

# DISCOURSE RELATIONAL DEVICES (LPTS2016)

Linguistic & Psycholinguistic Approaches  
To Text Structuring

<http://lpts2016.blogs.uv.es/>

## ABSTRACTS



Universitat de València (Spain)

24-26 January 2016

# DISCOURSE RELATIONAL DEVICES (LPTS2016)

Conference chair:: Maria Josep Cuenca (Universitat de València)

Organizing committee

Salvador Pons (Universitat de València); Catherine Bolly (Université catholique de Louvain); Liesbeth Degand (Université catholique de Louvain)

Local organizing committee

Adrián Cabedo (Universitat de València); Maria Estellés (Universitat de València); Laura Eliodoro (Universitat de València); Marta Albelda (Universitat de València); Ana Llopis (Universidad Católica de València);

Scientific committee

Karin Aijmer (Göteborg University); Nicholas Asher (CNRS, IRIT, Toulouse); Carla Bazzanella (Università di Torino); Margarita Borreguero (Universidad Complutense de Madrid); Antonio Briz (Universitat de València); Luis Cortés (Universidad de Almería); Manuel Casado Velarde (Universidad de Navarra); Laurence Danlos (Université Paris Diderot); Liesbeth Degand (Université catholique de Louvain); Gabriele Diewald (Leibniz Universität Hannover); Jacqueline Evers-Vermeul (Universiteit Utrecht); Kerstin Fischer (University of Southern Denmark); Catalina Fuentes (University of Seville); Volker Gast (University of Jena); Mikel Iruskieta (University of the Basque Country); Araceli López Serena (Universidad de Sevilla); Óscar Loureda (University of Heidelberg); Ana Cristina Macário Lopes (Universidade de Coimbra); David Martins de Matos (Universidade de Lisboa); Caterina Mauri (Università di Pavia); Amália Mendes (Universidade de Lisboa); Jiří Mírovský (Charles University Prague); Paola Pietrandrea (University of Tours); José Portolés (Universidad Autónoma de Madrid); Ted Sanders (Universiteit Utrecht); Wilbert Spooren (Radboud Universiteit Nijmegen); Manfred Stede (University of Potsdam); Jacqueline Visconti (University of Genoa); Richard Waltereit (Newcastle University); Sandrine Zufferey (University of Fribourg)

## PLENARY SESSIONS

DISCOURSE RELATIONAL DEVICES IN TEXTLINK: FROM (CATEGORIAL) DESCRIPTION TO CORPUS ANNOTATION, AND BACK AGAIN	
Liesbeth Degand.....	8
DEFINITIONS OF DISCOURSE MARKERS AND THEIR FUNCTIONS AS TEXT-STRUCTURING DEVICES	
Kerstin Fischer .....	8
COMBINATIONS OF DISCOURSE MARKERS	
Bruce Fraser.....	9
DOING CORPUS-BASED TYPOLOGY: CONCESSIVITY IN A CROSS-LINGUISTIC PERSPECTIVE	
Volker Gast.....	10
DISCOURSE PARTICLES AND COGNITION: AN EXPERIMENTAL APPROACH	
Óscar Loureda.....	11

## ABSTRACTS

MODAL AND EVIDENTIAL DEVICES AS STRUCTURE-SIGNALLING DEVICES IN EARLY ENGLISH TEXTS	
Francisco Alonso-Almeida and María Luisa Carrió Pastor.....	13
COMPARISON OF METADISOURSE MARKERS IN ARABIC AND ENGLISH RESEARCH ARTICLES IN THE INTRODUCTION AND CONCLUSION SECTIONS	
Hmoud Alotaibi .....	14
CONSTRUCCIONES CAUSALES IMPLÍCITAS EN ESPAÑOL E ITALIANO	
Ignacio Arroyo Hernández.....	15
DISCOURSE RELATIONS, DISCOURSE CONNECTIVES AND DISCOURSE SEGMENTATION INTERDEPENDENCY IN THE LIGHT OF CAUSALITY	
Caroline Atallah, Myriam Bras and Laure Vieu.....	16
MDMA: CORPUS-BASED PARAMETERS FOR THE IDENTIFICATION OF DISCOURSE MARKERS IN SPOKEN FRENCH	
Catherine Bolly, Ludivine Crible, Liesbeth Degand and Deniz Uygur-Distexhe .....	18
DISCOURSE MARKERS IN REPORTED SPEECH	
Margarita Borreguero Zuloaga; Ilaria Fiorentini; Elena Martínez Caro.....	19
'EN REALIDAD' AS A CROSSING GATE: MULTIFUNCTIONALITY AND POLYSEMY OF A DM	
Amalia Canes Nápoles and Nicole Delbecque .....	21
LES MARQUEURS DISCURSIFS DE LA RELATION D'AJOUT. UNE ÉTUDE CONTRASTIVE FRANÇAIS-ITALIEN	
Anna-Maria De Cesare .....	23

