

# MARCADORES Y RELACIONES DISCURSIVAS EN EL ÁMBITO MÉDICO: UN ESTUDIO EN ESPAÑOL Y EUSKERA

MIKEL IRUSKIETA

*Universidad del País Vasco (UPV-EHU)*

IRIA DA CUNHA

*Institut Universitari de Lingüística Aplicada (UPF)*

## RESUMEN

*El objetivo de este estudio es analizar las relaciones discursivas de la Rhetorical Structure Theory (RST) de Mann y Thompson (1988) empleadas en el discurso médico en español y euskera, y detectar los marcadores del discurso que las evidencian. Este análisis nos permite observar diferencias de comportamiento entre estas dos lenguas, tan diferentes entre sí, en el ámbito médico. La metodología del estudio consiste en contabilizar diversos aspectos discursivos sobre un corpus médico etiquetado con relaciones de la RST, para después llevar a cabo un análisis cualitativo de los mismos.*

Palabras clave: marcadores discursivos, relaciones discursivas, RST

## ABSTRACT

*The aim of this study is to analyze the discourse relations of the Rhetorical Structure Theory (RST) by Mann and Thompson (1988) that are used on the Spanish and Basque medical discourse, and to detect the discourse markers showing them. This analysis allows us to observe performance differences between these two so different languages in the medical domain. The methodology of the study consists of, on the one hand, the accounting of several discourse aspects of a medical corpus that has been annotated with RST relations and, on the other hand, the quantitative analysis of them.*

Keywords: discourse markers, discourse relations, RST

BUENO ALONSO, Jorge L., Dolores GONZÁLEZ ÁLVAREZ, Úrsula KIRSTEN TORRADO, Ana E. MARTÍNEZ INSUA, Javier PÉREZ GUERRA, Esperanza RAMA MARTÍNEZ & Rosalía RODRÍGUEZ VÁZQUEZ (Eds). 2010: *Analizar datos > Describir variación / Analysing data > Describing variation*. Vigo: Universidade de Vigo (Servizo de Publicacións); 146-159

## 1. INTRODUCCIÓN

El estudio de la estructura discursiva de los textos y de sus correspondientes marcadores discursivos es un tema de investigación que ha suscitado gran interés por parte de la comunidad científica internacional. Para el estudio de la estructura discursiva puede partirse de diversos marcos teóricos (véanse, entre otros, los trabajos de Austin 1962, 1970; Benveniste 1966; Labov y Waletzky 1967; Grice 1975; Halliday y Hasan 1976; van Dijk 1980, 1989; Anscombe y Ducrot 1983; Sperber y Wilson 1986). En este trabajo partimos de la *Rhetorical Structure Theory* (RST) de Mann y Thompson (1988).

El objetivo de nuestro trabajo es analizar las relaciones discursivas que se emplean en el discurso médico en español y en euskera (empleando para ello las relaciones de la RST) y observar los marcadores del discurso que las evidencian (a partir de la clasificación ofrecida por Portolés [2001]). Este análisis nos permite observar diferencias de comportamiento entre estas dos lenguas, tan diferentes entre sí, en el ámbito médico. Existen algunos trabajos sobre comparación de estructuras discursivas en diferentes lenguas (Marcu et al. 2000, da Cunha e Iruskietia 2009). Sin embargo, no existen trabajos previos que comparen las estructuras y marcadores del discurso de estas dos lenguas, ni siquiera a partir de corpus de ámbito general. Nuestro corpus de trabajo está formado por resúmenes de artículos médicos de investigación en español y en euskera.

En la sección 2 detallamos la metodología empleada en nuestro trabajo. En las secciones 3 y 4 mostramos el análisis cuantitativo y el análisis cualitativo realizados, respectivamente. En la sección 5 exponemos las conclusiones.

## 2. METODOLOGÍA

La metodología de este estudio incluye diversas fases. En primer lugar, seleccionamos el corpus de trabajo. Este corpus incluye 20 resúmenes de artículos médicos de investigación en español y en euskera extraídos de la Gaceta Médica de Bilbao<sup>1</sup> escritos entre los años 2000 y 2008 por diferentes especialistas. Esta revista solicita a los autores los resúmenes de sus artículos en español, euskera e inglés.

