

PEDECIBA Informática
Instituto de Computación – Facultad de Ingeniería
Universidad de la República
Montevideo, Uruguay

Reporte Técnico RT 09-12

**Factividad de los eventos referidos en
textos**

Dina Wonsever Marisa Malcuori Aila Rosá

2009

Factividad de los eventos referidos en textos
Wosever, Dina; Malcuori, Marisa, Rosá, Aiala
ISSN 0797-6410
Reporte Técnico RT 09-12
PEDECIBA
Instituto de Computación – Facultad de Ingeniería
Universidad de la República

Montevideo, Uruguay, marzo de 2009

Factividad de los eventos referidos en textos

Dina Wonsever

Instituto de Computación
Facultad de Ingeniería
Universidad de la República
wonsever@fing.edu.uy

Marisa Malcuori

Instituto de Lingüística
Facultad de Humanidades y Ciencias de la Educación
Universidad de la República
mmalc@chasque.net

Aiala Rosá

Instituto de Computación
Facultad de Ingeniería
Universidad de la República
aialar@fing.edu.uy

Marzo de 2009

Resumen

Un aspecto central en la comprensión computacional de un texto es la detección de los eventos o estados referidos en el mismo. Pero no basta con encontrar términos que refieran a algún tipo de evento para inferir que dicho evento ocurrió o está ocurriendo. Es necesario además interpretar estos términos en sus contextos de ocurrencia, donde pueden verse afectados por elementos de polaridad negativa, o por operadores modales, o por predicados que afectan su valor de veracidad, y combinaciones de todos ellos. Llamamos *factividad* a la propiedad de eventos de haber o no ocurrido o de estar ocurriendo. En este artículo se analiza el atributo factividad asociado a eventos, se propone un espectro de valores para el mismo y se analizan diversos factores que intervienen en la determinación de estos valores.

Palabras clave

Procesamiento de lenguaje natural, Discurso, Eventos, Factividad

1. Introducción

Un aspecto central en la comprensión computacional de un texto es la detección de los eventos o estados referidos en el mismo.

*a. Esto **dificulta** aún más el **diálogo** con el gobierno uruguayo quien **confirmó** ayer a través de la cancillería que no se **negociará** mientras permanezca algún **corte**.*

En el ejemplo anterior hemos marcado con negrita los eventos referidos¹ y hemos subrayado un operador aspectual sobre un evento. Debemos notar que mientras algunos eventos se presentan como ocurridos (*confirmó, dificulta, corte*) otros son dudosos (*diálogo*) y finalmente la eventual negociación (*negociará*) se presenta como futura y con polaridad negativa. Esto significa que no basta con encontrar términos que refieran a algún tipo de evento para inferir que dicho evento ocurrió o está ocurriendo. Es necesario además interpretar estos términos en sus contextos de ocurrencia, donde pueden verse afectados por elementos de polaridad negativa, o por operadores modales, o por predicados que afectan su valor de veracidad, y combinaciones de todos ellos. La propiedad de eventos de haber o no ocurrido o de estar ocurriendo no es entonces un dato evidente. De hecho, es necesario realizar algún tipo de inferencia textual para poder determinarla.

La presentación como realizados o no de los eventos lleva la marca de quien está emitiendo el texto. En general los eventos se presentan desde la perspectiva del autor pero es muy habitual, especialmente en textos periodísticos, que otras voces aparezcan en los textos, habitualmente utilizando el recurso de discurso reproducido. En nuestro trabajo definimos como factividad la actitud del **enunciador** (habitualmente el autor) respecto a la ocurrencia del evento referido. En el ejemplo anterior el evento de declaración del gobierno uruguayo (*el gobierno uruguayo confirmó ...*) se presenta como ocurrido sin lugar a dudas, mientras que no hay manifestación de certeza sobre la ocurrencia o no del diálogo (*... dificulta aún más el diálogo ...*) y se presenta como futura pero con certeza la no negociación (condicionado además a otro evento). Lo que se dice sobre una futura negociación no es además enunciación del autor de la noticia, sino lo que este dice sobre lo que el gobierno dijo.

En el contexto de un esquema de anotación para el registro de cuáles son los eventos referidos en un texto [WON 08] se propone un atributo FACTIVIDAD para modelar la ocurrencia o no de los mismos. En este artículo se analiza el atributo factividad y los valores propuestos para el mismo. La estructura del documento es la siguiente: luego de una motivación e introducción (sección 1, actual) presentamos y analizamos antecedentes de trabajos en procesamiento de lenguaje orientados al análisis de la modalidad y la factividad en textos (sección 2). En la sección 3 presentamos e ilustramos con ejemplos nuestra propuesta para la factividad. En la sección 4 presentamos los elementos lingüísticos pertinentes para la determinación de la factividad y en la sección 5 esbozamos líneas de trabajo futuras en este tema.

