

El Tratamiento de la Subcategorización Argumental en los Nominales Deverbales Árabes

Elisabeth Nebot¹, J.A.Alonso²

Resumen

Este artículo versa sobre una clase de nombres deverbales árabes, llamados "másdares" (o nombres verbales), cuyo comportamiento sintáctico se asemeja en gran medida al de los verbos finitos, y para los que, al igual que en éstos, es conveniente utilizar información de subcategorización argumental con el fin de tratarlos adecuadamente en aplicaciones de TA. El hecho de que numerosas construcciones que en árabe se hacen con "másdares" aparezcan traducidas con verbos finitos en la correspondiente frase en castellano, hace que sea vital tener un análisis correcto de la frase árabe en lo que concierne a adjunción de constituyentes y asignación de roles temáticos. Intentaremos explicar qué son y cómo se comportan estos deverbales árabes, cómo influye la información de subcategorización argumental durante su proceso de traducción y qué tratamiento se les está dando en el sistema de TA árabe-español que actualmente se está desarrollando en la Universidad de Barcelona.

1. Introducción

En la mayoría de aplicaciones de procesamiento de lenguaje natural que necesitan llevar a cabo un análisis sintáctico (sistemas de TA, interfaces en lenguaje natural, sistemas de consulta de información, etc.) es práctica común utilizar la información léxica de subcategorización argumental de los verbos para poder realizar correctamente una serie de pasos, como la adjunción de sintagmas nominales y preposicionales a sus núcleos correspondientes, la resolución de anáforas o la asignación de roles a los argumentos del verbo para la posterior desambiguación léxica.

¹ Departamento de Arabe (UB); email: nebot@lingua.fil.ub.es

² INCYTA S.L.; email: juan@incyta.es

Sin embargo, no existe un consenso en cuanto a la conveniencia de emplear información de subcategorización argumental en el caso de núcleos nominales y, en el supuesto de ser empleada, sobre cómo hacerlo.

En este artículo, veremos cómo en Árabe Estándar Moderno (MSA) existe una clase de nombres deverbales, llamados "*másdares*" o nombres verbales (cf. Anexo I) cuyo comportamiento sintáctico se asemeja en gran medida al de los verbos finitos, y para los que, al igual que en éstos pero con algunas diferencias de tratamiento, es conveniente utilizar información subcategorial para realizar un análisis correcto.

El hecho de que numerosas construcciones que en árabe se hacen con *másdares* aparezcan con verbos finitos en la correspondiente traducción en castellano hace que sea vital para un sistema de TA árabe-español tener un análisis correcto de la frase árabe en lo que concierne a adjunción de constituyentes y asignación de roles temáticos.

En la páginas que siguen intentaremos explicar qué son y cómo se comportan estos deverbales árabes, qué problemas plantea su traducción al castellano y qué tratamiento se les está dando en el sistema de TA árabe-español que actualmente se está desarrollando en la Universidad de Barcelona.

A lo largo de este artículo nos referiremos a la subcategorización argumental de núcleos verbales y nominales considerando tres niveles diferentes:

- Nivel semántico: roles temáticos profundos (agente, paciente, experimentador,..).
- Nivel funcional: función sintáctica de los argumentos (sujeto, objeto directo, complementos de los nombres).
- Nivel superficial de marca: categoría sintáctica, marca de caso y preposición.

Los núcleos verbales, y determinados núcleos nominales y adjetivales, subcategorizan argumentos que desempeñan roles temáticos, ocupan determinadas posiciones funcionales sintácticas y se marcan con un determinado caso o preposición.

2. ¿Qué son y cómo se comportan los *másdares* en árabe?

2.1 *Másdares: definición y tipos.*

Tradicionalmente, se considera que un *másdar* expresa la acción, estado o proceso, indicados por un verbo o el resultado de dicha acción, estado o proceso. De hecho, el vocablo árabe *másdar* significa, literalmente, "el lugar en el cual algo se origina", ya que los primeros gramáticos árabes derivaban la idea compuesta del verbo finito a partir de la idea simple de este sustantivo.

Todo *másdar* tiene dos interpretaciones: una como *nominal de proceso* y otra como *nominal de resultado*. Considerando los niveles funcional y superficial, los primeros tienen una estructura argumental de tipo "verbal" mientras que los segundos tienen una estructura argumental de tipo "nominal". No obstante, consideraremos que, a nivel semántico, ambos subcategorizan los mismos roles temáticos que el verbo del cual se derivan.