El hecho de que la revista exija a los autores incluir determinadas informaciones (propósito del estudio, procedimientos utilizados, hallazgos principales y conclusiones más relevantes) en sus resúmenes en las tres lenguas es una garantía de que dichos resúmenes incluyen la misma información y de que siguen una estructura similar. Este corpus paralelo fue previamente etiquetado con relaciones de la RST en el trabajo de da Cunha e Iruskieta (2009).

La RST es una teoría descriptiva de organización del texto, independiente de lengua, muy útil para describirlo caracterizando su estructura a partir de las relaciones que mantienen entre sí los elementos discursivos o retóricos<sup>2</sup> del mismo (Circunstancia, Elaboración, Motivación, Evidencia, Justificación, Causa, Propósito, Antítesis, Condición, entre otras). Por lo general, una de las unidades de la relación es el gobernante (*núcleo*), mientras que la otra (*satélite*) aporta cierta información retórica acerca de él. Estas relaciones se denominan relaciones *nucleares* o *núcleo-satélite*. En el caso de las relaciones que no presentan una unidad central con respecto a los propósitos del autor, la relación se denomina *multinuclear*.

En segundo lugar, contabilizamos diversos aspectos discursivos de este corpus. Entre otros aspectos analizamos cuantitativamente en cada subcorpus el porcentaje de relaciones y de marcadores, y las correlaciones entre relaciones y marcadores.

En tercer lugar, llevamos a cabo un análisis cualitativo del corpus. Analizamos en ambas lenguas el grado de ambigüedad de los marcadores y determinamos qué tipos de marcadores (partiendo de la ya mencionada clasificación de Portolés) reflejan cada relación.

La dispersión terminológica sobre las unidades lingüísticas que marcan relaciones entre unidades discursivas ha sido señalada en ambas lenguas (para el español, Martín Zorraquino y Portolés 1999; para el euskera, Iruskieta et al. 2009 y Esnal 2002)<sup>3</sup>. Además de la falta de acuerdo sobre la denominación, existe también una falta de acuerdo sobre los elementos que componen los marcadores, muchas veces determinada por el acercamiento teórico. Por ejemplo, para Schiffrin (1987), los marcadores del discurso son palabras como “well”, “now” u “oh”, que únicamente adquieren sentido discursivo en conversaciones y no en textos especializados. Consideramos que el

acercamiento de Portolés (2001) es más adecuado para el estudio de los textos médicos que componen nuestro corpus.

### 3. ANÁLISIS CUANTITATIVO

Hemos realizado un análisis cuantitativo de diversos factores presentes en el corpus bilingüe. En la Tabla 1 se muestran los resultados del análisis. En la primera columna (RELACIÓN) se indica el nombre de la relación y el tipo (NN = relación multinuclear, P = relación núcleo-satélite de presentación, C = relación núcleo-satélite de contenido). En la segunda columna (Nº EU) se muestra su frecuencia de aparición en euskera. En la tercera columna (MD EU), se ofrecen los marcadores que señalan cada relación en euskera. En la cuarta columna (Nº MD) se indica la frecuencia de aparición de dichos marcadores. En las tres columnas finales (Nº ES, MD ES, Nº MD) se ofrece la misma información para el español.

RELACIÓN	Nº EU	MD EU	Nº MD	Nº ES	MD ES	Nº MD
<b>Antítesis P</b>	2	baina (1)	1	2	pero (1)	1
<b>Causa C</b>	1	-	0	4	debido a (1) por lo que (2)	3
<b>Circunstancia C</b>	1	-	0	2	a la hora de (1)	1
<b>Concesión P</b>	3	-n arren (1) oraindik ere (1) baina (1)	3	1	aunque (1)	1
<b>Condición C</b>	2	-	0	2	-	0
<b>Contraste N-N</b>	11	aldiz (2) eta berriz (1) berriz (1) bestalde (2)	6	12	mientras que (2) y (2) o (1) por lo demás (1) frente a (1) por el contrario (1)	8
<b>Disyunción N-N</b>	1	edo (1)	1	0	-	0
<b>Elaboración C</b>	81	eta (5) era berean (2) baita ere (1)	17	56	como (2) en consecuencia (1) de ellos (1)	4