2. Antecedentes

Los antecedentes que reseñamos corresponden a desarrollos recientes (en algún caso sus aplicaciones están aún en proceso) para el inglés. Están estrechamente relacionados con nuestra propuesta. Realizamos una presentación resumida y, en la sección 3, comparamos nuestro enfoque con estas propuestas.

¹ La expresión del evento suele contener más de una palabra. Se anota (y se muestra resaltado) un término considerado el núcleo del evento.

2.1. SlinkET: A Partial Modal Parser for Events (Saurí, Verhagen, Pustejovsky).

SlinkET es un *parser* basado en *TimeML* [TIM 08, SAU 08]. Ha sido desarrollado en el *framework Tarsqui* [TAR 08]. Se apoya en un reconocedor de eventos (Evita, [SAU 05]) y tiene como objetivo calcular, para cada evento relevante, un valor que expresa el grado de certeza acerca de la ocurrencia del evento. Esto se expresa discriminando entre:

- si ocurrió (fáctico o contrafáctico)
- si está siendo reportado o atestiguado por alguien (evidencial)
- si está siendo introducido como posible (modal)

En términos del esquema de anotación *TimeML*, el rol de *SlinkET* es generar links del tipo SLINK, con sus correspondientes valores del atributo CLASE. Notar que esto es distinto a calcular el grado de certeza para todos los eventos (en *SlinkET* sólo se trabaja para los eventos considerados relevantes, estos serían los asociados a los arcos SLINK).

SlinkET se apoya en la detección de determinados predicados que seleccionan situaciones como argumento (*SSP, Situation Selecting Predicates*) y en la información que dan algunas estructuras sintácticas (aún no implementado al momento de escritura del artículo) para realizar su trabajo. Los SSPs son predicados que seleccionan un argumento que denota un evento o situación de algún tipo. Algunos ejemplos son verbos como *claim, suggest, offer, avoid*, nombres como *promise, hope* y adjetivos como *ready, eager, able*.

Reportan 70% de medida-F1 (92% precision, 56% recall). Las reglas son manuales y están implementadas por cálculo finito.

2.2 Computing relative polarity for textual inference, Rowan Nairn, Cleo Condoravdi, Lauri Karttunen (2006)

En un contexto de un sistema de análisis, representación semántica (lógica) e inferencias se desarrolla un modulo para predicados implicativos (*lograr, forzar, etc.*) con el objetivo de determinar la factividad de sus argumentos. Para eso se considera, además de las propiedades de los predicados (se toman en cuenta sólo verbos), el contexto de polaridad en el que se encuentra referido el evento, definiendo el concepto de polaridad relativa globalmente determinada (*globally determined relative polarity*). Notar que el término implicación en este contexto es amplio e incluye tanto implicación lógica como pragmática.

Karttunen [KAR 70] define como verbos implicativos (lógicos) a aquellos que inciden en el valor de verdad de sus argumentos proposicionales. Clasifica estos verbos según el tipo de implicación que habilitan (verbos de tipo *sii*, verbos de *condición necesaria* o de *condición suficiente*) y nota que la implicación puede ser positiva o negativa dependiendo de la polaridad del ámbito en que ocurre el verbo. Como las construcciones implicativas pueden estar anidadas, la polaridad del predicado debe ser determinada relativamente a la cadena de predicados u operadores sentenciales en que se encuentra. Se propone entonces un algoritmo de cómputo de la polaridad para dar

cuenta de cómo afectan los predicados implicacionales anidados y los modificadores de polaridad el carácter de realizado o no de un evento que se encuentre en su alcance.

La propuesta implica también la construcción de un léxico de predicados implicativos para el inglés, ya que este tipo de información no se encuentra registrada en recursos léxicos existentes (p.ej., *WORDNET* [WOR 08]). Para esto se consideraron los verbos en orden de frecuencia decreciente en el *British National Corpus* (BNC), prestando especial atención a los verbos que toman complementos eventivos. Los juicios sobre el carácter implicacional se basaron en la concordancia entre varios anotadores, utilizando recursos tales como *Google Search* [GOO 09] y el *Linguist's Search Engine* [LIN 09] para armar las construcciones relevantes. De los 1250 verbos relevantes se encontraron más de 100 con algún tipo de carácter implicacional (lógico o presuposicional).