CAUSAL CATEGORIES IN TURKISH CONNECTIVES; FIRST RESULTS FROM A CORPUS STUDY AND A JUDGEMENT TASK	
Derya Cokal, Deniz Zeyrek and Ted Sanders .....	25
WHAT DOES "CONNECTIVE" MEAN CROSS-LINGUISTICALLY?	
Margot Colinet, Tatjana Scheffler, Manfred Stede and Laurence Danlos .....	27
POSITION AND CONTEXT OF STRUCTURING DISCOURSE MARKERS: A CORPUS-BASED INVESTIGATION OF SCOPE AND COGNITIVE LOAD IN FRENCH AND ENGLISH	
Ludivine Crible.....	29
PROCESSING OF EXPECTED/UNEXPECTED FOCUS IN PRAGMATIC SCALES IN SPANISH: EXPERIMENTAL NOTES ABOUT THE SPANISH FOCUS PARTICLE INCLUSO	
Adriana Cruz Rubio .....	31
ANNOTATION OF MULTI-WORD DISCOURSE MARKERS IN SPOKEN SLOVENE	
Kaja Dobrovoljc.....	33
COHESIVE MARKERS OF CONTRAST IN ENGLISH AND FRENCH: WHAT, WHERE AND WHY?	
Maité Dupont .....	34
EL MARCADOR DISCURSIVO COLOQUIAL BUENO: FUNCIONES PRAGMÁTICAS EN CONVERSACIÓN COLOQUIAL Y DESCRIPCIÓN EN MANUALES DE E/LE DE NIVEL B2.	
Laura Eliodoro-Furió.....	35
APPLYING VISUALISATION TOOLS TO ANALYSE TURN-TAKING SYSTEMS IN INFORMAL CONVERSATIONS AND POLITICAL INTERVIEWS	
Guadalupe Espinosa-Guerri and Amparo Garcia-Ramon .....	36
THE PRESENCE AND ABSENCE OF DRDS IN BIBLE TRANSLATIONS: THE IMPLICIT AND EXPLICIT MARKING OF ADDITION, CONTRAST, AND CAUSE	
Péter Furkó .....	37
WHEN "SAME" MEANS SOMETHING DIFFERENT: THE SIGN SAME AS A DISCOURSE RELATIONAL DEVICE IN TWO SIGN LANGUAGES	
Sílvia Gabarró-López .....	38
THE TENA KICHWA =MI AS A MARKER OF SPEAKER AUTHORITY	
Karolina Grzech .....	39
SEGMENTING DISCOURSE UNITS: INCORPORATING INTERPRETATION INTO DECISION RULES?	
Jet Hoek, Jacqueline Evers-Vermeul and Ted Sanders.....	41
DISCOURSE RELATIONS AND STRATEGIC USE OF EXTRA-CLAUSAL CONSTITUENTS IN CO-CONSTRUCTED ARGUMENTATIVE DISCOURSE	
Carolin Hofmockel, Anita Fetzer and Robert Maier.....	42
DIFFERENT SYSTEMS OF ORGANIZATION OF SPANISH DISCOURSE MARKERS AND A NEW PROPOSAL FOR SPANISH AS AN L/2	
Anais Holgado Lage.....	43
MORPHOLOGICAL AND SYNTACTIC CHARACTERISTICS OF DISCOURSE CONNECTIVES IN CZECH	
Pavlina Jinova .....	44
COGNITIVE COMPLEXITY IN CAUSAL RELATIONS IN ADULT L2-WRITING	
Janne Laitinen .....	46