		hala Nola (3) horien artean (1) horren ondorioz (1) horien artetik (1) beraz (1) gainera (2)				
<b>Evidencia P</b>	2	-	0	1	como (1)	1
<b>Fondo P</b>	11	-	0	10	-	0
<b>Interpretación C</b>	14	beraz (1) ondorioz (1) eta (1)	3	13	lo que (1)	1
<b>Justificación P</b>	4	izan ere (1) horrela (1) horrela gauzak (1) -lako (1)	4		-	0
<b>Lista N-N</b>	40	eta (9) era berean (6) horrez gain (1)	16	59	y (16) así mismo (1) así como (3) aparte de (1) igualmente (2)	23
<b>Medio C</b>	11	besteak beste (1)	1	19	para ello (1) mediante (1)	2
<b>Motivación</b>	0	-	0	1	por lo que (1)	1
<b>Preparación P</b>	23	-	0	21	-	0
<b>Propósito C</b>	4	hau dela eta (1)	1	2	para (1) con el fin de (1)	2
<b>Resultado C</b>	14	eta (1) horien artean (1) horrenbestez (1)	3	17	-	0
<b>Secuencia</b>	0	-	0	2	-	0
<b>Total de RR</b>	<b>225</b>		<b>56</b>	<b>224</b>		<b>48</b>
<b>Total de N-N</b>	52		23	73		32
<b>Total de N-S</b>	173		33	151		16
<b>Total de presentación</b>	45		8	36		4
<b>Total de contenido</b>	128		25	115		13

Tabla 1. Datos de la evaluación cuantitativa

Como puede observarse, en el corpus en español se encuentran 224 relaciones de la RST y 48 marcadores que las evidencian, mientras que en euskera se detectan 225 relaciones y 56 marcadores. Esto quiere decir que en nuestro corpus en español hay un 21,42 % de relaciones discursivas evidenciadas por marcadores, mientras que en el corpus en euskera se marcan un 24,88 %.

En el corpus en español, hay 73 relaciones multinucleares (lo cual supone un 32,58 % de las relaciones totales) y 151 relaciones núcleo-satélite (es decir, un 67,41 % de las relaciones). A su vez, estas se dividen en 36 relaciones de presentación (16,07 %) y 115 de contenido (51,33 %). En cambio, en el corpus en euskera, se detectan 52 relaciones multinucleares (23,11 %) y 173 relaciones núcleo-satélite (76,88 %). De estas, 45 son relaciones de presentación (20 %) y 128 son relaciones de contenido (56,88 %).

El porcentaje de relaciones discursivas de cada tipo que se encuentran en los textos en español y en euskera son muy similares. La diferencia más representativa es la existencia de un porcentaje más elevado de relaciones de Elaboración en el corpus en euskera (36 %) que en el corpus en español (25 %). Quizás esta situación se deba al hecho de que el autor haya escrito en primer lugar el resumen en español y después lo haya traducido al euskera. En cambio, en los textos en español encontramos un número mayor de relaciones de Lista (26,33 %) que en los textos en euskera (17,77 %). Probablemente este hecho se deba a que en español el marcador “y” suele marcar una relación de Lista, no de Elaboración, mientras que en euskera, “eta” (que significa “y”) puede marcar indistintamente una relación de Lista o de Elaboración.

Asimismo, resulta curioso observar que en español no se ha marcado ninguna relación de Justificación, mientras que en euskera esta relación supone el 1,77 %. Probablemente se deba al empleo de los marcadores: “izan ere”, “horrela”, “horrela gauzak” y “-lako”, que marcan dicha relación. En español, sin embargo, en las tres primeras ocasiones no hay marcador explícito para expresar “de hecho”, “de ese modo” y “así las cosas”. En cambio, en la última, en español se encuentra el marcador “debido a”, que marca una relación de Causa.