2.3. Determining Modality and Factuality for Text Entailment. Roser Saurí, James Pustejovsky

Este trabajo es una continuación y refinamiento del trabajo reseñado en 2.1 (también es un trabajo en curso), situándolo en un contexto más amplio de implicación textual (*Text Entailment*). La implicación textual (o inferencias textuales) es un nuevo tipo de tarea en PLN, que ha sido objeto de varias conferencias competitivas y de demostración [REF]. Los autores argumentan que una interesante inferencia textual es el establecimiento del status de factividad de los eventos referidos. En este contexto se propone una representación y un método de cálculo de la factividad de los eventos. Se caracteriza la factividad como una categoría que representa si las expresiones que denotan eventos son presentadas como correspondiendo a situaciones reales en el mundo, situaciones que no han ocurrido o situaciones de status inseguro. La factividad se ve afectada por la polaridad (positiva o negativa) y por la modalidad (epistémica). Para representar de modo compacto esta información se utiliza un par $\langle \text{Modalidad, Polaridad} \rangle^2$.

La información mencionada de factividad se asocia a cada una de las fuentes intervinientes. Una fuente es o bien el propio autor o una secuencia de fuentes mencionadas (autor dice que fuente 1 dice que fuente 2 dice ... : $\langle \text{autor, f1, f2} \rangle$).

Se presenta la especificación de un algoritmo para un analizador de factividad (*factuality profiler*) que implica:

- manejo de información no local a la cláusula en la que se expresa el evento
- manejo de la fuente del evento. Distintos participantes en el discurso pueden tener distintas perspectivas sobre el status de factividad del mismo evento. En el ejemplo siguiente se muestra la factividad de un evento (Juan está enfermo) desde la perspectiva del autor (H1), de Juan (H2) y de María (H3) :

² Polaridad toma los valores positivo (+), negativo(-) y desconocido(U).

Modalidad toma los valores cierto(CT), probable(PR), posible(PS) y desconocido(U)

a. *María lamenta que Juan no sepa que está enfermo.*

H1 – Juan está enfermo. Verdadero

H2 – Juan sabe que está enfermo. Falso

H3 – María sabe que Juan está enfermo. Verdadero

b. *María sabe que Juan está enfermo.*

c. *María no sabe que Juan está enfermo.*

d. *Yo no sabía que Juan está enfermo.*

e. **Yo no sé que Juan está enfermo.*

‘Juan está enfermo’ es verdadero para María en b y desconocido en c. Siempre es conocido y dado como un hecho para el autor, esto está vinculado a la semántica del verbo *saber*. Notar que mientras d. es posible (hay un cambio en el conocimiento del autor) e. sólo puede tener sentido si se estableció previamente un contexto hipotético.

Los predicados epistémicos (*saber, decir,...*) son los que pueden introducir nuevas fuentes en el texto. Cada predicado de este tipo tiene su propio esquema para asignar valores de factividad a su argumento. Intervienen la factividad contextual (modalidad + polaridad) y el hecho de si se trata del autor (o rol similar, *anchor* en el texto) o de otra fuente.

3. El atributo factividad.

El atributo factividad representa la certeza del enunciador respecto a la ocurrencia del evento referido. El enunciador es variable (generalmente, es el autor del texto), el eje enunciación de nuestro esquema de modelado del discurso da cuenta de la atribución de autoría a cada referencia de evento. Esto significa que cuando el autor del artículo reporta lo que dice otro enunciador no necesariamente coincide o se expresa en cuanto a la factividad de los eventos que el otro enunciador menciona. Un recurso muy utilizado en prensa escrita para “alejarse” de lo dicho es atribuírselo a otro enunciador no especificado.

Los valores del atributo factividad representan entonces la actitud del enunciador respecto a la ocurrencia de los eventos referidos. La factividad presenta una intersección importante con la modalidad epistémica, o sea, con la atribución de posibilidad o necesidad física a los eventos. No son, sin embargo, conceptos coincidentes. En la oración *Esta tormenta tiene que terminar* se presenta “certeza” (o necesidad) como modalidad epistémica. Sin embargo, el fin de la tormenta es un evento no realizado, a lo sumo posible.

