, artic = no} ].

Nótese que estas reglas se aplicarán exclusivamente a sintagmas nominales que tengan nombres propios como núcleo (los comunes están marcados como clas-n = común, los pronombres, como clas-n = pron, etc.). Las reglas usan la barra inclinada para distinguir lo que es el contexto de aplicación de la regla (a la derecha de la barra) y el rasgo nuevo que la regla introduce (a la izquierda de la barra). Las dos reglas, pues, se aplican a estructuras de sintagmas nominales que sólo tienen una hija (es decir, un nombre), especificadas según los rasgos que aparecen en los dos nodos. Así pues, la regla FART PROP1 se aplicará solamente a árboles como el siguiente.

- (39)            cat = sn  
                   |  
                   cat = n  
                   clas-n = propio  
                   artic = sí.

Para tratar los nombres complejos podemos introducir reglas que tengan en cuenta las estructuras que éstos pueden tener. Por ejemplo, nos podemos fijar en estructuras del tipo *la Universidad de York* o *la ciudad de Londres*. En ellas encontramos como núcleo un nombre común (*universidad* y *ciudad* en los ejemplos) y un complemento preposicional (con *de*) en el que aparece un nombre propio como núcleo de su sintagma nominal (*York* y *Londres* en los ejemplos). Además, se da la circunstancia que el nombre común tiene que ser lo que podríamos llamar un clasificador del nombre propio (es decir, que tiene que haber una relación semántica de clase a miembro de la clase, o de tipo a ejemplar, entre la denotación del uno a la del otro). Los nombres clasificadores de nombres propios son nombres comunes que denotan conjuntos de entidades que suelen tener nombre propios en nuestra cultura: *río*, *ciudad*, *universidad*, *lago*, etcétera. Para estos casos, presentamos la regla siguiente:

- (40) ISES

FART PROP3 = {art = det/cat = sn}  
                   [ {/cat = n, sem = clasif-propio},  
                   {/cat = sp} [ {/cat = p},  
                   {/cat = sn, clas-n = propio} ]].

Naturalmente, con pocos cambios podemos preparar una regla que trate los casos de *el río Ebro* o *el lago Titicaca*:

- (41) ISES

FART PROP4 = {art = det/cat = sn}  
                   [ {/cat = n, sem = clasif-propio},  
                   {/cat = sn, clas-n = propio} ].

Otra de las condiciones internas al sintagma nominal que hay que tener en cuenta es la coaparición del artículo con otros determinantes. Uno de los casos más característicos es el de los cuantificadores *all* y *every* ingleses que pueden aparecer solos, sin ningún otro determinante, como se muestra en (42), mientras que su equivalente español *todos* tiene

que aparecer con otro determinante definido (habitualmente con el artículo determinado), como puede verse en (43).

- (42) a. all boys  
 b. all the boys  
 c. every boy  
 d. all our friends  
 e. all these books
- (43) a. todos los chicos (equivalente a 42a, b i c)  
 b. \*todos chicos  
 c. todos nuestros amigos (equivalente a 42d)  
 d. todos estos libros (equivalente a 42e)

Como por otra parte, tanto *all* como *todos* pueden ir seguidos por un determinante posesivo o por un demostrativo (casos en que, naturalmente, no aparece el artículo determinado), como se muestra en (42d-e) y (43c-d), las reglas de generación del artículo en relación con *todos* tienen que tener en cuenta si aparece algún otro determinante definido: si lo hay no se generará el artículo, mientras que si no lo hay se generará. Las reglas pueden tener la forma siguiente:

- (44) ISES
- FART TODOS1 = {art = det/cat = sn}  
 [+ {}],  
 {/cat = scuant} [ \*{}],  
 {/cat = cuant, ul-es = todos} ],  
 \*/cat \_ = sdem; sn} ].
- FART TODOS2 = {art = Odet/cat = sn}  
 [+ {}],  
 {/cat = scuant} [ \*{}],  
 {/cat = cuant, ul-es = todos} ],  
 \*{}],  
 {/cat = sdem; sn},  
 \*{} ].

La regla FART TODOS1 contempla el caso en que no haya en el sintagma nominal ningún demostrativo ni posesivo (tal como se indica en la última línea de la regla)<sup>8</sup>. Por esta razón se inserta el rasgo art = det

<sup>8</sup> Los pronombres posesivos son tratados en la gramática, al igual que los otros pronombres, como nombres. Como por otra parte éstos son los únicos sintagmas nominales

en el nodo del sintagma nominal. Por su parte, la regla FART TODOS2 contempla los casos en que entre las hijas del sintagma nominal aparece o un demostrativo o un posesivo; en ellos, el rasgo insertado en el nodo madre es el de art = Odet, para que en el proceso de generación no se cree ningún artículo. En las dos reglas aparecen nodos sin ningún rasgo especificado ( $\{ \}$ ), que indican cualquier tipo de nodo (es decir, un nodo cualquiera especificado por los rasgos que sea). Por otra parte, estos nodos poco especificados (y algún otro nodo) pueden llevar un operador \* o + — el primero (\*) indica cero o más apariciones de un nodo del tipo especificado, mientras que el segundo (+) indica una o más apariciones del nodo (implicando, por lo tanto, obligatoriedad de al menos uno). En general, los nodos poco especificados y los operadores \* y + sirven para mejorar la legibilidad de las reglas, al no tener que especificar todos sus posibles contextos.