#### 4. ANÁLISIS CUALITATIVO

Una vez realizado el análisis cuantitativo del corpus, nos centramos en el análisis de algunos aspectos más concretos, como el tipo de marcadores (partiendo de la clasificación de Portolés) que reflejan cada relación y el grado de ambigüedad de algunos marcadores.

Con respecto a los tipos de marcadores encontrados en ambos corpus, encontramos algunas diferencias. Por un lado, en euskera se observa un 14,28 % de estructuradores de la información, un 35,71 % de conectores aditivos, un 16,07 % de conectores consecutivos, un 17,85 % de conectores contrargumentativos, un 3,57 % de reformuladores y un 12,5 de operadores de concreción. Por otro lado, en español se observa un 8,33 % de operadores de concreción, un 52,08 % de conectores aditivos, un 12,5 % de conectores contraargumentativos, un 20,83 % de conectores consecutivos y un 6,25 % de ordenadores. De estos datos se concluye que en el discurso médico en euskera se emplean muchos más estructuradores y reformuladores que en el discurso médico en español, que no emplea ninguno de ellos.

Asimismo, y en relación con el hecho de que en español se encuentran muchas más relaciones de Lista que en euskera, se observa que en español se emplea un número más elevado de conectores aditivos que en euskera.

En la Tabla 2 se incluye la clasificación de los marcadores en euskera y en español encontrados en nuestro corpus bilingüe. En la primera columna (Clasificación Portolés general) se indica la tipología general de los marcadores. En la segunda columna (Clasificación Portolés específica) se muestra la tipología específica. En la tercera y cuarta columnas (MD ES y MD EU) se exponen, respectivamente, los correspondientes marcadores en español y en euskera. En la quinta y sexta columnas (Nº ES y Nº EU) se especifica la frecuencia de aparición de dichos marcadores en las dos lenguas. Los marcadores entre corchetes son traducciones realizadas por nosotros que no han aparecido en el corpus.

Clasificación Portolés general	Clasificación Portolés específica	MD ES	MD EU	Nº ES	Nº EU
Estructuradores de la información	Ordenador	así mismo	era berean	1	8
	Ordenador	por lo demás	[bestela]	1	0
	Ordenador	a la hora de	[-t(z)eko orduan]-	1	0
Conectores	Conector aditivo	y	eta	18	16
	Conector aditivo	[además]	gainera	0	2
	Conector aditivo	así como	[baita ere / halaber]	3	0
	Conector aditivo	aparte de	horrez gain	1	1
	Conector aditivo	o	edo	1	1
	Conector aditivo	igualmente	[halaber]	2	0
	Conector consecutivo	[debido a]	hau dela eta	0	1
	Conector consecutivo	con el fin de	[-t(z)eko helburuarekin]	1	0
	Conector consecutivo	[de hecho]	izan ere	0	1
	Conector consecutivo	en consecuencia	beraz	1	2
	Conector consecutivo	[por lo tanto]	ondorioz	0	1
	Conector consecutivo	[de ese modo]	horrela	0	1
	Conector consecutivo	[así las cosas]	horrela gauzak	0	1
	Conector consecutivo	[en consecuencia]	horren ondorioz	0	1
	Conector consecutivo	por lo que	[-(e)la eta]	3	0
	Conector consecutivo	lo que	[-(e)la eta]	1	0
	Conector consecutivo	para	[horretarako]	1	0
	Conector consecutivo	para ello	[horretarako]	1	0
	Conector consecutivo	mediante	[-(e)n bidez]	1	0
	Conector consecutivo	debido a	-(e)lako	1	1
	Conector aditivo + contraargumentativo	[y + en cambio]	eta + berriz	0	1
	Conector contraargumentativo	mientras que	aldiz	2	2
	Conector contraargumentativo	por el contrario	berriz	1	2
	Conector contraargumentativo	[sin embargo]	bestalde	0	2
	Conector contraargumentativo	frente a	[ostera]	1	0
	Conector contraargumentativo	pero	baina	1	2
	Conector contraargumentativo	aunque	arren	1	1
Reformuladores	Reformulador de distanciamiento	[aún todavía]	oraindik ere	0	1
	Reformulador recapitulativo	[en conclusión]	horrenbestez	0	1
Operadores discursivos	Operador de concreción	[por ejemplo]	hala nola	0	3
	Operador de concreción	como	[-(e)nez]	3	0