Por otra parte, un *másdar* puede traducirse en castellano como un deverbal, un infinitivo o como una forma finita del verbo correspondiente. Parece que el hecho de

que el *másdar* sea de proceso o de resultado determina en gran medida la forma en que debe traducirse (como verbo o como nombre).

Aunque las interpretaciones procesal y de resultado de los *másdares* árabes tienen diferente comportamiento sintáctico, veremos que sigue siendo útil utilizar la Información de subcategorización del verbo correspondiente (derivada mediante reglas léxicas) para, en la fase de transferencia, desambiguar las diferentes lecturas y conseguir un transfer léxico correcto.

2.2 Características morfológicas de los *másdares*.

Se considera que los *másdares* árabes ("nomina actionis" o "nomina verbi" para los orientistas occidentales) son nombres formados a partir de la raíz verbal aplicando unos determinados patrones. Así, la raíz verbal "ktb" en Forma I "kataba" ("escribir") forma un *másdar* con patrón 1a23 "katb" ("el hecho de escribir", "la escritura") y la misma raíz pero en Forma III "ka'taba" ("mantener correspondencia") forma el *másdar* correspondiente con patrón *mu1a'2a3a#* "muka'taba#" ("el hecho de mantener correspondencia", "la correspondencia").

El primer problema que se plantea con los *másdares de proceso* es cómo distinguirlos de los *másdares de resultado* ya que no hay diferencias morfológicas.

La comparación entre estos dos tipos de nombres deverbales muestra que difieren sustancialmente en sus propiedades y estructuras (cf. Fassi Fehri, *Issues in the Structure of Arabic Clauses and Words*).

Los principales rasgos distintivos de cada tipo pueden resumirse como sigue:

Másdares de Proceso:

- asignan marca de acusativo a sus objetos
- subcategorizan argumentos con posición funcional verbal y roles temáticos
- aceptan modificadores adverbiales de manera
- aceptan ser modificados por una cláusula de relativo
- pueden ser complementos en una estructura de control
- no admiten adjetivos como modificadores cuando asignan marca de caso acusativo
- no pueden flexionar en dual ni en plural
- no pueden ir precedidos de demostrativos
- no pueden formar diminutivo

Másdares de Resultado:

- aceptan modificadores adjetivales
- pueden flexionar en dual y plural
- aceptan demostrativos
- pueden formar diminutivo
- aceptan ser modificados por una cláusula de relativo
- no asignan marca de caso
- subcategorizan argumentos con posición funcional nominal y roles temáticos
- no admiten modificadores adverbiales de manera
- aceptan modificadores adjetivales

En los ejemplos que siguen se ilustran algunas de estas diferencias:

- (1) ?aqlaqa-niy ʔiḥtiqa'ra't-u-hu
preocupó-me desprecio-f-PL-NOM-su
"Sus desprecios me preocuparon"
- (2) *?aqlaqa-niy ʔiḥtiqa'ra't-u zayd-in yuwsuf-an
preocupó-me desprecio-f-PL-NOM Zayd-GEN Yusuf-ACC
**"Los desprecios de Zayd a Yusuf me preocuparon"
- (3) ?aqlaqa-niy ʔiḥtiqa'r-u zayd-in yuwsuf-an
preocupó-me desprecio-f-SG-NOM Zayd-GEN Yusuf-ACC
"Me preocupó (el hecho de) que Zayd despreciara a Yusuf"
- (4) yuriydu 'lḥuṣuwl-a ʕalaà 'lʔa'īza#-i bi-surʔat-in
quiere-él la-consecución-ACC sobre el-premio-GEN con-rapidez-GEN
"Quiere conseguir el premio rápidamente"
- (5) *yuriydu 'lḥuṣuwl-a't-a bi-surʔat-in
quiere-él la-consecución-PL-ACC con-rapidez-GEN
**"Quiere las consecuciones rápidamente"

En el ejemplo (1), el *másdar* expresa un *resultado* y se somete a pluralización, proceso éste puramente nominal y que es bloqueado por el *másdar de proceso* en (2). El núcleo del NP en (3) es un *másdar de proceso* que toma un agente y un paciente como argumentos y que, como se ve en (4), autoriza adverbios de manera que son modificadores verbales. En (5), el *másdar de resultado* bloquea la modificación adverbial.

2.3 Características sintácticas de los másdares.

Sintácticamente, el *másdar de proceso* puede regir verbalmente, es decir, en acusativo, un nombre como objeto directo. Normalmente a un verbo transitivo le corresponde un *másdar* transitivo y a un verbo intransitivo le corresponde un *másdar* intransitivo. Asimismo, en aquellos verbos que se integran semánticamente con una preposición (preposición de régimen), la construcción del *másdar* la conserva.