Naturalmente, tienen que prepararse otras reglas parecidas para los casos en que hay discrepancias en la coaparición de otros determinantes con el artículo, como por ejemplo los pares mostrados en (45).

- (45) a. half the cake  
 a'. la mitad del pastel  
 b. such a surprise  
 b'. tal/semajante/qué sorpresa  
 c. a lot of milk  
 c'. mucha leche.

No las formulamos todas aquí para no entorpecer demasiado el curso del artículo y porque en muchos casos las reglas concretas dependen de la manera en que se haya producido la traducción concreta de los cuantificadores en el módulo de transferencia, lo cual comportaría varias complicaciones (ya que no pueden traducirse de una manera uniforme).

Finalmente vamos a considerar algunos casos en que el núcleo del sintagma nominal es un nombre temporal. Consideremos los siguientes ejemplos:

- (46) a. nine o'clock  
 a'. las nueve  
 b. on monday  
 b'. el lunes.

---

que pueden aparecer en esta posición, la especificación de la categoría ya es suficiente.



En general, los días de la semana y las horas del día suelen aparecer sin artículo en inglés (46a, b) y con el artículo determinado en español (46a', b'). Este fenómeno puede ser tratado de una manera bastante sencilla mediante una regla como la siguiente:

(47) ISES

FART TIEMPO = {art = det/cat = sn, valor adv = tiempo}  
 [ {/cat = n, valor-adv = tiempo, sem = hora; día},  
 \*{} ].

Esta regla considera los sintagmas nominales con un nombre con valor temporal como núcleo (es decir, los que tienen valor adverbial temporal), cuyo valor semántico está explícitamente especificado como indicando día o hora (sem = hora; día). En estos contextos, se inserta el rasgo art = det que va a producir la generación del artículo en el módulo de sintaxis de constituyentes español.

#### 6. *Condiciones externas al sintagma nominal para la traducción del artículo*

En los siguientes párrafos vamos a presentar el conjunto de reglas que se hacen necesarias para la correcta generación del artículo dentro del marco oracional. Para ello, la función sintáctica constituye una información relevante que condiciona distintos comportamientos. No menos importante será también el tratamiento de diversos tipos de verbos que determinan la actualización o no del sintagma que les acompaña. Hay que tener en cuenta que estas posibilidades son características de cada lengua y las regularidades que aparecen en una no tienen por qué aparecer en la otra. El siguiente grupo de oraciones ejemplifican los casos que vamos a tratar:

- (48) Por delante de nuestra casa pasan vacas
- (49) Hay leones con la cola larga
- (50) Juan prefiere las manzanas a las naranjas
- (51) Tiene los ojos grandes
- (52) Es profesor de universidad
- (53) Se ha roto la pierna
- (54) Echó agua en el vaso.

Todos estos casos han de tenerse en cuenta expresamente a causa de su comportamiento específico en las dos lenguas.

### 6.1. *Verbos presentativos*

En castellano, puede generalizarse que todo sujeto ha de estar actualizado<sup>9</sup>. Esta regla se opone a la regla general del inglés que permitirá que los sintagmas no referenciales aparezcan sin actualizador:

- (55) a. Los niños no lloran  
b. Children do not cry

Los sintagmas nominales de (55) son de tipo genérico y ejemplifican el hecho de que en castellano las oraciones genéricas han de aparecer con un sintagma en función de sujeto con el rasgo art = det; todo lo contrario de lo que sucede en inglés. Esto es así incluso cuando el sintagma no denote entidades sino la clase entera o el tipo. Un contraejemplo a esta regla del castellano lo constituyen los sujetos superficiales de los verbos de tipo presentativo (o inacusativo). La oración bajo (48) es un buen ejemplo del comportamiento que han de seguir estos sintagmas sujetos junto a verbos de este tipo. Marcando léxicamente el verbo y aplicando la siguiente regla obtendremos la generación correcta de estas oraciones.

(56) ISES

FART PRESENT = {/cat = o}  
[{/cat = v, subcat = arg2, tipo-v = present},  
{art = Odet/cat = sn, rol = arg2},  
\* {}].

Como se recordará, estos sintagmas llegan al nivel de interficie del español con el valor de art sin especificar, puesto que en inglés el valor era Odet.

