

Anotación Discursiva de Corpus en Español

Apéndice I: procedimientos de decisión para anotación

enero 2004

Resumen

Proporcionamos un procedimiento de decisión pautado para determinar si cada una de las características posibles para la descripción de las relaciones entre segmentos discursivos se aplica en cada caso.

1 Introducción

Este apéndice está motivado por la necesidad de mayor concreción en los criterios de anotación discursiva de corpus, dados los bajos valores de acuerdo κ entre jueces (nunca más altos de .65) en los trabajos preliminares de anotación.

Se basa en las observaciones realizadas durante el trabajo preliminar de anotación de corpus llevado a cabo en el último trimestre de 2003 por Ezequiel Andújar Hinojosa y Roberto Sola Salvatierra. Daniel Alonso i Alemany contribuyó a la delimitación del nivel de anotación, y al diseño final de la visualización en XML. También ha hecho aportaciones substanciales Henk Zeevat, sobretodo respecto a la característica de *revisión*.

2 Árboles de decisión

Se aplica una metodología de árboles de decisión que ha sido utilizada con frecuencia en anotación discursiva de corpus, para consolidar la consistencia y reproducibilidad de la anotación. Los árboles de decisión ofrecen un procedimiento a seguir para cada decisión a tomar, como se ve en el ejemplo 1.

En nuestro caso, se trata de árboles que permiten pasar del caso por defecto a casos más específicos. Es decir, en cada decisión que nos encontremos existen varias posibilidades, una de las cuales es el caso por defecto, y lo aplicaremos si no se cumple ninguna de las condiciones para aplicar los demás casos.

3 Árboles de decisión para características discursivas

Los casos por defecto se aplicarán siempre que el ejemplo no contradiga la definición de la característica.

3.1 Simétrico, Asimétrico

El caso por defecto es *simétrico*.

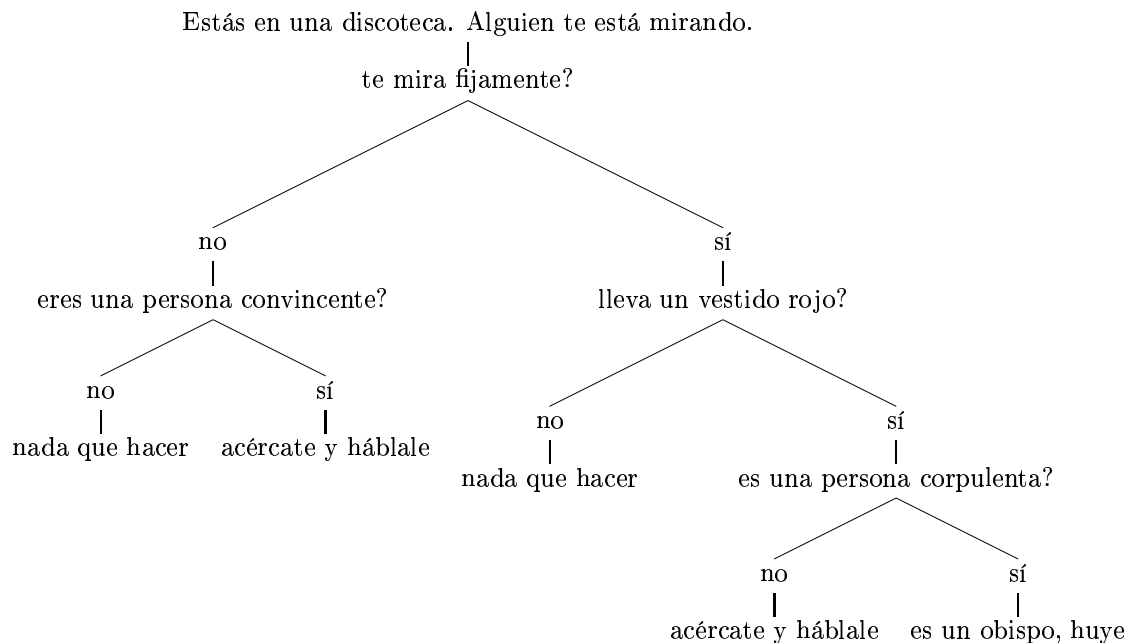


Figura 1: Ejemplo de árbol de decisión

simétrico una unidad discursiva se encuentra en el mismo nivel estructural que la unidad a la que se adjunta.