	Operador de concreción	de ellos	horien artean	1	2
	Operador de concreción	[de ellos]	horien artetik	0	1
	Operador de concreción	[entre otros]	besteak beste	0	1
	<b>Total</b>			<b>48</b>	<b>56</b>

Tabla 2. Clasificación de los marcadores en español y euskera

Somos conscientes de que algunos de los elementos incluidos en la Tabla 2 no entrarían en la categoría de marcador de Portolés, como por ejemplo “de ellos”, “entre otros”, “a la hora de”, etc., ya que este autor solo contempla elementos invariables que son periféricos a la predicación (Portolés, 2001). Sin embargo hemos considerado adecuado incluir estas unidades en nuestro estudio, ya que son claros marcadores de algunas de las relaciones discursivas de la RST. El marco de Montolío (2001) es más amplio y estas unidades cabrían en su categorización.

Una vez analizados los tipos de marcadores que reflejan cada relación en nuestro corpus, observamos también que, tal y como apunta van Dijk (1984: 101), hay determinados marcadores que pueden resultar ambiguos:

“Uno de los problemas de la semántica de los conectivos naturales es su posible ambigüedad: el mismo conectivo puede expresar diferentes clases de conexión, y una clase de conexión puede ser expresada por varios conectivos.”

Además de la conocida problemática de que las relaciones no están siempre señaladas (Taboada 2006), en el marco de nuestro trabajo la ambigüedad de los marcadores se hace notable cuanto más se utilizan. Tanto en español como en euskera los marcadores más frecuentes son los más ambiguos. En este corpus se utilizan más marcadores ambiguos en euskera (57,13 %) que en español (52,08 %), por lo que la ambigüedad aumenta en euskera un 5,05 %. Esta característica no es específica únicamente de los textos especializados, como los que integran el corpus de este estudio, sino que también es corriente en textos periodísticos (Iruskietta et al. en prensa).

En las Tablas 3 y 4 se muestran los marcadores ambiguos detectados en el corpus en euskera y en español, respectivamente. En la primera columna (MD) se incluyen dichos marcadores; en la segunda (Clasificación Portolés), su tipo; en la tercera (Nº), su frecuencia de aparición; en la cuarta (%), su porcentaje de aparición; en la quinta (Relación RST), el tipo de relación de la RST que marcan, y, en la sexta (Nº), el número de ocasiones que marcan cada relación.

MD	Clasificación Portolés	Nº	%	Relación RST	Nº
baina	Conector contraargumentativo	2	3,57	Antítesis Concesión	1 1
beraz	Conector consecutivo	2	3,57	Interpretación Elaboración	1 1
berriz	Conector contraargumentativo	2	3,57	Contraste Elaboración	1 1
era berean	Ordenador	8	14,28	Lista Elaboración	3 5
eta	Conector aditivo	16	28,57	Elaboración Lista Interpretación Resultado	4 10 1 1
horien artean	Ordenador	2	3,57	Elaboración Resultado	1 1
<b>Total</b>		<b>32</b>	<b>57,13</b>		<b>32</b>

Tabla 3. Marcadores ambiguos en el corpus en euskera

MD	Clasificación Portolés	Nº	%	Relación RST	Nº
como	Operador de concreción	3	6,25	Evidencia Elaboración	1 2
y	Conector aditivo	19	39,58	Lista Contraste	17 2
por lo que	Conector consecutivo	3	6,25	Motivación Causa	1 2
<b>Total</b>		<b>25</b>	<b>52,08</b>		<b>25</b>

Tabla 4. Marcadores ambiguos en el corpus en español

Las principales fuentes de ambigüedad detectadas son tres:

1. Las conjunciones “y” en español y “eta” en euskera son los marcadores más ambiguos en las dos lenguas con mucha diferencia.
2. A veces la detección de algunos marcadores no es determinante para que el anotador etiquete otra relación

- diferente a la Elaboración, como es el caso de (“beraz” y “berriz”).
3. Existen pequeñas diferencias entre relaciones como, por ejemplo, Motivación y Causa (ambas evidenciadas por el marcador “por lo que”) o Antítesis y Concesión (evidenciadas por el marcador “baina”).