Vamos a ver con más detalle el comportamiento sintáctico de los *másdares* de proceso, especialmente en lo que respecta a su argumento externo (correspondiente al sujeto sintáctico del verbo finito) y a sus argumentos internos (correspondientes a los objetos del verbo finito).

2.3.1 Argumento Externo:

El *másdar de proceso* puede o no tener argumento externo. El término alternativo "complemento agente" no nos parece el apropiado dado que el rol temático de dicho argumento no siempre es el de 'agente' sino que puede ser el de 'paciente' o el de 'experimentador' no siendo, por consiguiente, necesariamente actorial.

El argumento externo, si aparece, suele estar en genitivo, a menos que siga al objeto directo, en cuyo caso estará en nominativo. En otras palabras, el *másdar*, como todo

nombre, puede poner en genitivo al nombre que le sigue ("¿ida'fa#" o anexión) y éste, como veremos más adelante, es uno de los problemas con que nos enfrentamos a la hora de traducir un NP constituido por un núcleo nominal -que podría ser un *másdar-* y un nombre en anexión.

- (6) fa'ya?a-niy ¿i'tira'f-u 'lqa'til-i
sorprendió-me confesión-NOM el-asesino-GEN
 (i) "Me sorprendió que el asesino confesara"
 (ii) "Me sorprendió la confesión del asesino"

- (7) fa'ya?a-niy ¿i'tira'f-u 'lqa'til-i bi-'lqanb-i
sorprendió-me confesión-NOM el-asesino-GEN con-el-crímen-GEN
 "Me sorprendió que el asesino confesara el crimen"

En (6) el *másdar* tiene dos interpretaciones posibles: una (i) como *proceso* y otra (ii) como *resultado*, mientras que en (7) sólo es posible una lectura como *proceso* en la que el *másdar* "¿i'tira'f" subcategoriza un paciente con función de objeto preposicional (con preposición de régimen "bi"), además de un complemento agente. El *másdar de resultado* no subcategoriza argumentos internos con posiciones funcionales verbales (sólo roles temáticos) y, en cualquier caso, nunca está permitida la construcción "*másdar de resultado* + subj + dobj". O bien se da la construcción "*másdar de resultado* + subj", o bien se da la construcción "*másdar de resultado* + dobj".

2.3.2 Argumentos Internos:

Cuando sólo aparece un objeto siguiendo inmediatamente al *másdar de proceso*, el objeto recibe marca de caso genitivo:

- dobj (np GEN)

- (8) ha'wala ¿inka'r-a 'lwa'qi'a#-i
intentó-él negación-ACC la-realidad-GEN
 "Intentó negar la realidad (lo acaecido)"

Cuando el argumento externo (sujeto) es explícito, el objeto recibirá marca de caso acusativo si le sigue y marca de genitivo si le precede.

- subj (np GEN) + dobj(np ACC)
- dobj (np GEN) + subj(np NOM)

- (9) a. ?aqlaqa-niy ¿inka'r-u ¿ariq-in 'lwa'qi'a#-a
preocupó-me negación-NOM Tariq-GEN la-realidad-ACC
 "Me preocupó que Tariq negara la realidad"
- b. ?aqlaqa-niy ¿inka'r-u 'lwa'qi'a#-i ¿ariq-un
preocupó-me negación-NOM la-realidad-GEN Tariq-NOM
 "Me preocupó que Tariq negara la realidad"

En predicados de dos argumentos, se permiten dos construcciones:

- iobj (np GEN) + dobj (np ACC)
- iobj (np ACC) + dobj (np ACC)

(10) ta^ʕliym-u 'lʔulla'b-i 'l^ʕuluwm-a
enseñanza-NOM los-alumnos-GEN las-Ciencias-ACC
"El hecho de enseñar (la enseñanza de) las Ciencias a los alumnos"

(11) ta^ʕliym-u 'lʔulla'b-a 'l^ʕuluwm-a
enseñanza-NOM los-alumnos-ACC las-Ciencias-ACC
"El hecho de enseñar (la enseñanza de) las Ciencias a los alumnos"

Otras construcciones permitidas con *másdares de proceso* son las siguientes:

- dobj (cci: cláusula completiva) ⇨ (12)
- dobj (mm = "maf^ʕuwl muʔla'q" o complemento absoluto) ⇨ (13)
- pobj1 (prep. régimen + np GEN) ⇨ (14)
- pobj1 (prep. de neutralización "li" + np GEN) ⇨ (15)