THE FUNCTIONAL EQUIVALENTS OF FRENCH POURTANT IN ENGLISH AND PORTUGUESE	
Pierre Lejeune .....	47
NEGATIVE EMOTION PRAGMATIC MARKERS – CORPORA AND CROSS-LINGUISTIC CONTRASTS	
Barbara Lewandowska-Tomaszczyk and Paul A. Wilson .....	49
LAS HIPÓTESIS SOBRE LA GRAMATICALIZACIÓN DE POR CIERTO COMO MD EPISTÉMICO EN ESPAÑOL: ORALIDAD, ESCRITURALIDAD Y PERSPECTIVA COGNITIVO-FUNCIONAL	
Araceli López-Serena .....	51
WHAT DOES CONTEXT MEAN WHEN REFERRING DISCOURSE MARKERS?	
Ana Llopis Cardona .....	52
PRESUPPOSITION AND SEMANTIC CHANGE: THE CASE OF PERO	
Patrícia Matos Amaral and Susana Rodríguez Rosique. ....	53
THE ROLE OF EXEMPLIFICATION IN DISCOURSE	
Caterina Mauri and Alessandra Barotto .....	55
PROCESSING CAUSALITY IN SPANISH: THE CASE OF "POR TANTO" AND "POR ESO"	
Laura Nadal and Inés Recio Fernández .....	56
RHETORICAL RELATIONS MARKED BY 'YOU KNOW' IN THE CONTEXT OF HUMOR	
Anna Nagy.....	58
HESITATIONS, REPETITIONS, RESTARTS AND SELF-CORRECTIONS: TOWARDS A SYSTEMATIZATION OF DISCOURSE PLANNING PHENOMENA IN SPANISH COLLOQUIAL CONVERSATIONS	
Elena Pascual Aliaga .....	59
MARCADORES Y PARTÍCULAS DISCURSIVAS INTERACTIVAS EN TWITTER	
María Soledad Padilla Herrada .....	61
FROM ANAPHORA TO DISCOURSE RELATIONAL DEVICES	
Bernardo E. Pérez Alvarez .....	62
CONNECTIVE MOVEMENT (A DEMONSTRATION ON CZECH)	
Lucie Poláková.....	64
THE COMBINATION OF DISCOURSE MARKERS AND DISCOURSE UNITS: NOT A MORGANATIC MARRIAGE	
Salvador Pons.....	66
"RAZÓN DE MÁS" COMO INVERSOR ARGUMENTATIVO	
José Portolés .....	67
WHEN SEMANTICS AND PRAGMATICS DIVERGE: THE TRANSLATION OF ADVERBIAL CONNECTORS	
Beatriz Sánchez-Cárdenas .....	69
THE USE OF A SPANISH LEXICON OF SUBJECTIVITY FOR THE ANALYSIS OF DISCOURSE RELATIONAL DEVICES (DRDS) IN SPANISH ACADEMIC TEXTS	
Andrea Santana Covarrubias .....	71
"ON THE ONE HAND" AS A CUE TO ANTICIPATE UPCOMING DISCOURSE STRUCTURE	
Merel Scholman; Hannah Rohde and Vera Demberg .....	73

LAS CONJUNCIONES ILATIVAS ASÍ QUE Y CONQUE COMO MARCADORES DE MODALIDAD EVIDENCIAL Eugenia Sainz .....	75
DISCOURSE MARKERS IN THE VAL.ES.CO. SEGMENTATION MODEL: FUNCTIONS AND USES OF "PUES ESO" IN SPANISH COLOQUIAL CONVERSATIONS Shima Salameh Jiménez .....	77
UNIFYING DIMENSIONS IN DISCOURSE RELATIONS: HOW VARIOUS ANNOTATION SCHEMES ARE RELATED Ted Sanders, Vera Demberg, Jacqueline Evers-Vermeul, Jet Hoek, Merel Scholman, Fatemeh Torabi Asr and Sandrine Zufferey .....	79
CONTRIBUTOS PARA A ANÁLISE DO CONECTOR "SIM" EM PORTUGUÊS EUROPEU CONTEMPORÂNEO Sara Sousa .....	81
CONCESSIVE MARKING REVISITED: 'BUT' IN COURTROOM TALK Magdalena Szczyrbak.....	83
PHATIC MARKERS AS CONVERSATION REGULATORS IN INTERVIEWS AND INFORMAL CONVERSATION Gloria Uclés Ramada .....	84
ADVERSATIVE CONJUNCTIONS IN LITHUANIAN AND ENGLISH Giedre Valūnaite-Oleskeviciene; Prof. Dr. Jolita Sliogeriene; Vilma Asijaviciute. ....	85
PRAGMATIC FUNCTIONS OF DISCOURSE RELATIONAL DEVICES IN ONLINE CUSTOMER REVIEWS Tuija Virtanen .....	86
PRESENT DAY ITALIAN ANZI IN A CROSS-LINGUISTIC PERSPECTIVE Jacqueline Visconti.....	87
PERSPECTIVE MARKING AND SUBJECTIVITY IN COHERENCE RELATIONS: A COLLOCATION ANALYSIS OF CHINESE CONNECTIVES Yipu Wei and Jacqueline Evers-Vermeul.....	89

# PLENARY SESSIONS

## Discourse Relational Devices in TextLink: From (categorical) description to corpus annotation, and back again

Liesbeth Degand

The presentation will give an overview of the challenges facing the members of the COST TextLink Action (IS1312). The aim of the network is to interconnect the multilingual discourse annotated corpora developed (or under development) across Europe in order to, among other things, delineate the dimensions and properties of discourse annotation across corpora; organise these properties into a sharable taxonomy; encourage the use of this taxonomy in subsequent discourse annotation and in cross-lingual search and studies of devices that relate and structure discourse.

First results will be presented and illustrated on the basis of spoken and written data.

### Definitions of Discourse Markers and their Functions as Text-Structuring Devices

Kerstin Fischer

In previous papers (Fischer 2006, 2014), I have suggested that the definition of discourse particles/ discourse markers should be based on a set of dimensions that describe the characteristics of the items under consideration. These dimensions include the degree of integration of a discourse marker, its context dependency, its affinity to written versus spoken contexts and its functional spectrum. Such a dimension-based approach accounts naturally for the facts that discourse particles/markers form a continuum instead of discrete classes, that many of them participate in more than one class, that historically they migrate from class to class, and that often one particle instance can fulfill several different functions at the same time. Furthermore, it turns out that the dimensions according to which discourse particles can be described are highly correlated; therefore, these dimensions contribute to systematizing the field.