Millones de personas están consumiendo estas drogas en Estados Unidos y en Europa y no sabemos realmente nada sobre los efectos que producen.

asimétrico una unidad discursiva se encuentra en un nivel estructural distinto que la unidad a la que se adjunta.

Hay que leer porque la literatura es un gran plagio.

3.2 Progresión, Elaboración

El caso por defecto es *progresión*.

Atención!!

Para cada segmento que introduce una progresión, hay que resumir en pocas palabras qué contenido u objetivo introduce.

progresión una unidad discursiva aporta un elemento nuevo a la línea argumentativa del texto.

Para freír un huevo hay que seguir los siguientes pasos: primero, calentar el aceite en una sartén. Luego, romper el huevo encima de la sartén.

escribid → “romper el huevo”

La relación entre el segmento B y el segmento A es...

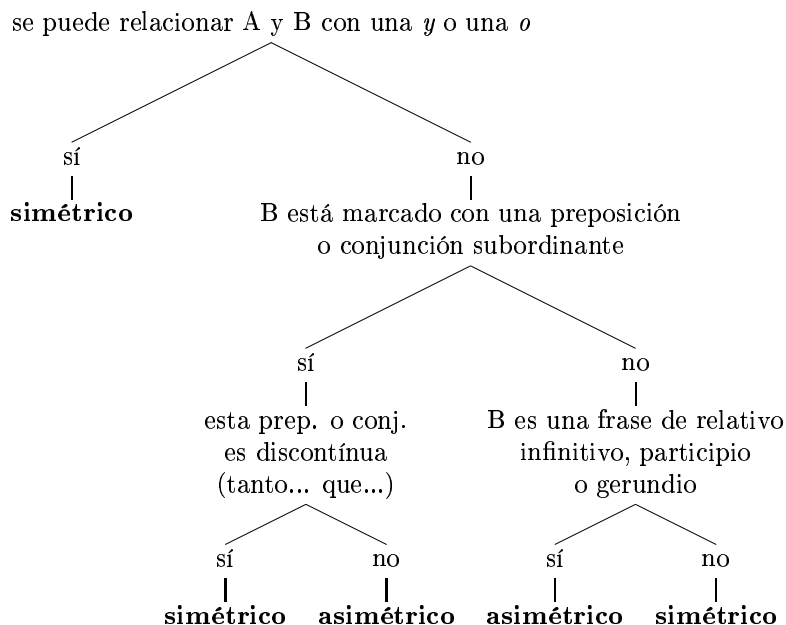


Figura 2: Árbol de decisión para las características simétrico y asimétrico, donde *simétrico* es el caso por defecto.

elaboración una unidad discursiva aporta información adicional sobre un elemento ya introducido en la línea argumentativa del texto.

amplía información

Luego, romper el huevo encima de la sartén. Esto puede ser muy peligroso.
escribid → *nada o bien "romper el huevo"*

desarrolla un objetivo

Luego, romper el huevo encima de la sartén. Para ello, deberéis golpear el huevo contra alguna superficie dura con contundencia pero delicadamente.
escribid → *nada o bien "romper el huevo"*

3.3 Revisión, Causalidad, Paralelismo, Contexto

El caso por defecto es *ninguno*.

La relación entre el segmento B y el segmento A es...

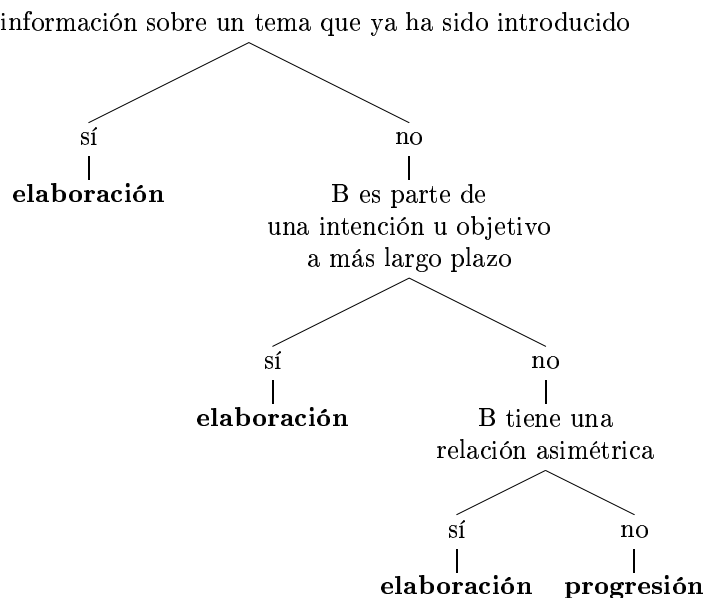


Figura 3: Árbol de decisión para las características progresión y elaboración, donde *progresión* es el caso por defecto.