La regla FART PRESENT especifica que el sintagma nominal ha de desempeñar el rol de arg2. Dentro del marco lingüístico de Eurotra, los sujetos de los verbos inacusativos se identifican en cuanto a su rol con los objetos de los verbos transitivos, de ahí que el valor de rol sea el

<sup>9</sup> Por supuesto en español pueden encontrarse sintagmas nominales no actualizados en función de sujeto. Son muy comunes en frases hechas y refranes (*Perro ladrador, poco mordedor*, y en expresiones comparativas con un marcador negativo del tipo *Hombres como éste no actúan a tontas y a locas*).

de arg2. El verbo, además de estar marcado en el léxico mediante el rasgo tipo-v, presenta un espectro limitado de estructuras argumentales o subcategorizaciones (en nuestra regla sólo especificamos uno de ellos, mediante el rasgo subcat, aunque pueden ampliarse). La parte de la izquierda de nuestra regla asigna, pues, ausencia de artículo a los sintagmas que cumplen las condiciones de la parte de la derecha.

## 6.2. *Verbos existenciales*

La oración de (49) ha de ser también objeto de tratamiento especial ya que en inglés este sintagma no contiene ningún actualizador, por lo que se hace necesario generarlo para el español. Obsérvense las siguientes oraciones:

- (57) a. Hay leones con la cola larga  
 b. Hay unos leones corriendo por el jardín  
 c. Unos perros están ladrando en el jardín.

Las traducciones al inglés de estas oraciones son las siguientes:

- (58) a. There are lions with a long tail  
 b. There are lions running in the garden  
 c. Dogs are barking in the garden.

Hay que hacer notar que las oraciones inglesas muestran dos tipos de predicación sobre individuos. (58a) adscribe propiedades más o menos permanentes a un número relevante de individuos. (58b) y (58c), por el contrario, predicán algo sobre unos individuos durante un corto período de tiempo. La diferencia entre la primera oración y las dos últimas se encuentra en la diferenciación que Carlson (1977) introdujo entre predicados de individuos (*individual predicates*) y predicados contingentes (*stage predicates*). Mientras en los primeros se predica una propiedad sobre un individuo que es la totalidad de sus partes (58a), en el segundo caso se predica sobre individuos entendiéndolos como porciones temporales de la existencia de esos individuos (*stages of individuals* en la terminología de Carlson) o realizaciones de esos individuos (58b y 58c). Haciendo uso de este análisis, podremos generar correctamente el artículo en unas oraciones y otras, ya que las gramáticas del inglés proporcionarán la información adecuada para estos dos tipos de predicados (atendiendo, por ejemplo, al tiempo y el aspecto verbal, tipo aspectual de los modificados).

res, etcétera). Esta información se copia de la lengua fuente a la lengua destino por defecto.

Las reglas que proponemos para la traducción de (58) contemplan dos tipos de situaciones. En primer lugar, nos referimos a aquellas oraciones en las que la predicación es introducida por verbos de tipo existencial (haber, existir, etc.). Estos verbos recibirán en el léxico el valor de exist en el rasgo tipo-v. La aplicación de estas reglas puede controlarse, pues, a través del tipo al que pertenece el verbo (FART EXIS 1 y FART EXIST 2). En segundo lugar, nos referiremos a aquellas oraciones en las que no hay ninguna marca mediante la cual podamos controlar en qué oraciones ha de aplicarse la regla FART EXIST 3. Dada la ambigüedad existente entre los sintagmas nominales ingleses en plural y las distintas predicaciones que introducen, se hace necesario un análisis más profundo que proporcione a las lenguas de salida (en nuestro caso, al español) el tipo de predicación implícito. Véanse, pues, las siguientes reglas en las que los valores de tipo-pred indican el tipo de predicación a la que aludíamos antes. La información necesaria, tal y como puede verse en las reglas que a continuación proponemos, se calcula en la predicación contenida en los modificadores (FART EXIST 1 y FART EXIST 2) o en el nodo principal atendiendo, como ya hemos indicado, al tiempo, al aspecto y a los modificadores que actúan en general sobre el nodo oracional (FART EXIST 3) <sup>10</sup>.

(59) ISES

FART EXIST 1 = {/cat = o}  
 [ {/cat = v, subcat = arg2, tipo-v = exist}  
 {art = Odet/cat = sn, rol = arg2},  
 {/tipo-pred = ind},  
 \* {} ].

FART EXIST 2 = {/cat = o}  
 [ {/cat = v, subcat = arg2, tipo-v = exist}  
 {art = indet/cat = sn, rol = arg2},  
 {/tipo-pred = stage},  
 \* {} ].

FART EXIST 3 = {/cat = o, tipo-pred = stage}  
 [ {/cat = v}  
 {art = indet/cat = sn, rol = arg1},  
 \* {} ].

<sup>10</sup> Un interesante estudio contrastivo sobre el uso del artículo en español y en inglés en oraciones genéricas y habituales puede encontrarse en Pease-Gorrisen (1980).