Now, crucially, also the text-structuring functions these particles/markers fulfill correlate with the dimensions identified since these functions depend on what units discourse marker uses refer to, i.e. on their domain of reference, which is related to the characteristics described by the dimensions. Thus, the dimensions identified also contribute indirectly to an understanding of the types of text-structuring functions a given marker fulfills.

Finally, we need to consider the impact of the language-specific resources available on the background of which the functions of discourse markers are possible. That is, languages deal differently with the different dimensions of integration and context-dependency, which has an impact on what kinds of units discourse markers structure. Construction Grammar has been proven useful to capture the generalizations observable and to describe the particular slices from the functional spectrum that particular language-specific structural properties single out.



## Combinations of Discourse Markers

Bruce Fraser

I start from the assumption that there is a group of lexical expressions in every language called Discourse Markers. These expressions may occur as a part of a discourse segment, have meaning, but are not part of the propositional content of the message conveyed. Members of this class typically have the following properties: they are free morphemes, discourse segment-initial, and they signal a specific message either about or in addition to the interpretation of the segment.

Discourse Markers divide up into two sub-groups, with additional divisions:

### Discourse Conjoining Markers

1. Contrastive (but, instead, rather)
2. Elaborative (and, furthermore, in addition)
3. Implicative (so, thus, as a result)

### Discourse Structuring Markers

1. Discourse Management Markers (in summary, I add, actually, all in all)
2. Topic Orientation Markers (back to my point, as I was saying, by the way)
3. Attention Markers (look anyway, then, oh, ok, so)

In this lecture I will present an analysis of the combining properties of these two groups of Discourse Markers: the order in which they occur in combination, where they occur (sentence-initial/sentence-final/separated), the interpretation of the combinations, and what properties account for those that do combine. For example,

He wanted to go to the movies but instead, he stayed at home.

She couldn't find a pen. So instead, she used a pencil.

I don't like your argument, but, as I was saying, let's move on.

A: The project didn't work out. B: So, all in all, it was a complete failure.

A: John won first prize in the race. B: Oh, wow, Ok, let's go congratulate him.

Sarah was discouraged at the result. But, after all, what could she expect at her age.

I'm glad you're here. And oh, can you help me this

A: The house is empty and smells. B: Well, as I was saying, you shouldn't expect too much.

Although I do not expect to touch on every instance of Discourse Marker combinations, I hope to provide the basis of an analysis which others can fill in.

## Doing corpus-based typology: Concessivity in a cross-linguistic perspective

Volker Gast

Semantically speaking, concessivity is known to be one of the most complex discourse relations, e.g. insofar as it implies background expectations and the cancellation of such expectations. This semantic complexity is reflected in the fact that concessive relations are often encoded 'compositionally', e.g. by combining a scalar additive operator with a conditional clause ('Even if it rains, I will go out'), or a non-specific free relative clause with a coreferential pronoun ('Whatever you are selling, I'll buy it', see for instance Haspelmath & König 1998). For a corpus-based approach to discourse linkers, such 'distribution' of information over various components of a sentence represents a non-trivial challenge, as there is no specific markable to which the expression of concessivity can be attributed. At the same time, the compositionality of strategies of concessive marking offers an interesting way of classifying these strategies crosslinguistically, e.g. in the spirit of Haspelmath & König (1998). In this talk I will propose a way of annotating examples of 'distributed concessivity' in natural language corpora, using data from a Mayan language (Tzotzil) and various European languages. The annotations will be implemented with GraphAnno, a tool for manual multi-level annotation (cf. Gast et al. 2015). On the basis of a distributed annotation scheme, I will moreover propose a way of identifying instances of distributed concessivity in corpora of English.

### References

- Gast, V., L. Bierkandt and C. Rzymiski (2015). [Annotating modals with GraphAnno, a configurable lightweight tool for multi-level annotation](#). In M. Nissim & P. Pietrandrea (eds.), Proceedings of the Workshop on Models for Modality Annotation, 19-28. Stroudsburg, PA : Association for Computational Linguistics (ACL).
- Haspelmath, M. & E. König (1998). Concessive conditionals in the languages of Europe. In: van der Auwera, Johan (ed.), Adverbial Constructions in the Languages of Europe. Berlin: Mouton de Gruyter, 563-640.

## Discourse particles and cognition: an experimental approach

Óscar Loureda

It is well accepted in linguistics in general and specifically in pragmatics, that—for a number of reasons—not all utterances require the same processing effort. To approach the study of the dynamics of communication, the cognitive-oriented Relevance Theory provides a suitable frame: utterances are underdetermined stimuli communicated ostensively by a speaker and enriched by a hearer/reader, who builds a mental representation thereof. Such mental representations arise, however, from the interaction of conceptual meaning, i.e. the lexical information of the propositional content of utterances, and procedural meaning, that is to say, instructions on how to process conceptual meanings and carry out the needed inferential computations to arrive to the communicated assumption.