(12) ʔi^ʕla'n-u 'lmudiy-r-i ʔanna 'lnaʔa'h-a qariyb-un
anuncio-NOM el-director-GEN que el-éxito-ACC (está) cercano-NOM
"El anuncio del director de que el éxito está cerca"
("El hecho de que el director anuncie que el éxito está cerca")

(13) ɖarb-u zayd-in ɖarb-an šadiyd-an
golpe-NOM Zayd-GEN un-golpe-ACC fuerte-ACC
"El hecho de que Zayd golpee fuertemente"

(14) yuqliqu-niy baḥt-u-hu ʕan 'lnaʔa'h-i
preocupa-me búsqueda-NOM-su el-éxito-GEN
"Me preocupa que (él) persiga el éxito (su búsqueda del éxito)"

(15) našr-u 'lraʔul-i li-lkita'b-i ʔamr-un gariyb-un
publicación-NOM el-hombre-GEN para-el-libro-GEN (es) un-asunto-NOM
extraño-NOM
"Que el hombre publique el libro es extraño"
("La publicación del libro por parte del hombre es un asunto extraño")

También son posibles las combinaciones entre ellas (aunque siguiendo algunas reglas).

2.3.3 Neutralización de marcas temáticas superficiales

Hay dos construcciones diferentes con *másdares de proceso*: en una de ellas, el *másdar* asigna directamente marca de acusativo al argumento interno mientras que en la otra, esta rección se neutraliza recurriendo a la preposición "li".

Se habrá observado que, internamente, algunas construcciones con *másdares de proceso* parecen a primera vista más "verbales" que otras:

- (16) fa'ya?a-niy qawl-u zayd-in 'lhaqq-a
sorprendió-me el-decir-NOM Zayd-GEN la-verdad-ACC
 "Me sorprendió que Zayd dijera la verdad"
- (17) fa'ya?a-niy qawl-u zayd-in li-lhaqq-i
sorprendió-me el-decir-NOM Zayd-GEN para-la-verdad-GEN
 "Me sorprendió que Zayd dijera la verdad"

En ambas oraciones, el *másdar* expresa un proceso o evento y podría ser modificado por un adverbio de manera. Pero en (16), el objeto recibe marca de acusativo, lo que le hace parecerse más al verbo transitivo del cual se deriva y en (17) el objeto recibe caso genitivo mediante la preposición (no de régimen) "li".

En Arabe Estándar Moderno (MSA), la rección casual en acusativo por parte del *másdar* se está perdiendo progresivamente en favor de la construcción del *másdar* con la preposición "li" (excepcionalmente la preposición "ʿilaa") que rige genitivo.

Esta neutralización obedece, entre otras, a la tendencia natural de los que se expresan en lengua árabe a facilitar la comprensión de la lengua escrita evitando este tipo de construcciones nominales ambiguas mediante la introducción de marcas superficiales, como la preposición "li" que indica, de forma excluyente, que la posición funcional del argumento subcategorizado por el *másdar* mediante dicha preposición es la de objeto.

3. La subcategorización argumental de los *másdares*

3.1 Ejemplos

Veamos los siguientes ejemplos:

- (18) a. ?amara 'lmalik-u 'lruḍaʔ-a bi-'lwaziyr-i bi-'ljuruwý-i
ordenó el-rey-NOM los-satisfechos-ACC con el-visir-GEN con la-salida-GEN
 "El rey ordenó salir a los que estaban satisfechos con el visir"
 "**El rey ordenó el visir a los satisfechos con la salida"
- b. ?amr-u 'lmalik-i 'lruḍaʔ-a bi-'lwaziyr-i bi-'ljuruwý-i
orden-NOM el-rey-GEN los-satisfechos-ACC con-el-visir-GEN con-la-salida-GEN
 "El hecho de que el rey ordene salir a los que están satisfechos con el visir"
 ("La orden por parte del rey de salir a los que están satisfechos con el visir")
 "**El hecho de que el rey ordenara el visir a los satisfechos con la salida"
- (19) a. 'lṭa'lib-u yaruddu ʿalaà-lmudarris-i
el-alumno-NOM responde sobre-el-profesor-GEN
 "El alumno responde al profesor"
- b. radd-u 'lṭa'lib-i ʿalaà-lmudarris-i
respuesta-NOM el-alumno-GEN sobre-el-profesor-GEN
 "El hecho de responder el alumno al profesor"
 ("La respuesta del alumno al profesor")
- c. 'lṭa'lib-u yaruddu ʿalaà-ltawā'liy
el-alumno-NOM responde sobre-la-sucesión-GEN
 "El alumno responde ininterrumpidamente"