Los valores de tipo-pred son ind para *individual predicate* ((58a) y (57a)), y stage para *stage predicate* ((58b-c) y (57b-c)). Para el primer tipo de predicado se genera ausencia de artículo, mientras que para los segundos se genera el artículo indeterminado.

Por otra parte, las oraciones genéricas del tipo ya mencionado en (55), que se relacionan a su vez con los predicados de individuos, recibirán un tratamiento en un estadio posterior de la generación (cf. las reglas bajo (80)).

### 6.3 Verbos afectivos

Otra subclase de verbos que en castellano ha de recibir un tratamiento especial debido a que la lengua de origen de la que partimos los caracteriza sin actualizador son los verbos de tipo afectivo (*odiar, detestar, preferir, amar, respetar, etcétera*).

- (60) a. John hates oranges  
 a'. John detesta las naranjas  
 b. John prefers oranges to apples  
 b'. John prefiere las naranjas a las manzanas.

La regla bajo (61) proporciona la información necesaria que permitirá decidir el valor del rasgo art:

- (61) ISES  
 FART AFECT = {/cat = o}  
           [ {/cat = v, tipo-v = afect}  
           {/cat = sn, rol = arg1},  
           {art = det/cat = sn, rol = arg2},  
           ^ {art = det/cat = sn, rol = arg3},  
           \* {} ].

El símbolo ^ significa opcionalidad del nodo sn que ocupa la posición de tercer argumento. En relación con este nodo, la regla describe una situación que, por otra parte, es bastante general: si nos atenemos a una lista ordenada de complementos, el hecho de que el objeto aparezca con algún tipo de determinante, obliga a que el complemento oblicuo que le sigue aparezca también actualizado. Aunque en el caso que nos ocupa, la regla determina que en los verbos afectivos el español exige la presen-

cia del artículo determinado en todos sus argumentos. Nótese que los sintagmas subcategorizados por estos verbos no pueden recibir el tratamiento general que proponemos para los sintagmas en función de objeto directo que se han recibido del inglés sin actualizador. Sobre este tratamiento y las implicaciones que hemos comentado de paso, puede consultarse la regla FART OBJ de (80).

#### 6.4 *Verbos de posesión*

Verbos como *tener* o *llevar*, en la acepción de *vestir*, obligan a su complemento a un comportamiento muy diferente al que presenta el complemento de este verbo en inglés:

- (62) a. She has a nose  
 b. She has a long nose  
 c. She has a broken leg  
 d. She has a beautiful long nose  
 e. She has green eyes  
 f. She was wearing broken trousers.

Como hemos dicho anteriormente (ver secc. 4), el artículo indeterminado se copia de una lengua a otra a menos que el sintagma cumpla el rol de atributo del sujeto o del objeto, en cuyo caso no se copia (véase la regla FART 2 de (36)). Ahora bien, en los verbos de posesión, el artículo indeterminado de un sintagma cuyo núcleo es un sustantivo singular relacionado con las partes del cuerpo (62a) ha traducirse en español sin artículo siempre y cuando el sintagma no esté modificado (63a). Si el sintagma al que hacemos referencia apareciera modificado (62b-c), independientemente de si la parte del cuerpo a la que se hace referencia es única (*nariz*) o no (*pierna*), parece que lo mejor será mantener en español la indeterminación del artículo inglés (63b-c) puesto que tanto en inglés como en español, el artículo indeterminado es uno de los mecanismos para introducir nuevas entidades que no están disponibles en el discurso. Si junto a un sintagma de estas características aparece otro adjetivo modificador, más concretamente, un adjetivo valorativo, habrá de mantenerse dicho artículo indeterminado (cf. (62d) y (63d)). En (62e) se produce un hecho diferente: un sintagma plural sin actualizador cuyo núcleo refiere la clase semántica de parte del cuerpo, o no (62f), seguido de un modificador ha de producir en español un artículo determinado.

- (63) a. Tiene nariz  
 b. Tiene una nariz larga  
 c. Tiene una pierna rota  
 d. Tiene una bonita nariz larga  
 e. Tiene los ojos verdes  
 f. Llevaba los pantalones rotos.

Confróntese (63e) con el siguiente par de oraciones, en las que la ausencia de artículos se mantiene de una a otra lengua:

- (64) People have eyes  
 La gente tiene ojos.

Por otra parte, no puede olvidarse la relación que se establece en la lengua inglesa entre los usos de los pronombres posesivos y las partes del cuerpo. La utilización del artículo indeterminado puede darse siempre y cuando la parte del cuerpo no sea única (65a-b). El artículo definido es permitido en construcciones preposicionales donde el pronombre personal o el sustantivo que hace referencia al poseedor de la parte del cuerpo es el objeto y la parte del cuerpo es un complemento preposicional (65c).

- (65) a. He has broken a leg  
 Se ha roto una pierna  
 b. \*He has hurt a nose  
 \*Se ha lastimado una nariz  
 c. I shook him by the hand  
 Lo cogí de la mano.