Proponemos cinco valores posibles para el atributo factividad

- R – evento realizado
- RN – evento no realizado
- FP – futuro programado
- FN – futuro negado
- IND – indefinido

Los valores R y RN corresponden a certeza de la fuente respecto al evento referido. En el caso R, el evento se presenta como ocurrido u ocurriendo, en el caso RN, como no ocurrido. Todos los otros valores corresponden a eventos no realizados. Existe una graduación en lo que refiere a la posibilidad de ocurrencia según la fuente de eventos no realizados (en general, futuros) o cuya realización en el pasado es desconocida. En nuestro esquema de anotación hemos resuelto separar los valores de futuro (positivo y negativo) del terreno general de los indefinidos. Notar que en principio a : “María va a venir”, siendo futuro, y por lo tanto no factivo, le correspondería el valor INDEF (no es indudable ni que ocurra ni que no ocurra). A “María vendría” o “María podría venir” le correspondería también el valor INDEF, lo que no parece atinado. Por esa razón resolvimos distinguir los valores de “futuro asegurado o programado” (FP, FN).

Presentamos a continuación ejemplos que ilustran estos valores. Se marcan en negrita los núcleos de los eventos, en una versión simplificada respecto al esquema de anotación que hemos definido. Se indica a continuación el valor del atributo FACTIVIDAD para cada evento.

*a. Juan Peirano Basso **volvió a ingresar** a la sede judicial de Misiones, donde **declara** otra vez ante la jueza Graciela Gatti. La fiscalía **pedirá** hoy el **procesamiento** de Juan Peirano Basso, aunque la magistrada **tendrá** tiempo hasta mañana para **determinar** si **queda** en libertad.*

volvió a ingresar – R
declara – R
pedirá – FP
procesamiento – FP
tendrá – FP
determinar – IND
queda– IND

*b. Si Uruguay le **hubiera ganado** a Ecuador **estaría segundo** en la tabla de posiciones, ya que Brasil no **pudo** como local ante Bolivia y Argentina **empató** con Perú 1 a 1 en Lima. El próximo mes **jugamos** dos partidos como visitantes.*

hubiera ganado – RN
estaría segundo – RN
pudo – RN
empató – R
jugamos – FP

4. Elementos lingüísticos que inciden en la factividad

Existen distintos tipos de indicadores o elementos que influyen en el valor de factividad de un evento referido en el texto. En lo que sigue analizamos brevemente algunos de estos elementos.

4.1 Modo verbal

Los contextos de selección modal suelen estar determinados por algún inductor de naturaleza léxica (verbos, adjetivos, sustantivos) o sintáctica (las modalidades oracionales, como la negación o la interrogación) en función de que se exprese o no un estado de cosas percibido, afirmado o interpretado como cierto. En general, el modo verbal es sensible al hecho de que los estados de cosas se presenten como conocidos, imaginados, deseados, logrados, negados, etc. La selección modal puede ser obligatoria, como es el caso, por ejemplo, de los predicados volitivos o yusivos (*mandar, ordenar, pedir*, etc.) que seleccionan subjuntivo o de los predicados que expresan o presentan informaciones (*afirmar, señalar, indicar, asegurar*, etc.) que seleccionan indicativo (entre otros muchos casos). También puede darse alternancia modal con un mismo predicado (*sospecha que es mentira / sospecha que sea mentira*) que se corresponde con el grado de compromiso del hablante con respecto a la verdad del contenido proposicional: el subjuntivo mitiga la responsabilidad del hablante con respecto al contenido de la proposición.

Las nociones semánticas de aserción / no aserción, asociadas con el indicativo y subjuntivo respectivamente, han sido destacadas por numerosos autores para explicar la alternancia modal en español. En efecto, los predicados que inducen el subjuntivo presentan estados de cosas concebidos como una evaluación, una emoción, una intención o deseo, pero en lo que respecta a la factividad el comportamiento del subjuntivo no es uniforme. Hay construcciones con subjuntivo llamadas precisamente factivas que presuponen la ocurrencia del evento referido, como es el caso de los predicados valorativos o de afección (*lamentó que viniera, fue necesario que renunciara*) y el de los realizativos que inducen este modo verbal (*logró que salvaran el examen, consiguió que le rebajaran el precio*) que tomarán por lo tanto el valor R (evento realizado) y hay otros casos de predicados que expresan duda, incertidumbre, posibilidad, sospecha, mandato, que también inducen el subjuntivo, y no determinan la factividad del evento referido mediante este modo, pueden, por lo tanto, tomar valores IND (indefinido) (*duda de que lleguen temprano, temía que no terminaran a tiempo, es posible que no respondan, la orden de que se presentaran al otro día*).