- d. radd-u 'lta'lib-i 'alaà-Itawa'liy
 respuesta-NOM el-alumno-GEN sobre-la-sucesión-GEN
 "El hecho de responder el alumno ininterrumpidamente"

En (18) tenemos un problema de adjunción de PP tanto con el verbo finito (18a) como con el *másdar* (18b). El verbo y el *másdar* tienen una preposición de régimen "bi" pero el complemento objeto de persona ("lruḍa'f") es, a su vez, un deverbial (participio activo) con preposición de régimen "bi". Sólo la información acerca de las estructuras argumentales correspondientes puede ayudarnos a analizar bien esta frase. En ambos casos, verbo y *másdar*, las construcciones son paralelas y los problemas también: el argumento preposicional, con preposición "bi" regido por el verbo "ʔamara" debe ser no humano, y por lo tanto sólo puede ser "bi-l-juruwý-i" ("con la salida") y nunca podría ser "bi-l-waziyr-i" ("con el visir").

Lo mismo sucede en (19a,b) y (19c,d), en los que el hecho de asignar o no rol temático al complemento preposicional, cambia el significado.

Para poder analizar bien estas frases, necesitamos, pues, utilizar información sobre la subcategorización argumental del *másdar*. No obstante, los argumentos de los *másdares* tienen particularidades que no tienen los de los verbos, de manera que el tratamiento no puede ser el mismo aunque la estructura argumental sí lo sea.

3.2 Diferencias entre *másdares* y verbos finitos

3.2.1 Obligatoriedad de la subcategorización argumental.

Los verbos, si subcategorizan argumentos, deben ocurrir siempre con ellos. Los *másdares*, sólo opcionalmente. (cf. Fassi Fehri)

- (20) a. *wahaba Zayd-un
 dio-él Zayd-NOM
 **Zayd donó
- b. šakara wahb-a zayd-in
 agradeció-él donación-ACC Zayd-GEN
 "Agradeció la donación de Zayd"

Aquí, mientras que (20a) es agramatical, (20b) no lo es. La construcción sin objeto directo del verbo finito es del todo incorrecta pero no así la construcción con el *másdar* correspondiente.

3.2.2 Neutralización de marcas morfológicas en los *másdares*.

a) Neutralización mediante la preposición "li":

- (21) a. qawl-u zayd-in 'lhaqq-a
 el-decir-NOM Zayd-GEN la-verdad-ACC
 "El hecho de que Zayd diga la verdad"
- b. qa'la zayd-un 'lhaqq-a
 dijo-él Zayd-NOM la-verdad-ACC
 "Zayd dijo la verdad"

- c. *qawl-u zayd-in li-lḥaqq-i*
el-decir-NOM Zayd-GEN para-la-verdad-GEN
 "El hecho de que Zayd diga la verdad"
- *d. *qa'la zayd-un li-lḥaqq-i*
dijo-él Zayd-NOM para-la verdad-GEN
 "Zayd dijo la verdad"

El ejemplo (a) ilustra una frase con un *másdar* del verbo "qa'la" cuyo objeto directo aparece marcado en acusativo, de igual forma que en la correspondiente versión con la forma finita del verbo (b). El ejemplo (c) es otra versión posible de (a) en la cual la marca de objeto directo pasa de ser caso acusativo a ser la preposición "li". Sin embargo, esta neutralización es imposible con la forma finita del verbo (d).

b) Neutralización mediante la construcción en anexión que hace indistinguibles el genitivo objetivo del genitivo subjetivo:

- (22) a. *ḥintaqada 'irayul-u*
criticó-él el-hombre-NOM_(AGT)
 "El hombre criticó"
- b. *ḥintiqa'd-u 'irayul-i*
crítica-NOM el-hombre-GEN_(AGT/PAT)
 (i) "La crítica del hombre (hecha al hombre)"
 (ii) "La crítica del hombre (hecha por el hombre)"

Mientras que en (22a) no hay duda de que "'irayulu" es el sujeto o agente de la acción por el orden superficial de la frase (VSO) y la marca de nominativo, en la construcción con *másdar* (22b) no podemos saber si se trata de un genitivo objetivo o de un genitivo subjetivo y tendremos la posibilidad de asignar a "'irayuli" rol de paciente (i) o de agente/experimentador (ii). Teniendo en cuenta que el sistema trabaja con árabe no vocalizado, además de la dicotomía 'genitivo subjetivo/genitivo objetivo', que es la que aquí se ilustra, cabría la posibilidad de que "'irayul", tanto en (22a) como en (22b) fuera objeto y estuviera marcado en acusativo.