En todos los ejemplos anteriores puede también aparecer el pronombre posesivo. El problema que se plantea, pues, es determinar cuál es la traducción al inglés del artículo determinado español junto a las partes del cuerpo. Somos de la opinión de que sólo las predicaciones contenidas en el discurso pueden proporcionarnos la traducción correcta.

- (66) a. Estábamos en el bar. Entraron cuatro hombres. Dos de ellos tenían (?el) un brazo roto  
 We were in the bar. Four men came in. Two of them had a broken arm  
 b. Estábamos en el bar. Entraron dos hombres mancos. Tenían el brazo roto  
 We were in the bar. Two one-armed men came in. Their arm was broken.

(66b) muestra que, puesto que los hombres eran mancos, el hecho de tener una sola mano es una información que está disponible, de tal forma que el uso del artículo determinado hace mención a una entidad ya disponible. Contrástese con (66a) donde para una persona con dos brazos, la mención con artículo determinado resulta al menos un poco extraña puesto que no tenemos disponible a cuál de los dos brazos nos referimos. Nótese el mecanismo mediante el cual el inglés evita la utilización del artículo determinado junto a una parte del cuerpo disponible en el discurso haciéndose servir para ello del pronombre posesivo.

Ante esta situación, lo más pragmático será mantener el artículo indeterminado siempre y cuando el sustantivo no sea una parte del cuerpo que aparezca sin modificadores. Este es el caso ya ejemplificado en (62a) y (63a). Por otra parte, el pronombre posesivo se traducirá por el artículo determinado. Compárense las siguientes oraciones, donde la proyección al artículo indeterminado en español es claramente agramatical en el sentido de que pierde la referencialidad presente en la oración inglesa:

- (67) He has his paper under his arm  
 Tiene el periódico debajo del brazo  
 \*Tiene un periódico debajo de un brazo.

Puesto que en el traductor isen-ises se copia el artículo indeterminado de una lengua a otra, hemos de marcar estos sintagmas objeto directo como pertenecientes a un tipo, al de los de posesión (pos).

(68) ISES

FART POS = {/cat = o}  
 [ {/cat = v, tipo-v = pos}  
 {/cat = sn, rol = arg1},  
 {tipo-n = pos/cat = n, rol = arg2},  
 \* {} ].

La regla de (68) constituye el paso previo para la aplicación de las reglas que aparecen bajo (69). Puesto que el cometido de las reglas será borrar un valor —los de art— e insertar uno nuevo, han de escribirse en un traductor que es el único componente lícito para ello.



(69) ISES-ERSES

FART POS1 = {cat = sn, tipo-n = pos, num = sing, art = indet}  
 [ {cat = n, sem = part-cuerpo},  
 \*{cat = sa; sp} ]  
 ⇒ {art = Odet}.

FART POS2 = {cat = sn, tipo-n = pos, num = pl, art = Odet}  
 [ {cat = n},  
 {cat = sa; sp},  
 \*{ } ]  
 ⇒ {art = det}.

La regla FART POS1 se aplicará exclusivamente a sintagmas cuyos núcleos sean sustantivos pertenecientes a la clase de partes del cuerpo; esto elimina a verbos como *llevar*; donde una oración como *she wore a hat* habrá recibido su traducción adecuada por la regla FART 2 de (36). El hecho de que la palabra *sombrero* no pertenezca a la clase semántica expresada en la regla no la hace candidata para que la regla se aplique sobre ella <sup>11</sup>.

<sup>11</sup> En español el verbo *tener* tiene un comportamiento especial pues en muchas ocasiones los sintagmas que aparecen junto a él han sufrido un proceso de incorporación, por lo que normalmente aparecen sin artículo.

- (i) a. Tiene pasaporte francés  
(He has a French passport)
- b. El apartamento tiene teléfono  
(The apartment has a telephone)
- c. El apartamento tiene un teléfono  
(The apartment has one telephone)

(i) nos muestra que en inglés es mucho más corriente el uso del artículo indeterminado junto al verbo al que nos referimos. Sin embargo, es difícil determinar cuándo el español elige incorporar su sintagma y cuándo no. Parece que son circunstancias externas a la oración y al discurso, esto es cuestiones socio-culturales, las que deciden este comportamiento. En una sociedad habituada a tener como animal de compañía a las serpientes, podría ser usual decir *tengo serpiente*. Ante esta situación, parece complicado determinar qué sustantivos pueden sufrir incorporación, pero una vez que estos se marcaran en el léxico con un rasgo especial, sería muy sencillo generar la ausencia de artículo donde el inglés presenta el artículo indeterminado, y mantenerlo en los demás casos. Con *llevar* también se producen incorporaciones. Nótese que en español junto a *llevaba un sombrero* puede aparecer también *llevaba sombrero*. Sin embargo, con otras prendas de vestir no se produce incorporación (*\*llevaba jersey, blusa, pantalón...*). Para esta cuestión, pueden consultarse los artículos de Klein (1976) y el ya clásico de Alonso (1974).