4.2 Tiempo verbal

El tiempo verbal es una categoría gramatical déictica mediante la cual se expresa la localización de una situación o evento en términos de anterioridad, simultaneidad o posterioridad con respecto a la relación que mantiene con diversos puntos temporales de anclaje directa o indirectamente orientados con respecto al momento de la enunciación.

En términos muy generales, puede pensarse que las formas del indicativo que expresan anterioridad suelen señalar un evento como realizado, las que expresan posterioridad como no realizado o incierto, dependiendo del punto temporal de anclaje, y las que expresan simultaneidad, como realizado o incierto dependiendo

también del punto temporal de anclaje. Sin embargo, formas verbales como el presente y el futuro de indicativo pueden dar cuenta de eventos realizados en los usos llamados “históricos”; el futuro y el condicional pueden interpretarse como “presente de conjetura” y “pasado de conjetura” respectivamente en usos modalizados como *En este momento serán las tres de la tarde* y *Serían las tres de la tarde cuando golpearon a la puerta*; el imperfecto no tiene valor de pasado en usos de cortesía como *Quería pedirle un favor*, etc. Notar, asimismo, el uso del condicional modalizado (*María estaría embarazada cuando se casó.*), en este caso sabemos que el estado de embarazo tuvo lugar, sólo hay una imprecisión en cuanto a la fecha exacta. Corresponde entonces un valor R de factividad. Reparar en la diferencia con el uso de este tiempo en los llamados períodos condicionales (*Si ganara la lotería, Juan se compraría una casa.*) donde corresponde un valor de factividad de indefinido. Un tercer valor posible para el condicional es el de un futuro en el pasado, que se encuentra fundamentalmente en relatos históricos o novelas (*Abandonó el país en 1968, un año después se instalaría definitivamente en Madrid.*), al cual corresponde un valor de factividad R.

En cuanto al subjuntivo, el español establece en este modo menos distinciones temporales que en el indicativo y todas sus formas pueden expresar acciones futuras en alguna de sus interpretaciones y, por lo tanto, no realizadas o inciertas. Sin embargo, como ya se ha dicho al hablar del modo, las construcciones factivas presuponen la ocurrencia del evento referido y también se presenta como realizado un evento referido en imperfecto de subjuntivo en los llamados “usos indicativos” de este tiempo verbal que conllevan información que se presenta como conocida: *J:P., que fuera presidente de la institución durante los últimos cinco años, renunció ayer a su cargo.*

4.3 Determinación.

En el caso de eventos nominales, la determinación definida puede ser un índice de realización.

Señala Leonetti [LEO 99] que los nombres eventivos difícilmente aceptan la presencia del artículo indefinido. Esto está vinculado al hecho de que si aparecen los complementos argumentales en construcción referencial se refuerzan las condiciones para una interpretación unívoca de todo el sintagma.

Presenta los ejemplos:

*{La/*Una} captura de los infiltrados.*
*{La/*Una} publicación de este libro.*

Pero también son posibles construcciones con artículo indefinido cuando se selecciona de un conjunto o cuando se realiza una lectura resultativa

{La/Una} rápida obtención de beneficios fue el resultado de esta operación.
{La/Una} declaración del ministro en el día de ayer fue objeto de controversia.

Podemos concluir que la presencia de un artículo indefinido no excluye que nos encontremos ante un nombre eventivo. En todo caso, dadas ciertas condiciones sintácticas, la presencia del artículo indefinido puede ser un elemento más que permita descartar un sustantivo como nombrador de evento. Por otra parte, tanto los artículos indefinidos como los definidos con nombres eventivos no parecen influir en el status de factividad del evento.

4.4 Polaridad negativa

La presencia de elementos de polaridad negativa afecta la factividad. En principio, el valor de factividad R de un evento (*Juan fue a la reunión*) es transformado en el valor RN (*Juan **no** fue a la reunión*). Cuando estamos ante un valor de no certeza, la presencia de un elemento de polaridad negativa no afecta el valor del atributo factividad

- a. Juan (no) fue a la reunión. (no realizado)
- b. Juan no vendrá mañana. (no realizado)
- c. Juan no llega a las 10 (no realizado)

4.5 Presuposición e implicación– Existen predicados que implican ocurrencia o no ocurrencia de su argumento. Dos grandes tipos de relaciones lógicas se pueden dar: relación de factividad y de implicación (Kiparsky y Kiparsky 1970 [KIP 70], Karttunen 1970 [KAR 70]).