Aunque puede darse el caso de que una construcción con *másdar* omita la preposición de régimen del verbo correspondiente y aparezca en anexión con el objeto (en genitivo) -siendo éste un tipo de neutralización menos frecuente- la lengua árabe prefiere conservar siempre que sea posible dicha preposición para evitar muchas ambigüedades.

- (23) *'riḍāà 'an-'lmuṣlimiyna*
la-satisfacción acerca-los-musulmanes-GEN_(PAT)
 "La satisfacción por los musulmanes (a causa de ellos)"
- (24) *riḍāà 'lmuṣlimiyna*
satisfacción los-musulmanes-GEN_(PAT/EXP)
 (i) "La satisfacción por los musulmanes (a causa de ellos)"
 (ii) "La satisfacción de los musulmanes (sentida por ellos)"

Así, mientras que en (23) el PP subcategorizado por el *másdar* sólo puede ser paciente (PAT), en (24) cabría una interpretación del NP "Imuslimiyna" como paciente (PAT) en (i) y otra como experimentador (EXP) en (ii).

3.2.3 Neutralización de la oposición activa/pasiva en los másdares.

El hecho de que en los *másdares* se neutralice la oposición activa/pasiva (cosa que no ocurre en los verbos) da lugar a ciertas ambigüedades que únicamente el contexto puede contribuir a desambiguar.

El siguiente NP con el *másdar* "ta^cliym":

(25) ta^cliym-u 'Iraÿul-i 'f'uluwm-a
enseñanza-NOM el-hombre-GEN_(GO/AGT) las-Ciencias-ACC_(PAT)

puede tener dos posibles traducciones:

- (i) "La enseñanza de las Ciencias por parte del hombre"
- (ii) "La enseñanza de las Ciencias al hombre".

según exprese una relación predicativa activa (i):

(25a) ^callama 'Iraÿul-u 'f'uluwm-a
enseñó-él el-hombre-NOM_(AGT) las-Ciencias-ACC_(PAT)
"El hombre enseñó Ciencias"

o exprese una relación predicativa pasiva (ii):

(25b) ^cullima 'Iraÿul-u 'f'uluwm-a
fue-enseñado-él el-hombre-NOM_(GO) las-Ciencias-ACC_(PAT)
"El hombre fue enseñado Ciencias"

3.2.4 Argumentos pronominales

Los *másdares* pueden tomar argumentos pronominales internos y externos. Los verbos sólo admiten argumentos internos pronominales. Así, un pronombre personal clítico sufijado a un verbo finito puede corresponder únicamente a su objeto directo o indirecto. En cambio los clíticos sufijados a un *másdar* pueden corresponder tanto al objeto (directo o indirecto) como al sujeto.

- (26) a. ¿intaqada-hu
criticó-lo_(PAT)
"Le criticó"
- b. ¿intiqad-u-hu
crítica-NOM-su_(AGT/PAT)
(i) "El hecho de que se le critique (su crítica)"
(ii) "El hecho de que él critique (su crítica)"

- (27) a. * ζ intaqada-hu zayd-an
criticó-él-su-(AGT) Zayd-ACC
 "Criticó-él/su a Zayd"
- b. ζ intiqā'd-u-hu zayd-an
crítica-NOM-su_(AGT) Zayd-ACC_(PAT)
 "Su crítica a Zayd (que él criticara a Zayd)"

En (26a) el pronombre es un objeto directo que desempeña el rol de paciente pero en (26b) tanto puede ser un sujeto que desempeña el rol de agente (ii) como un objeto que desempeña el rol de paciente (i).

La función de sujeto con rol agente del clítico pronominal sufijado al *másdar* en (27b) se ve reforzado por la posibilidad de ir seguido por un NP con marca de acusativo que desempeña el rol de paciente.

Contrariamente a lo que ocurre con las construcciones con formas verbales finitas, los *másdares*, si subcategorizan un objeto pronominal, no admiten después el sujeto explícito (28). En otras palabras, un *másdar* no admite el orden M(ásdar)-O_{pron}-S (el verbo sí lo admite: V-O_{pron}-S).