## 5. CONCLUSIONES

Las conclusiones del estudio confirman los resultados de otros trabajos (en euskera Iruskieta et al. [en prensa], en español [da Cunha 2007], en inglés Taboada [2006]), al obtener un porcentaje de marcadores por relación de entre el 20 y el 25 % (en concreto, un 21,42 % en español y un 24,88 % en euskera). Sin embargo, el porcentaje de marcadores es algo mayor en euskera que en español, lo cual se debe probablemente a que en euskera las oraciones son más cortas (por lo que existe un mayor número de oraciones: 167 en castellano y 193 en euskera, es decir un 15,5 % más) y, al haber más unidades discursivas, estas deben cohesionarse mediante más marcadores.

Observamos también que hay determinados marcadores que pueden resultar ambiguos. En el marco de nuestro trabajo, esta situación es más evidente en euskera que en español.

Finalmente, concluimos que el porcentaje de relaciones discursivas de cada tipo que se encuentran en los textos en español y en euskera son muy similares, mientras que sí existen diferencias más relevantes con respecto a los tipos de marcadores encontrados en ambos corpus. Este hecho se debe principalmente a la utilización de diferentes estrategias lingüísticas para señalar una relación.

En este trabajo, realizamos un análisis riguroso y exhaustivo sobre nuestro corpus de análisis formado por textos paralelos del ámbito médico en español y euskera. Este corpus, aunque no es muy extenso, sí es suficiente para hacernos una idea de las diferencias y similitudes que existen entre las estructuras discursivas de ambas lenguas, en concreto en cuanto a las relaciones retóricas y a los marcadores del discurso incluidos en ellas.

## AGRADECIMIENTOS

Este trabajo ha sido realizado en el marco de las siguientes ayudas y proyectos: ayuda de movilidad posdoctoral otorgada por el Ministerio de Ciencia e Innovación de España (Programa Nacional de Movilidad de Recursos Humanos de Investigación; Plan Nacional de Investigación Científica, Desarrollo e Innovación 2008-2011) a Iria da Cunha; proyectos *Berbatek: Herramientas y Tecnologías para el impulso de la Industria de las Lenguas* (IE09-262), financiado por el Gobierno Vasco, y *OPENMT-2 Traducción automática híbrida y evaluación avanzada* (EHU) (TIN2009-14675-C03-01), financiado por el Ministerio de Ciencia e Innovación, en los que participa Mikel Iruskietia.

## NOTAS

<sup>1</sup> <http://www.gacetamedicabilbao.org/web/es/>

<sup>2</sup> En este trabajo utilizamos los términos “retórico” y “discursivo” como sinónimos.

<sup>3</sup> Para el inglés, véase Jucker y Ziv (1998).

<sup>4</sup> Entre el francés y el inglés hay un 10% de falta de emparejamiento o no coincidencia, tal y como indican Marcu et al. (2000: 9): “Yet, even between languages as close as English and French, there is a 10% mismatch in number of sentences -what is said in two sentences in one language is said in only one, or in three, in the other (Gale and Church 1993).”

## REFERENCIAS BIBLIOGRÁFICAS

- Croft, W.; Cruse, D.A. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Evans V.; Green, M. 2006. *Cognitive Linguistics. An Introduction*. Edinburgh: EUP.
- Eynon, T. 2002. “Cognitive linguistics”. *Advances in Psychiatric Treatment* 8. 399-407.
- Anscombe, J.C.; Ducrot, O. 1983. *L’argumentation dans la langue*. Bruselas: Mardaga.
- Austin, J.L. 1962. *How to do things with words*. Oxford: Clarendon Press.