4.5.1- Presuposición

- a. Es **asombroso** que haga frío.
- b. Es **posible** que haga frío.

En a se presupone que hace frío, mientras que el predicado no factivo “es posible” no induce presuposición de este tipo. La misma diferencia se aprecia en :

- c. Juan está **seguro** de que María va a venir.
- d. Juan es **consciente** de que María va a venir.

Notar que “está seguro” implica un estado mental de Juan, no da indicaciones sobre la certeza del evento. Pero “es consciente” presupone la verdad de su argumento.

Una propiedad interesante de la factividad de un evento deducida por presuposición es que es impermeable a la negación del predicado que opera sobre este.

- d'. Juan **no** es **consciente** de que María va a venir.

La presuposición de que María va a venir se mantiene. O sea, el valor de factividad no cambia a pesar de la polaridad negativa. Lo mismo ocurre con operadores modales afectando al predicado matriz (*Juan sería consciente ...*, *Es posible que Juan sea consciente ...*, etc.). También con la interrogación : *¿Juan es consciente de que María va a venir?*

4.5.2 – Implicación

Existen predicados que establecen relaciones de implicación lógica con su argumento. La implicación puede ser positiva (*lograr*) o negativa (*evitar*). Al mismo tiempo, puede ser de tipo “si y solo si” (condición necesaria y suficiente), de tipo “si” (condición suficiente) y de tipo “solo si” (condición necesaria). Las relaciones de implicación, a diferencia de la presuposición, son sensibles a la polaridad del contexto en el que se encuentra el predicado matriz.

- a. Juan logró abrir la puerta.
- b. Juan evitó que se abriera la puerta.
- c. Juan lo forzó a abrir la puerta.
- d. Juan rehusó abrir la puerta.
- e. Juan intentó abrir la puerta.
- f. Juan dudó en abrir la puerta.

a. y b. son ejemplos de implicación “si y solo si”, positiva y negativa respectivamente. En efecto, en a. deducimos que se abrió la puerta así como en a'. que no se abrió la puerta. En b. deducimos que no se abrió la puerta y en b'. que se abrió.

- a'. Juan no logró abrir la puerta.
- b'. Juan no evitó que se abriera la puerta.

c. es un ejemplo de implicación por condición suficiente. Podemos deducir que la puerta se abrió pero no ocurre lo contrario en un contexto negativo (c'.). Del mismo modo en d'. deduciríamos que se abrió la puerta pero no podemos deducir nada en d.

- c'. Juan no lo forzó a abrir la puerta.
- d'. Juan no rehusó abrir la puerta.

Se presenta la situación inversa con los ejemplos e. y f. Mientras que en e. no podemos deducir nada respecto a la apertura de la puerta, en e'. deducimos que no se abrió la puerta. (Observar que podemos decir *Juan intentó abrir la puerta y lo logró* o *Juan intentó abrir la puerta y no lo logró*, por lo tanto no podemos deducir nada. En cambio en e' no tiene sentido *Juan no intentó abrir la puerta y logró abrirla*). En f', a su vez, deducimos que se abrió la puerta mientras que en f. no podemos deducir nada.

- e'. Juan no intentó abrir la puerta.
- f'. Juan no dudó en abrir la puerta.

4.6 Predicados introductores de hipótesis.

Dentro de los elementos léxicos que influyen en la factividad del argumento podemos definir una tercera clase considerando a los predicados que explícitamente señalan un valor de hipotético en su argumento. Forman parte de esta clase los verbos *suponer*, *hipotetizar*, ...

- a. El tren llegó a las 4.
- b. Supongamos que el tren llegó a las 4, ...

Es claramente distinta la referencia a la llegada del tren en a. y en b. Mientras que en a. se presenta como un evento realizado, en b. existe la voluntad expresa de no considerar la realización o no de la llegada del tren sino sus eventuales consecuencias.

4.7 Modalidad.

Nos referiremos brevemente a la modalidad de la oración y más en detalle a las expresiones asociadas a la posibilidad, necesidad o deseo de un evento (modalidades proposicionales).

4.7.1- Modalidad de la oración.

Nos interesamos fundamentalmente por la oración aseverativa o declarativa, que es la situación más frecuente en textos del tipo que tratamos. De todos modos, sí se presentan en este tipo de textos interrogativas indirectas, con modalidad interrogativa indirecta. También, especialmente en las citas directas, pueden presentarse interrogaciones directas, órdenes o exclamaciones.