- (28) a. ζ intaqada-hu zayd-un
criticó-éHo zayd-NOM_(AGT)
 "Zayd le criticó"
- *b. intiqā'd-u-hu zayd-un(in)
crítica-NOM-su_(PAT) zayd-NOM(GEN)_(AGT)
 "La crítica de Zayd a él (El hecho de que Zayd le critique)"

Todo lo anterior muestra que las marcas superficiales en la estructura argumental de un verbo finito están claras (doj=acc, subj=nom...) pero las de los argumentos de un *másdar* pueden ser ambiguas ya que varios argumentos diferentes pueden ir marcados en genitivo. La identificación de los argumentos en el caso de los *másdares* es, por consiguiente, más problemática y, además de la información morfológica superficial (categoría, orden, caso y preposiciones), necesitaremos utilizar información semántica.

4. Tratamiento computacional

La traducción de estos nombres deverbales plantea dos problemas computacionales básicos en lo que respecta a la TA:

- *la desambiguación léxica:*

- (29) $\text{ḍarb-u zayd-in } ^{\text{c}}\text{amr-an}$
golpe-NOM Zayd-GEN_(AGT) Amr-ACC_(PAT)
 "El (hecho de) golpear Zayd a Amr"
- (30) $\text{ḍarb-u zayd-in } ^{\text{h}}\text{uwd-a}$
tañido-NOM Zayd-GEN_(AGT) eHaúd-ACC_(PAT)
 "El tañido de Zayd del láud (el hecho de tocar Zayd el láud)"

• **la elección de categoría sintáctica:**

(31) fa'ya?a-niy hadm-u 'l-?amiyr-i 'l-madiyna#-a"
sorpren*dió-me* destrucción-NOM el-emir-GEN la ciudad-ACC
"Me sorprendió que el emir destruyera la ciudad"

(32) hadma'ni 'l-madiyna#-i"
destrucción-DUAL-NOM la-ciudad-GEN
"Las dos destrucciones de la ciudad"

4.1 Desambiguación léxica

Para la *desambiguación léxica* necesitaremos información sobre los roles de los argumentos que aparecen ligados al *másdar*. Así, en el ejemplo (29), si el *másdar* "ḍarb" aparece con un rol paciente de tipo "humano"; lo traduciremos al castellano como "golpear" mientras que si el rol paciente tiene tipo semántico "instrumento musical" lo traduciremos por "tocar" o "tañer".

Esta información de subcategorización argumental la encontraremos en el léxico monolingüe árabe (LMA) y se utilizará durante el análisis sintáctico para asignar a los constituyentes relevantes las funciones sintácticas y los roles temáticos adecuados.

El formato actual de la estructura argumental de un verbo en el LMA es el siguiente:

```
frame1(arg1(np,subj,agt(hum)),
        arg2(pp("bi"),pobj1,pat(inanim)))
```

Los *másdars*, tanto si son *nominales de proceso* como si son *nominales de resultado*, no constituyen por sí mismos entradas léxicas del LMA, sino que se derivan de las entradas de los verbos correspondientes mediante reglas léxicas. Por lo tanto, su estructura de subcategorización es la del verbo correspondiente, con las diferencias indicadas en el apartado 3. Por otro lado, existen *másdars lexicalizados* que son tratados como nombres comunes y constituyen entrada léxica en el LMA.

4.2 Categoría sintáctica en la traducción.

Respecto a la *categoría sintáctica*, para decidir si debemos traducir el *másdar* como un nombre o como un verbo (ya sea en forma finita o en infinitivo) necesitaremos información que el léxico no nos proporciona y que tendremos que calcular durante o después del proceso de análisis sintáctico.

Por un lado, tendremos que decidir si el *másdar* en cuestión es de *proceso* o de *resultado*. Esta información la podemos deducir durante el análisis sintáctico, dado que algunas configuraciones sintácticas bloquearán la gramaticalidad de un *másdar de resultado* mientras que otras bloquearán su interpretación como *másdar de proceso*. Por otra parte, en una fase posterior al análisis pero previa a la transferencia léxica, deberemos poner la información necesaria en los nodos correspondientes para que en el acceso al Léxico Bilingüe se pueda utilizar esta información para seleccionar la categoría de la traducción.