FART POS1 da cuenta de las diferencias entre (62a) y (63a). Además la regla explícita que junto al sustantivo no puede aparecer un modificador, puesto que éste posibilita otras proyecciones.

La regla FART POS2 da cuenta del contraste entre las oraciones de (62e-f) y (63e-f). En este contexto sólo se tiene en cuenta que el sintagma sea complemento de un verbo tipo-v = pos sin artículo, con número plural y un modificador.

En (70) se ejemplifica todos los casos en los que no es necesario explicitar una regla ya que el artículo se mantiene de una lengua a otra.

- (70) a. She has a long nose  
Tiene una nariz larga
- b. She has a beautiful long nose  
Ella tiene una (\*la) bonita nariz larga
- c. She has a very fat leg  
Tiene una pierna muy gorda
- d. She was wearing a blue/broken skirt  
Llevaba una falda azul/rota.

Para los posesivos, véase secc. 6.6.

### 6.5. *Atributos*

Los atributos de sujeto y objeto, en relación con la determinación, han de tratarse también como casos particulares debido a las diferencias entre una y otra lengua. Para ello se han diseñado las siguientes reglas:

(71) ISEN-ISES

FART ATR 1 = {cat = sn, rol = argAS; argAO, art = indet}  
[ {cat = n, sem = n-ag},  
\*{ } ]  
⇒ {art = indet}.

FART ATR 2 = {cat = sn, rol = argAS; argAO, art = indet}  
[ {cat = n, sem = n-ag},  
^ {cat = sa; sp, tipo-m = clas} ]  
⇒ {art = Odet}.

El inglés es mucho más regular en el uso del artículo indeterminado en sintagmas que desempeñan estas funciones sintácticas que el español. La regla FART ATR 1 se aplica a cualquier sintagma, que desempeñando los roles que se especifican (argAS o atributo del sujeto; argAO o atribu-

to del objeto), tengan artículo indeterminado. El núcleo de estos sintagmas ha de pertenecer a una clase semántica diferente a la de los *nomina agentis* (n-ag), es decir sustantivos no relacionados con profesiones. Esta regla se aplicará a oraciones como las siguientes:

- (72) a. This is a table  
 a'. Esto es una mesa  
 b. This is a large table  
 b'. Esto es una mesa larga.

En realidad, la regla FART ATR 1 copia el valor de art de una lengua a otra. Por el contrario FART ATR 2 transforma el valor del artículo indeterminado en la ausencia de artículo si el núcleo del atributo pertenece a la clase semántica de las profesiones.

- (73) a. She is a doctor  
 a'. (Ella) es médico  
 b. She is a forensic surgeon  
 b'. (Ella) es médico forense.

Los modificadores de tipo clas son aquellos que caracterizan el sustantivo como miembro de una clase (*médico de niños, profesor de universidad, etcétera*). Pero lo que en la regla determina la ausencia de artículo es el núcleo del sintagma ya que el sustantivo pertenece a la clase de las profesiones.

Por otra parte, cuando un sustantivo perteneciente a la clase n-ag aparece junto a un modificador que no es de tipo clasificador, el sintagma tendrá que presentar artículo indeterminado, sea o no modificado a su vez por un modificador de tipo clas. Véanse las siguientes oraciones y la regla bajo (75).

- (74) a. She is a good doctor  
 a'. (Ella) es una buena médico  
 b. She is a good forensic surgeon  
 b'. (Ella) es una buena médico forense.

(75) ISEN-ISES

FART ATR 3 = {cat = sn, rol = argAS; argAO, art = indet}  
 [ {cat = n, sem = n-ag},  
 ^ {cat = sa; sp, tipo-m = clas},  
 {cat = sa; sp, tipo-m = no clas} ]  
 ⇒ {art = indet}.

Las reglas anteriores se aplican únicamente a sintagmas singulares, ya que en plural ambas lenguas se comportan igual.

- (76) a. They are (good) doctors  
 a'. Son (buenos) médicos  
 b. They are forensic surgeons  
 b'. Son médicos forense.

Los sintagmas de las oraciones anteriores llegarán hasta el nivel de estructura de los constituyentes del español sin un valor para art. Una regla general especificará que todo sintagma sin artículo tendrá el valor Odet.

#### 6.6. *Posesivos*

Como es bien sabido, el inglés hace un uso mucho más extenso de los pronombres posesivos que el español. En inglés únicamente se permite la sustitución del pronombre posesivo por el artículo determinado cuando el pronombre personal o el nombre que hace referencia al poseedor aparece en la función de objeto directo y la parte del cuerpo es un complemento preposicional.

- (77) a. Every one gave us a pat on the back  
 a'. Todos nos dieron una patada en la espalda  
 b. I shook him by the hand  
 b'. Lo cogí de la mano.