Discourse particles display precisely a mainly procedural meaning: they serve as linguistic cues to guide the inferential processes in communication and contribute to the interpretation of utterances. They are intentionally used to carry out formulation operations (e.g. that is to say, namely), at a structural level as information-organizing devices (firstly, secondly, finally), as argumentative devices (therefore, however, after all) or as information-structuring operators (even, too, at least).

Furthermore, and given such semantic and pragmatic features, it seems plausible that discourse particles constitute an attentional focus during processing. In this talk, I will address this question and will provide experimental results that help to ascertain the contribution of discourse particles for the dynamics of communication. Specifically, I will refer to results obtained in eye-tracking processing and reading comprehension experiments in which the role of the Spanish information-structuring operators *incluso* ('even') and *también* ('also') and the argumentative connectives *por tanto* ('therefore') and *sin embargo* ('however') was analyzed. Results show that discourse particles contribute to constrain the inferential processes in communication and to homogenize information processing patterns.

# ABSTRACTS

## Modal and evidential devices as structure-signalling devices in early English texts

Francisco Alonso-Almeida and María Luisa Carrió Pastor

### Abstract

The present paper seeks to identify modal and epistemic devices with a discourse organising function. For this purpose, I will focus on the Corpus of eighteenth century History texts (CHET), compiled as a subsection of the Coruña Corpus, currently under development at the University of Coruña (Spain) (Moskowich and Crespo, 2007). The corpus is to be used with its own corpus tool, i.e. the Coruña Corpus Tool (CCT) for text retrieval and analysis. Although the study obviously relies on corpus methodology, manual inspection is also performed to disambiguate modal and evidential cases. My approach to evidentiality is disjunctive (cf. Dendale and Tasmowski, 2001) in the sense that it stands as a distinct category from epistemic modality, even if functional overlapping may result from truth-value readings of particular samples. I want to show how productive modal and evidential items are for structuring and developing the text with a focus on cohesiveness (Alonso Almeida, 2015: 138). Furthermore, I want to detect their frequency of usage with this purpose, and also whether they own further particular discourse and pragmatic functions. Conclusions will report on the categorisation of these discourse and pragmatic functions, and on the interactive nature of evidentials and modals to elaborate and construe meaning.

### References

- Alonso Almeida, F (2015) The functions of seem and parecer in early medical writing. *Discourse Studies* 17.2; 121–140.
- Dendale P and Tasmowski L (2001) Introduction: Evidentiality and related notions. *Journal of Pragmatics* 33: 339–348.
- Moskowich, I and B Crespo (2007) Presenting the Coruña Corpus: A Collection of Samples for the Historical Study of English Scientific Writing. In Pérez Guerra, Javier et al. (eds.) *Of Varying Language and Opposing Creed: New Insights into Late Modern English*. 341–357. Bern: Peter Lang.

## Comparison of Metadiscourse Markers in Arabic and English Research Articles in the Introduction and Conclusion Sections

Hmoud Alotaibi

### Abstract

Metadiscourse is known as language about language. It can be conceptualized as a framework or a model of analysis which determines particular rhetorical devices or strategies used by writers. Ädel (2006) explained metadiscourse as 'text about the evolving text, or the writer's explicit commentary on her own ongoing discourse, displaying an awareness of the current text or its language per se, and of the current writer and reader qua writer and reader' (p. 183). Little is known about metadiscourse in Arabic and English research articles written by Arabic-speaking writers. Therefore, this paper uses Ädel's (2006) framework to investigate metadiscourse markers in the introduction and conclusion sections of research articles written in Arabic and in English by native speakers of Arabic. The findings indicated that the text-oriented metadiscourse was more preferable across corpora and particularly was more prominent in introductions compared to conclusions. The findings also revealed cross-linguistic variations as English texts used a higher number of metadiscourse markers than their Arabic counterparts. The study closes with several recommendations for future studies and addresses pedagogical implications.

## Construcciones causales implícitas en español e italiano

Ignacio Arroyo Hernández

### Abstract

Se han sugerido numerosos marcos teóricos para describir las relaciones causales marcadas explícitamente, con frecuencia en compañía de otros tipos semánticos de cláusulas, como condicionales y temporales. Las construcciones causales donde las relaciones cohesivas entre preferencias no están explícitamente marcadas han recibido menos atención. Pocos son los estudios disponibles que afrontan la motivación tras la marcación de las relaciones causales en el discurso (Gohl 2000, Corminboeuf 2010, Diessel y Hetterle 2011, Chang y Su 2012), puesto que, por lo que a la implícitación se refiere, el interés parece recaer básicamente en las cuestiones psicolingüísticas que se refieren al esfuerzo de procesamiento, al tiempo de procesamiento o a la calidad de los free recalls.