Una primera observación simple es que en las oraciones interrogativas totales los eventos referidos no son afirmados como verdaderos o negados. Así en:

- a. ¿Ya salió el último tren?

No hay afirmación alguna respecto a si el tren salió o no salió. En nuestro esquema esto implicaría un valor de factividad INDEF.

Por otra parte, se mantienen otros mecanismos que influyen en la factividad de los eventos. En particular lo que no es directamente interrogado en las interrogativas parciales es presupuesto (esto sería una condición de “felicidad” en la enunciación de la pregunta).

- b. ¿Cuándo cerraron la puerta?

Así, en b. podemos inferir que la puerta fue cerrada. En general, en una interrogación se podrían presuponer todos los elementos sobre los que no se está preguntando.

4.7.2 Modalidades proposicionales.

Refieren a actitudes del hablante respecto a la posibilidad (de ocurrencia, de permiso) o necesidad de ocurrencia de un evento. Se habla de modalidades epistémicas (apreciación del mundo físico), deónticas (permiso, obligación). Se incluyen a veces en este tema a las modalidades volitivas (*querer, rechazar*). Existen diversos elementos que introducen modalidades en el texto: adjetivos (a), adverbios (b), verbos en perífrasis (c).

- a. Es posible que llueva.
- b. Posiblemente llueva.
- c. Puede llover.

Otro modo de introducción de este tipo de modalidad es a través del modo en la conjugación verbal, al cual ya hicimos referencia en el punto 4.1. Así en d. y en e. expresamos distintas actitudes respecto a la ocurrencia de la llegada del tren.

- d. El tren llega / llegó a las 4.
- e. Ojalá que el tren llegue a las 4.

Las manifestaciones de modalidad que afectan la expresión de un evento tienen claras consecuencias sobre el valor del atributo factividad del evento. En ninguno de los ejemplos a-c y e podemos concluir que lloverá o que el tren llegó.

4.8 Contextos sintácticos

El hecho de encontrarse en el alcance de algunas construcciones sintácticas influye en el valor del atributo factividad asociado a un evento. Algunas construcciones implican presuposición de verdadero (relativas, b.), adverbiales de causa o explicación (c.) o de falso (condicionales contrafácticas, d.), otras señalan el carácter hipotético del evento (condicionales con si, a.). La factividad de los eventos referidos se ve influenciada así por las construcciones retóricas del texto.

- a. Si el tren llegó a las 4, no hay razón para seguir esperando.
- b. María vino en el tren que llegó a las 4.
- c. Como el tren llegó a las 4, María debe estar ya en su casa.
- d. Si el tren hubiera llegado a las 4, María ya estaría acá.

En el sistema SlinkET se reportan los siguientes contextos sintácticos:

- cláusulas relativas : el evento denotado por la cláusula relativa se presupone verdadero
- cláusulas *cleft* (*fue María la que llegó tarde*) : el evento se presupone verdadero
- subordinadas temporales : el evento en la cláusula temporal se presupone verdadero
- cláusulas de propósito : el evento de la cláusula es intensional (posible en algún futuro)
- construcciones condicionales : el evento denotado por el consecuente es intensional y dependiente de la factividad del antecedente, también intensional.

5. Determinación de la factividad en textos

La determinación de la factividad requiere previamente haber realizado la detección de los eventos referidos en el texto. En tal sentido, y dentro de un esquema general de trabajo, la propuesta es realizar la anotación de un volumen significativo de textos y luego la inferencia, por aprendizaje automático o por construcción directa, de un conjunto de reglas o de un método de cálculo del valor de factividad. Cualquiera sea la metodología, es importante el análisis previo de qué situaciones se pueden presentar y cuáles son los atributos relevantes.

Del estudio de los atributos que influyen en la factividad se deduce que es importante conocer el alcance de expresiones que afectan la factividad: negaciones, operadores modales, predicados subordinantes. Los sistemas que hemos mencionado en la sección 2 fueron diseñados para el inglés; para esta lengua se dispone de analizadores sintácticos que pueden dar la información de alcance de los distintos operadores. La situación es distinta para el español; sólo se dispone de analizadores de superficie, tales como FreeLing [FRE 09] que construyen estructuras sintácticas parciales. Se planea utilizar un segmentador en cláusulas de desarrollo propio [CLA 04, WON 01] ya que si bien no realiza análisis sintáctico completo, marca adecuadamente las subordinaciones y los alcances de los operadores que nos interesan.