Atención!!

Para cada segmento que introduce una relación de este tipo, hay que escribir:

revisión la prueba que nos permite saber de qué tipo de revisión se trata

causalidad la conjunción de causa con que podemos relacionar los segmentos

Conjunciones de causa: porque, para que, por eso, para eso

paralelismo la clase a la que pertenecen los dos elementos relacionados

contexto el pronombre circunstancial con que podemos preguntar por el segmento

revisión una unidad discursiva revisa una información que se desprende de la unidad a que se adjunta, ya sea porque la información está contenida en ella o porque está implicada por ella.

concesiva

Hoy llueve pero saldré a dar una vuelta.

escribid → “cuando llueve me quedo en casa”

Pepe ha pedido un visado para los EEUU pero no es ningún terrorista.

escribid → “si pides un visado para los EEUU, eres un terrorista”

adversativa

Es muy trabajador, pero le cuesta entender las cosas.

escribid → “queremos hacer el trabajo con él”

contrastiva

Pepe es simpático, pero Lolo es un imbécil.

escribid → “simpático - imbécil”

No es blanco, sino negro.

escribid → “blanco - negro”

causa una unidad discursiva establece una relación de causa con la unidad a la que se adjunta.

Me gustan las manzanas, me voy a comer una.

(*Me gustan las manzanas, y por eso me voy a comer una.*)

escribid → “por eso”

paralelismo una unidad discursiva establece una relación de equivalencia con la unidad a que se adjunta, del tipo *A es (modalidad) B*, por ejemplo: *A es un tipo de B*, *A es un ejemplo de B*, *A es como B*, *A es lo contrario de B*.

Nos compraremos un piso. Pediremos una hipoteca, viviremos con el agua al cuello toda la vida y cuando no nos podamos ya mover, el piso será nuestro.

escribid → “lista”

No me gustan los hombres. Tú, menos.

escribid → “hombres”

contexto una unidad discursiva sitúa la unidad a la que se adjunta en un contexto determinado.

Si Colón no se hubiera pasado de listo, ahora habría más indios en América.

escribid → “si”

No se dirigieron la palabra durante todo el viaje.

escribid → “cuando”

La relación entre el segmento B y el segmento A es...