- Austin, J.L. 1970. *Philosophical Papers*. Londres: Oxford University Press.
- Benveniste, E. 1966. *Problèmes de linguistique générale I*. París: Gallimard.
- da Cunha, I. 2007. “Importancia del marcaje de las relaciones discursivas para la generación automática de resúmenes”. En Cano, P.; Fernández, I.; González, M.; Prego, G.; Souto, M. (eds.) *Actas del VI Congreso de Lingüística General*. Vol 1 (*Métodos y aplicaciones de la lingüística*). Santiago de Compostela: Arco Libros. 835-846.
- da Cunha, I.; Iruskieta, M. 2009. “La influencia del anotador y las técnicas de traducción en el desarrollo de árboles retóricos. Un estudio en español y euskera”. En *Proceedings of the 7th Brazilian Symposium in Information and Human Language Technology*. São Carlos, Brasil: Universidade de São Paulo.
- Esnal, P. 2002. “Testu-antolatzaileak”. *Hizpide* 50. 25-36.
- Gale, W.A.; Church, K.W. 1993. “A program for aligning sentences in bilingual corpora”. *Computational Linguistics* 19 (1). 75-102.
- Grice, P. 1975. “Logic and conversation”. En Cole, P.; Morgan, J.L. (eds.). *Syntax and Semantics 3: Speech Acts*. Nueva York: Academic Press. 41-58.
- Halliday, M.A.K.; Hasan, R. 1976. *Cohesion in English*. Essex: Longman.
- Iruskieta, M.; Diaz de Ilarraza, A.; Lersundi, M. (en prensa). “Correlaciones en euskera entre las relaciones retóricas y los marcadores del discurso”. En *Actas del XXVII Congreso Internacional de la Asociación Española de Lingüística Aplicada (AESLA)*. Ciudad Real.
- Iruskieta, M.; Diaz de Ilarraza, A.; Lersundi, M. 2009. “Análisis de los marcadores del discurso para el euskera: Denominación, clases, relaciones semánticas y tipos de ambigüedad”. En Bretones, M.C. et al. (eds). *Applied Linguistics Now: Understanding Language and Mind*. Almería: Universidad de Almería. 1271-1282.
- Jucker, A.H.; Ziv, Y. 1998. *Discourse Markers: Descriptions and theory*. Amsterdam: John Benjamins.

- Labov, W.; Waletzky, J. 1967. "Narrative analysis: Oral versions of personal experience". *Essays on the verbal and visual arts*. Seattle: University of Washington Press. 12-44.
- Mann, W. C.; Thompson, S.A. 1988. "Rhetorical structure theory: Toward a functional theory of text organization". *Text* 8 (3). 243-281.
- Marcu, D. 2000. *The Theory and Practice of Discourse Parsing and Summarization*. Cambridge, MA: The MIT Press.
- Marcu, D.; Carlson, L.; Watanabe, M. 2000. "The automatic translation of discourse structures". En *Proceedings of the First Annual Meeting of the North American Chapter of the Association for Computational Linguistics*. 9-17.
- Martín Zorraquino, M.A.; Portolés, J. 1999. "Los marcadores del discurso". En Bosque, I.; Demonte, V. (eds). *Gramática descriptiva de la lengua española*. Vol. 3. 4051-4213.
- Montolío, E. (2001). *Conectores de la lengua escrita. Contraargumentativos, consecutivos, aditivos y organizadores de la información*. Barcelona: Ariel.
- Portolés, J. 2001. *Marcadores del discurso*. Barcelona: Ariel.
- Sperber, D.; Wilson, D. 1986. *Relevance*. Oxford: Blackwell.
- Taboada, M. 2006. "Discourse markers as signals (or not) of rhetorical relations". *Journal of Pragmatics* 38. 567-592.
- van Dijk, T.A. 1980. *Macro-Structures. An interdisciplinary study of global structures in discourse, cognitions and interaction*. Hillsdale NJ: Erlbaum.
- van Dijk, T.A. 1984. *Texto y contexto (Semántica y pragmática del discurso)*. Madrid: Cátedra.
- van Dijk, T.A. 1989. *La ciencia del texto*. Barcelona: Paidós.