Bibliografía

[BOS 90] Bosque, Ignacio. "Las bases gramaticales de la alternancia modal. Repaso y balance". En I. Bosque (ed.). *Indicativo y Subjuntivo*. Madrid: Taurus, 1990, p. 13-65.

[BOS 99] Bosque, Ignacio. "El nombre común". En: I. Bosque y V. Demonte (dirs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. Vol. 1, p. 3-75.

[CLA 04] Proyecto *Clatex*, *Un segmentador en cláusulas para el español*. InCo, Grupo de Lenguaje Natural, 2004

[CRO 03] Crouch Dick, Condoravdi Cleo, De Paiva Valeria, Stolle Reinhard, Bobrow Daniel G. *Entailment, Intensionality and Text Understanding* In Proceedings of the Human Language Technology Conference, Workshop on Text Meaning (2003)

[DEM 99] De Miguel, Elena. "El aspecto léxico". En: I. Bosque y V. Demonte (dirs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. Vol. 2, p. 2977-3060.

[FRE 09] Freeling 2.1 An Open Source Suite of Language Analyzers, <http://garraf.epsevg.upc.es/freeling/> (consulta febrero 2009)

[GOO 09] Google search, <http://www.google.com/coop/cse/>, (consulta febrero 2009)

[GRA 01] Grassi, M., M. Malcuori, J.Couto, J.J. Prada, D.Wonsever, "Corpus informatizado: textos del español del Uruguay (CORIN)", *SLPLT-2 - Second International Workshop on Spanish Language Processing and Language Technologies - Jaén*. España. 2001.

[KAR 70] Karttunen. L.1971. Some observations on factivity. *Papers in Linguistics*, pages 55–69

- [KIP 70] Kiparsky P., Kiparsky C. 1970. Fact. Manfred Bierwisch and Karl Heidolph Eds, *Progress in Linguistics. A Collection of Papers*, pages 143-173. Mouton, The Hague, Paris.
- [LEO 99] Leonetti, Manuel. "El artículo". En: I. Bosque y V. Demonte (dirs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. Vol. 1, p. 787-890.
- [LIN 09] Linguistic Search Engine, <http://lse.umiacs.umd.edu:8080/>, (consulta febrero 2009)
- [ROJ 99] Rojo, Guillermo. "El tiempo verbal. Los tiempos simples". En: I. Bosque y V. Demonte (dirs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. Vol. 2, p. 2867-2934.
- [PIC 99] Picallo, M. Carme. "La estructura del sintagma nominal: las nominalizaciones y otros sustantivos con complementos argumentales". En: I. Bosque y V. Demonte (dirs.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. Vol. 1, p. 363-393.
- [PUS 03] Pustejovsky, James, Patrick Hanks, Roser Saurí, Andrew See, Robert Gaizauskas, Andrea Setzer, Dragomir Radev, Beth Sundheim, David Day, Lisa Ferro and Marcia Lazo. "The TIMEBANK Corpus". En *Proceedings of Corpus Linguistics*, 2003. P. 647-656.
- [ROS 08] Rosá, Aiala. "Identification Automatique de marques d'opinion dans des textes", *Actas de JEP-TALN-RECITAL '08*. Avignon, 2008.
- [SAU 05] Saurí, Roser, Robert Knippen, Marc Verhagen y James Pustejovsky. "Evita: A Robust Event Recognizer for QA Systems". En *Proceedings of HLT/EMNLP*, 2005. P. 700-707.
- [SAU 08] Saurí, Roser, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer y James Pustejovsky. *TimeML Annotation Guidelines Version 1.2.1* [en línea]. Enero 2006 [consulta: setiembre 2008].
- [TAR 08] Proyecto Tarsqi, <http://www.timeml.org/site/tarsqi/index.html>, consulta setiembre 2008
- [TIM 08] TimeML Working Group. *TimeML 1.2.1. A Formal Specification Language for Events and Temporal Expressions* [en línea]. Octubre 2005. [consulta: setiembre 2008].
- [WON 08] Wonsever D., Malcuori M., Rosá A., *SIBILA: Esquema de anotación de eventos*, Reporte Técnico RT 08-11 – InCo UDELAR, Montevideo, 2008
- [WON 01] Wonsever D., Minel J-L. "Contextual Rules for Text Analysis", *Lecture Notes for Computer Science* 2004, pages 509-523, 2001

[WOR 08] Wordnet, a lexical database for the English language
<http://wordnet.princeton.edu/> (consulta febrero 2009)