Por lo que nos consta, a día de hoy no existen estudios específicos acerca de las relaciones causales implícitas en español e italiano- Concentrándonos en estas dos lenguas, emergen preguntas de investigación como:

1) ¿Cuándo podemos decidir que dos o más segmentos forman una relación causal cohesiva?, 2) ¿Cómo se señala esta cohesión?, 3) ¿Hacen los hablantes italianos y españoles un uso similar (en términos de tipología de las estructuras, frecuencias, contextos discursivos, etc.?, 4) ¿Tiene la variable modalidad oral vs. escrito incidencia en la dicotomía implícito-explicito?, 5) Si consideramos que la presencia de una marca explícita vehicula verosímelmente funciones pragmáticas, ¿cuál es la relación entre las construcciones implícitas las y construcciones explícitas o paráfrasis que se proponen frecuentemente como legítimas "versiones"? ¿Cuándo y por qué deciden los hablantes añadir marcas explícitas?

Se recogerán datos procedentes de varios corpus contemporáneos: datos orales del corpus C-ORAL-ROM (Corpus di parlato spontaneo delle principali lingue romanze), del corpus CLIPS y del corpus C-OR-Dial, y datos escritos del corpus CREA, del Corpus del español (Mark Davis) y del corpus CORIS/CODIS.

La indeterminación semántica prevalente en las relaciones causales implícitas requiere de los hablantes un fuerte enriquecimiento contextual que puede ser examinado a través de varios paradigmas teóricos bien conocidos (Grice, modelos neo-griceanos, Teoría de la Relevancia, etc.), pero el papel central que la Gramática Metaoperacional, propuesta por Adamczewski, asigna a la interacción enunciador-coenunciador en la construcción de significado sugiere que puede constituir un paradigma útil para nuestra investigación.

### Referencias

- Adamczewski, H., *Genèse et développement d'une théorie linguistique, suivi de Les dix composantes de la grammaire métaopérative de l'anglais*, La Tilv Éd, Perros-Guirec, 1996.
- Chang M.S, Su, H., To mark or not to mark the cause, that is the question: Causal marking in Taiwanese conversation, *Journal of Pragmatics*, Volume 44, Issue 13, October 2012, 1743-1763.
- Corminboeuf, G., 2000. La causalité sans les connecteurs 'causaux'. Préalables épistémologiques, in Johnsen L. A., Corminboeuf G. & Conti V. (eds), *Entre syntaxe et discours*, Linx.
- Diessel, H., & Hetterle, K. 2011. Causal clauses: A cross-linguistic investigation of their structure, meaning, and use. *Linguistic universals and language variation*, 21-52.
- Gohl, Christine, 2000. Causal relations in spoken discourse: asyndetic constructions as a means for giving reasons. In: Couper-Kuhlen, E., Kortmann, B. (Eds.), *Cause, Condition, Concession, Contrast*. Mouton de Gruyter, New York, pp. 83-110.
- Ford, Cecilia E., Mori, J., 1994. Causal markers in Japanese and English conversations: a cross-linguistic study of interactional grammar. *Pragmatics* 4 (1), 31-62.

## Discourse relations, discourse connectives and discourse segmentation interdependency in the light of causality

Caroline Atallah, Myriam Bras and Laure Vieu

### Abstract

Many discourse theories are based on the existence of discourse relations (DRs) linking discourse units. These DRs can be triggered by discourse markers, like connectives. For example, in French, the connective 'parce que' ('because') triggers a causal DR:

(1) [Pierre est parti] $\alpha$  [parce que Marie a crié.] $\beta$

'[Pierre left] $\alpha$  [because Marie yelled.] $\beta$ '

Identifying discourse relational devices (DRDs) associated to a DR (or a group of DRs) implies to know when there is a DR. As a DR holds between two different discourse segments –  $\alpha$  and  $\beta$  in (1) –, identifying a DR implies first segmenting. Therefore, DRDs, DRs inference and discourse segmentation are interdependent issues that must be analyzed together. However, segmentation issues are rarely focused on in the literature on DRDs and DRs.

In this presentation, we will discuss the problem of segmentation across the causality spectrum in the well-formalized theoretical framework of SDRT (Segmented Discourse Representation Theory, Asher & Lascarides, 2003). We will illustrate it with the following examples:

(2) Les cris de Marie ont causé le départ de Pierre.

'Marie's yells caused Pierre's leaving.'

(3) Pierre est parti à cause des cris de Marie.

'Pierre left because of Marie's yells.'

(4) Marie a crié, causant le départ de Pierre.

'Marie yelled, causing Pierre's leaving.'

(5) Marie a crié, ce qui a causé le départ de Pierre.

'Marie yelled, which caused Pierre's leaving.'

(6) Marie a crié. Ces cris ont causé le départ de Pierre.