4.3 Ejemplo

En la frase siguiente:

(33) sa'mi^htu ɖarab-a'ti 'f^huwd-i
 escuché-yo tañido-PLU-ACC el-laúd-GEN_(PAT)
 "Escuché los tañidos del laúd"

para poder traducir correctamente el másdar "ɖarb" como "tañido", y no como "golpe" o como formas finitas de los respectivos verbos ("tañer" y "golpear"), necesitamos seguir los siguientes pasos:

- Durante el análisis, se debe asignar correctamente la función de objeto directo y el rol de paciente (PAT) a "uwd" (laúd). Para hacer eso, hay que utilizar la información léxica de subcategorización argumental del verbo "daraba", del cual se deriva "ɖarb".
- Durante el análisis, se debe identificar si el *másdar* en cuestión funciona como un *másdar de proceso* o como un *másdar de resultado* y se debe poner dicha información en los nodos relevantes. En este caso, dado que está en plural, sólo puede ser un *másdar de resultado*.
- Durante la fase preparatoria antes de la transferencia léxica, se debe averiguar si el *másdar* en cuestión se va a traducir como un verbo o como un nombre. En (33), el hecho de que sea un *másdar de resultado* exige que se traduzca en castellano como un nombre.
- Finalmente, con toda esta información presente en los nodos adecuados del árbol, podemos efectuar la transferencia léxica y proseguir con la fase de generación.

Las entradas para "ɖarb" que tendremos en el Léxico Bilingüe serán (entre otras):

```
=====
ɖarb --> golpe
Test:
Rol Paciente (No Abstracto)
Categoría de salida (Nombre)
=====
```

```
ɖarb --> tañido
Test:
Rol Paciente (Instrumento Musical)
Categoría de salida (Nombre)
=====
```

```
ɖarb --> golpear
Test:
Rol Paciente (No Abstracto)
Categoría de salida (Verbo)
=====
```

```
ɖarb --> tañer
Test:
Rol Paciente (Instrumento Musical)
Categoría de salida (Verbo)
=====
```

5. Conclusiones

En este artículo hemos pretendido justificar que, para tratar computacionalmente durante la fase de análisis y de transferencia léxica de un sistema de T.A. un tipo de nombres deverbales árabes llamados *másdares*, es conveniente utilizar la información de subcategorización argumental de los verbos de los que aquéllos se derivan.

A pesar de que existen diferencias sustanciales en el comportamiento sintáctico y semántico de los *másdares* y de los verbos finitos correspondientes, la información sobre roles temáticos de los argumentos del *másdar* es vital para desambiguar su traducción y generar la categoría sintáctica correcta en castellano.

6. Bibliografía

1. Blachère, R. & Gaudefroy-Demonbynes, M. 1952. *Grammaire de l'Arabe Classique*.
2. Blake, B.J. 1994. *Case*, Cambridge University Press.
3. Corriente, F. 1984. *Gramática Árabe*. Instituto Hispano-Árabe de Cultura. Madrid.
4. Chalabi, A. 1992. *Automatic Parsing of Arabic Sentences*, en: Proceedings of the 3rd International Conference and Exhibition on Multilingual Computing (Arabic and Roman Script), University of Durham, 10-12 December, 1992.
5. Ditters, E. 1992. *A Formal Approach to Arabic Syntax*, Ph.D., Nijmegen.
6. Fassi Fehri, A. 1993. *Issues in the Structure of Arabic Clauses and Words*, Kluwer Academic Publishers.
7. Fleisch, H. 1979. *Traité de Philologie Arabe*, Vol. II, Dar El-Machreq eds. Beirut.
8. Grimshaw, J. 1992. *Argument Structure*, MIT.
9. Hindle, D. 1990. *Noun Classification from Predicate-Argument Structures*, en: Proceedings of the ACL 1990.
10. Levin, B. & Pinker, S. 1992. *Lexical and Conceptual Semantics*, Blackwell Publishers.
11. Owens, J. 1988. *The Foundations of Grammar. An Introduction to Medieval Arabic Grammatical Theory*, en: Amsterdam Studies in the Theory and History of Linguistic Science, E.F. Konrad Koerner (ed.), vol.45., John Benjamins Publishing Company, Amsterdam/Philadelphia.
12. Palmer, F.R. 1994. *Grammatical Roles and Relations*, Cambridge University Press.

ANEXO II

<u>Transliteracion</u>	<u>Arabe</u>	<u>Transliteracion</u>	<u>Arabe</u>
o	ا	g	غ
b	ب	f	ف
t	ت	q	ق
ʔ	ث	k	ك
ŷ	ج	l	ل
h	ح	m	م
j	خ	n	ن
d	د	h	هـ
q	ذ	w	و
r	ر	y	ي
z	ز	#	ة
s	س	?	أ
ʃ	ش	؛	إ
s.	ص	^	ء
q.	ض	û	ؤ
t.	ط	î	ئ
n.	ظ	â	ى
e	ع	am	آ