En estos casos, y a efectos de traducción, el sintagma nominal mantendrá el valor del atributo art del inglés dada la regla FART 1 de (36). Sin embargo, la situación más normal es aquella en la que aparece el sintagma nominal que designa lo poseído precedido del pronombre posesivo correspondiente:

- (78) a. He broke his leg  
 b. Se ha roto la pierna.

Para estos casos, se propone la siguiente regla:

(79) ISEN-ISES

FART POSESIVO = S: {cat = o}

[V: {}],

SUJ: {pers = P, gen = G, num = N},

OBJ: {cat = sn}

[N: {cat = n},

-: {cat = sn, clas-n = pron, tipo = pos,  
pers = P, gen = G, num = N},

MOD: \*{rol = mod} ],

MODS: \*{ } ]

⇒

S < V,

SUJ,

OBJ:{art = det} < N, MOD > ,

{cat = sn, clas-n = cli, pers = P},

MODS > .

En cuanto a las condiciones que ha de verificar la oración inglesa, nótese que el sintagma en función de sujeto (el nodo SUJ) ha de concordar en número, persona y género con el sintagma pronominal (caracterizado por los valores de clas-n que determina que es un pronombre y por el valor de tipo que indica el tipo de pronombre: posesivo en este caso). Por otra parte, el sintagma nominal cuyo núcleo es un pronombre va precedido del símbolo  $-:$ , que significa que ha de borrarse en el objeto que obtendremos en español, cuya estructura se determina a la derecha de la regla. La parte de la derecha, pues, especifica que el valor det ha de insertarse en el nodo OBJ (que puede ser un sintagma nominal en función de objeto o un complemento oblicuo, ya que las preposiciones fuertemente regidas se contienen en el nodo sn). También se especifica que ha de insertarse en el objeto español un clítico (la forma *se*) que conserve el valor referencial de persona pero no los de género y número. El número no ha de copiarse nunca en las reglas de los posesivos porque mientras el castellano permite la concordancia en número del objeto poseído con el pronombre, el inglés sólo presenta una forma tanto para el singular como para el plural (*his book* vs. *his books*). Serán las gramáticas del español las que se ocupen de asignar el número correcto al pronombre. Existirían reglas similares para casos como *I put your hat on the table* (*Te he puesto el sombrero encima de la mesa*) o *He has got his paper under his arm* (*Tenía el periódico bajo el brazo*).

### 6.7. *Funciones sintácticas*

A lo largo de este artículo hemos mencionado en varias ocasiones que el español exige la determinación en los sintagmas nominales en función de sujeto. Hemos contemplado además los casos que no podían tratarse con la regla general que a continuación proponemos (verbos presentativos y existenciales). Puesto que las excepciones han sido tratadas en el nivel de interficie del español, estamos ya en condiciones de poder establecer reglas generales para aquellos sintagmas que aún no han recibido un valor en el atributo art. Se trata ya solamente de aquellos sintagmas que en inglés han aparecido sin actualizador: los sujetos genéricos y los sintagmas en función de objeto. Estos casos son tratados por las reglas bajo (80).

(80) ERSES

FART SUJ = {art = det/cat = sn, fs = suj; gen-suj}.

fart obj = {art = Odet/cat = sn, fs = obj; gen-obj}.

La regla FART SUJ da cuenta de las siguientes oraciones:

- (81) a. Japanese eat raw fish  
 a'. Los japoneses comen pescado crudo  
 b. Dogs bark  
 b'. Los perros ladran
- (82) a. The arrival of trains  
 a'. La llegada de los trenes
- (83) a. People have eyes  
 a'. La gente tiene ojos  
 b. The production of cars  
 b'. La producción de coches.

En las reglas de (80) se especifica que la función sintáctica del sintagma ha de ser bien sujeto de una oración (suj, como en las oraciones de (81)), bien sujeto de un sintagma nominal (gen-suj, como en el sintagma (82)), o bien objeto (obj, como en las oraciones (83a) y (81a), o gen-obj como en (83b)). Esta regla filtrará oraciones que todavía no han recibido un valor en la determinación a causa de que en inglés aparecen

sin actualizador. Piénsese en oraciones como *John likes oranges*. Tras un proceso de transformación, esta oración recibiría un objeto de salida como el siguiente:

(84) O [ V, SN[art = , num = pl, fs = suj], SN[fs = iobj] ]

La representación de (84) se adecua a la descripción de la regla FART SUJ de (80) de tal forma que, al aplicarse sobre ella, el sintagma sujeto recibirá artículo determinado.

La regla FART OBJ de (80) determina que un sintagma en función de objeto, ya sea de una oración ya sea de un sustantivo predicativo (gen-obj), que haya llegado a este nivel relacional sin especificación para el atributo art, ha de ser Odet. En realidad, a este nivel sólo han podido llegar los sintagmas sin actualizador que en inglés aparecían sin él. Esta regla afectará también a oraciones como las siguientes:

- (85) a. He poured water into the glass  
 a'. Echó agua en el vaso  
 b. Houses are sold  
 b'. Se venden casas.