'Marie yelled. These yells caused Pierre's leaving.'

(7) Marie a crié. Cela a causé le départ de Pierre.

'Marie yelled. This caused Pierre's leaving.'



In each case, like in (1), a causal relation holds between the same two events: 'les cris de Marie' ('Marie's yelling') and 'le départ de Pierre' ('Pierre's leaving'). The aim of this presentation will be to determine whether the causal relation is the same in all cases and specifically whether this relation is a DR.

If no precise definition is given to characterize an elementary discourse unit (EDU), most theories agree on the fact that this unit approximately matches a syntactic clause. According to this criterion, (2) and (3) should not be segmented. So, how can we account for the causal relation involved in these examples? Should we consider that they are causal DRs and segment the discourse despite of the syntactic criterion violation? Or should we consider that they are not causal discourse relations but just semantic relations between events? According to the first option, 'à cause de' ('because of') and 'ont causé' ('caused') would play a discourse marker role, like 'parce que' in (1), i.e. they would trigger a causal DR.

Unlike (2) and (3), examples (4), (5), (6) and (7) count two syntactic clauses separated by a punctuation mark. If we follow the common syntactic criterion, we have to segment them into two EDUs and thus to assume that a DR holds between the two EDUs. As these examples do not involve phrases traditionally known as causal connectives, one might wonder about the nature of the DR involved.

Therefore, the characterization of EDUs and that of DRs are strictly interdependent tasks. They also question what is a discourse marker and more specifically a discourse connective. In this presentation, we will offer some criteria to distinguish DRs from strictly semantic relations as well as clear segmentation rules. We will also discuss the nature of a variety of causality markers in French, distinguishing DRDs among them.

#### References

ASHER, N., LASCARIDES, A. (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.

## MDMA: corpus-based parameters for the identification of discourse markers in spoken French

Catherine Bolly, Ludivine Crible, Liesbeth Degand and Deniz Uygur-Distexhe

### Abstract

Starting from the common observation that there is no recognized closed class of Discourse Markers (DMs) and that their definition may vary from one theoretical framework to another (Schourup 1999: 228), the MDMA project ("Model for Discourse Marker Annotation") has established an empirical corpus-based method for the identification and annotation of DMs in spoken French. Central to our proposal is that DMs may be described as clusters of features that, in specific patterns of combination, allow distinguishing between more or less prototypical uses of DMs in context. The general goal of MDMA is to cover every step of the analysis of DMs from identification to parameter and functional description, while its ultimate perspective is to provide reliable clusters to feed machine-learning systems for semi-automatic identification of DMs in authentic data (cf. Fischer 2000 for a similar approach to functional category assignment).

We proceeded in three steps : (i) manual identification of all so-called "potential" DMs in a balanced corpus of spoken French (5,000 words; Belgium and France) ; (ii) automatic extraction from the corpus of every token corresponding to the candidate DMs previously identified (1,181 tokens) ; and (iii) parameter analysis of a random sample of 200 potential DMs (syntactic, formal and semantic-pragmatic variables). In line with the objectives of this conference, our hypothesis is that statistical analysis – based on the distributional constraints of the potential DMs at stake – should uncover a certain hierarchy between the different features under scrutiny, regarding their relevance, reliability, and generalizability (or even specificity). Our results show that syntactic position is the most predictive feature that correctly identifies a potential DM (which is encouraging for NLP purposes), and that the statistically most prototypical profile of DM is the following: autonomous, post-final, pre-initial or indeterminable position; with a procedural meaning; in initial or medial position in the conversational turn; not expressing its coded meaning; not mobile. This pattern is exemplified below:

(1) bon moi heureusement je n'ai pas eu de choses très graves mais enfin quand même / on sent / on commence à se sentir qu'on / qu'on / qu'on diminue / qu'on ne sait pas faire ce qu'on veut (Corpage : corpus ageNM1)

Less prototypical profiles either correspond to propositional equivalents of DM tokens (example 2), or borderline expressions at the edges of several pragmatic categories such as modal particles or response signals (example 3):

(2) le loup s'empresse/ chemin pour arriver chez la mère-grand avant le petit chaperon rouge mais là faut mettre pas pris le bon chemin /euh/ (Clapi : corpus Chaperon Rouge, "Jean-Pierre et Magali")

(3) ah à cause des sans papiers à Saint Bernard CEC voilà Saint Bernard CG d'accord (Clapi : corpus Adi étudiants)

In this poster, we first describe the annotation scheme and procedure, then we illustrate several issues of inter-rater agreement, and finally present the main results from the statistical analyses. Pointers to some (functional and multimodal) perspectives of the MDMA project will also be mentioned.