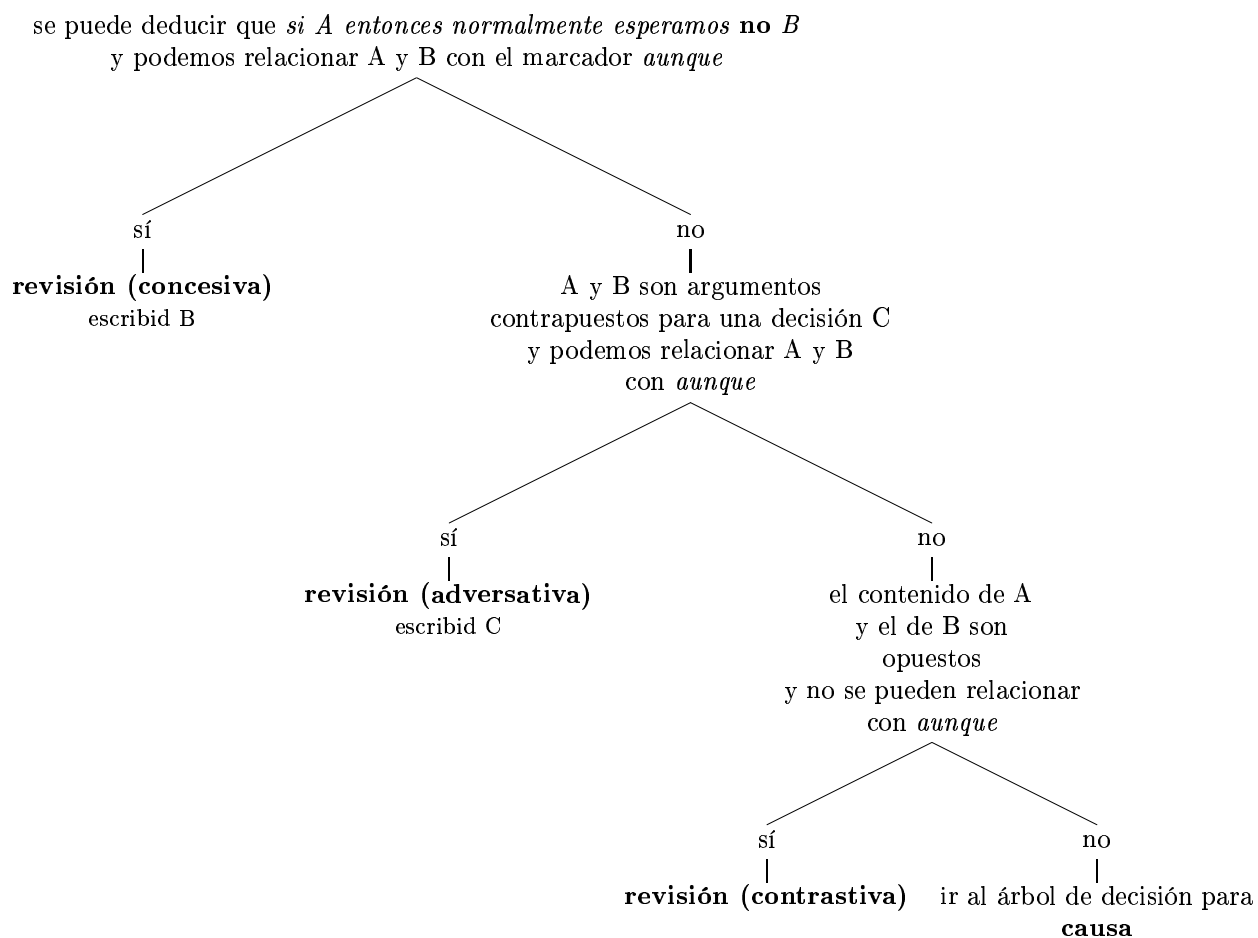


Figura 4: Árbol de decisión para la característica de revisión, donde no hay caso por defecto.

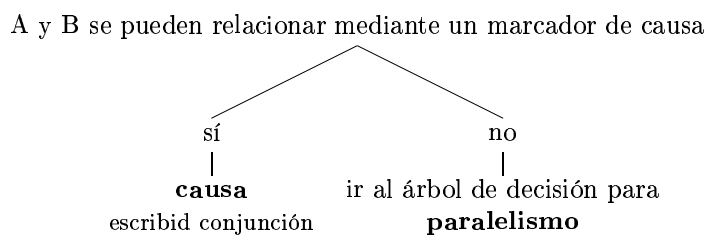


Figura 5: Árbol de decisión para la característica de causa, donde no hay caso por defecto.

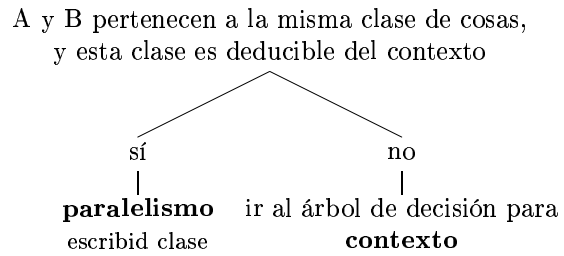


Figura 6: Árbol de decisión para la característica de paralelismo, donde no hay caso por defecto.

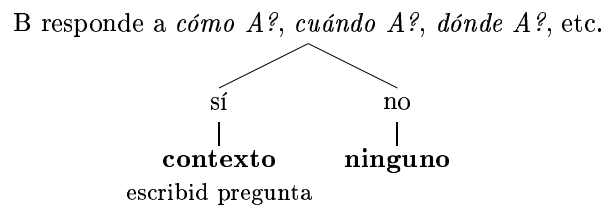


Figura 7: Árbol de decisión para la característica de contexto, donde no hay caso por defecto.