(85b) se verá afectada por esta regla ya que la pasiva refleja —o cualquier otra pasiva— es deshecha en los niveles superiores por lo que el sujeto superficial recibe la función de objeto. Recuérdese que si este sintagma tuviera un actualizador, pasaría por las reglas del traductor ISEN-ISES FART 1 y FART 2 de (36).

Por otra parte, hay que hacer notar que el español admite con mayor facilidad la ausencia de artículo en sintagmas en función de objeto directo (en singular si el sustantivo es de masa; en plural si el sustantivo es contable) que en función de sujeto. Si bien es cierto que el comportamiento del artículo en otras funciones es algo irregular, podemos decir que hay una tendencia general por la que los objetos indirectos suelen ir con artículo pero, necesariamente, lo exigirán si el objeto directo lleva algún tipo de determinante. En cuanto a los complementos oblicuos, parece que cuanto más ligado esté, menos obligado es el uso del artículo y, por lo tanto, depende más de las características internas de ese complemento. Compárense las siguientes oraciones:

- (86) a. \*Entregué cartas a estudiantes  
b. \*Entregué las cartas a estudiantes  
c. \*Enviamos mensajes a planetas  
d. \*Recogimos agua de suelo
- (87) a. Entregué cartas a unos/los estudiantes  
b. Entregué las cartas a unos/los estudiantes  
c. Enviamos mensajes a los planetas  
d. Recogimos agua del suelo.

(86a) y (87a) muestra la preferencia del objeto indirecto por la determinación. (86b) muestra que si el objeto está determinado el objeto indirecto debe estarlo también. Los dos últimos ejemplos indican que el sintagma preposicional al ser un complemento que indica un lugar o un destino necesitan la actualización, al margen de que los sustantivos sean de tipo masa o contables.

## 7. Conclusiones

Las proyecciones propuestas del artículo inglés al español están basadas principalmente en razonamientos de tipo sintáctico, por las razones ya mencionadas en la introducción. Empezamos proponiendo una regla general que copia el artículo del inglés al español, a excepción de los casos mencionados (atributos y coaparición de algún cuantificador). Puesto que el mayor número de irregularidades se presentan en la ausencia de artículo en inglés y la presencia en español de artículo determinado, hemos propuesto una proyección que incidía en los aspectos internos del sintagma nominal (características especiales de algunos sustantivos, coaparición de otros determinantes, etc.) y en los externos (factores que determinan la actualización o no del sintagma atendiendo a las funciones sintácticas, tipos de verbos, etc.).

El tratamiento propuesto tiene la ventaja de poder ser ampliado sin que ello afecte a la propuesta general. Sin embargo presenta limitaciones en lo concerniente a la interpretación semántica y pragmática ya que ésta requeriría el estudio de estos aspectos del sintagma nominal junto a los de otros constituyentes oracionales.



## BIBLIOGRAFÍA

- Allegranza, V., P. Bennett, J. Durand, F. van Eynde, L. Humphreys, P. Schmidt y E. Steiner (1991): «Linguistics for Machine Translation: The Eurotra Linguistic Specifications», en Copeland, C., J. Durand, S. Krauwer y B. Maa-gaard (eds.), *The Eurotra Linguistic Specifications*, Office for Official Publications of the European Communities, Luxemburgo.
- Alonso, A. (1974): «Estilística y gramática del artículo en español», *Estudios Lingüísticos*, Madrid, Gredós, págs. 125-160.
- Carlson, G. N. (1977): «A unified analysis of the English bare plural», *Linguistics and Philosophy*, 1, págs. 413-457.
- Cerdá, R., T. Badia, S. Balari y J. Vidal (1990): «El tratamiento computacional de la lengua española en el proyecto de traducción automática Eurotra», en *Actas del Congreso de la Sociedad Española de Lingüística. XX Aniversario*, Tenerife, 2-6 de abril de 1990, págs. 137-147.
- Durand, J., P. Bennett, V. Allegranza, F. van Eynde, L. Humphreys, P. Schmidt y E. Steiner (1991): «The Eurotra Linguistic Specifications: An Overview», *Machine Translation*, 6, págs. 103-147.
- Van Eynde, F. (1988): «The analysis of tense and aspect in Eurotra», en *Proceedings of the 12th International Conference on Computational Linguistics. Coling '88*, John von Newman Society for Computing Sciences, Budapest, págs. 699-704.
- Klein, F. (1976): «Same vs. different crosslinguistically: the 'articles' in English and in Spanish», *Papers from the Twelfth Regional Meeting of the Chicago Linguistic Society*, 12, págs. 413-424.
- Paese-Gorrissen, M. (1980): «The use of the article in Spanish habitual and Generic Sentences», *Lingua*, 51, págs. 311-336.
- Quirk, R., S. Greenbaum, G. Leech y Svartvik, J. (1985): *A Comprehensive Grammar of the English Language*, Londres, Longman.