### References

- Balthasar L. & Bert M. 2005. La plateforme « Corpus de langues parlées en interaction » (CLAPI), *Lidil*, 31. URL : <http://lidil.revues.org/139>
- Bolly C. T., Masse M. & Meire P. 2012. Corpage. Reference corpus for the elderly's language. Louvain-la-Neuve: Université catholique de Louvain.
- Fischer, K. 2000. From cognitive semantics to lexical pragmatics. Berlin: Mouton de Gruyter.
- Schourup, L. 1999. Discourse markers. *Lingua* 107: 227-265.

## Discourse markers in reported speech

Margarita Borreguero Zuloaga; Ilaria Fiorentini; Elena Martínez Caro

### Abstract

The present contribution proposes a first survey of the use of Italian and Spanish discourse markers (henceforth DMs) in reported speech. One of the features of DMs is that they can be deleted in indirect speech without affecting the propositional content of the utterance. This is mainly due to the fact that DMs typically do not add anything to the propositional content of the utterance and therefore can be omitted in reported speech (cf. Bazzanella 1995, Molinelli 2014).

By indirect speech we mean the reproduction of utterances belonging to another locutionary act (Maldonado 1991, Reyes 2002, Güldeman/von Roncador 2002, Mandelli 2010a), without quoting them directly, as in the following examples, where the DMs *guarda* 'look' in (1) can be omitted in (2):

Guarda, ti sbagli

Ha detto che mi stavo sbagliando

However, DMs, especially interactional and intersubjective markers are accepted in direct (reported) speech (Mandelli 2010b): in this case, they can be employed as quotation markers (metatextual function, Bazzanella 2006), in order to be more similar to the original speech, as in (3-4):

infatti me l'ha presentato dice guarda quando tu hai bisogno visto visto che c'hai il deposito qui a a allo Sheraton se hai bisogno dice ti rivolgi qui a al dottor Bruschi (Ghezzi/Molinelli 2015)

llega un momento que vas aguantando y que las cosas se juntan y que dices/ PUES NO / TENGO QUE PARARME / Y – Y DECIDIR (Briz/Val.Es.Co. 2002)

In some cases, it is difficult to determine whether the DM belongs to the speaker's direct speech or to direct reported speech, as in (5), and this ambiguity is closely intertwined with the type of DM:

*ha detto che* i movimenti erano, *per così dire*, involontari, che il cervello era, *per così dire*, passivo

Thus, we will try to assess which DM appear more frequently in direct reported speech and which their functions are, as well as which DM are systematically excluded in this particular use. We carry out a corpus-based research collecting our data from different spoken corpora (LIP, C-Oral-Rom and CLIPS for Italian, Val.Es.Co and COLA for Spanish). In our analysis, we will adopt a contrastive perspective, comparing DM with interactional and metatextual functions such as It. *guarda*, Sp. *mira* 'look', It. *beh*, Sp. *bueno* 'well', It. *ma*, Sp. *pues* 'but', It. *per così dire*, Sp. *o sea* 'that is to say, so to say'.

Our claim is that these particular DMs are employed in order to convey the original illocutionary force of the reported utterances and therefore they cannot be deleted without affecting their discourse meaning.

## References

- Bazzanella, C. (1995): I segnali discorsivi. In Renzi, L., Salvi, G., Cardinaletti, A. (eds.), *Grande grammatica italiana di consultazione*, vol. III. Bologna, Il Mulino, 225-257.
- Bazzanella, C. (2006): Discourse Markers in Italian: towards a 'compositional' meaning. In Fischer, K. (ed.), *Approaches to discourse particles*. Amsterdam, Elsevier, 449-464.
- Ghezzi, C., Molinelli, P. (2015): Segnali allocutivi di richiamo: percorsi pragmatici e sviluppi diacronici tra latino e italiano. *Cuadernos de Filología italiana* 22, 21-47.
- Güldemann, T., von Roncador, M. (eds.) (2002): *Reported discourse: a meeting ground for different linguistic domains*. Amsterdam, John Benjamins
- Maldonado, C. (1991): *Discurso directo y discurso indirecto*. Madrid, Taurus.
- Mandelli, M. (2010a): Discorso indiretto. In *Enciclopedia dell'italiano*. Roma, Istituto della Enciclopedia italiana Giovanni Treccani
- Mandelli, M. (2010b): Discorso diretto. In *Enciclopedia dell'italiano*. Roma, Istituto della Enciclopedia italiana Giovanni Treccani
- Molinelli, P. (2014): Orientarsi nel discorso. Segnali discorsivi e segnali pragmatici in italiano. In Pirvu, E. (ed.), *Discorso e cultura nella lingua e nella letteratura italiana*. Firenze, Franco Cesati, 195-208.
- Reyes, G. (2002): *Los procedimientos de cita: estilo directo y estilo indirecto*. Madrid, Arco Libros.

## 'En realidad' as a crossing gate: multifunctionality and polysemy of a DM

Amalia Canes Nápoles and Nicole Delbecque

### Abstract

One of the main problems in the description of DMs appears to be their multifunctionality and polysemy. Yet, the semantic and functional analysis of DMs is characterized by a lack of attention for the multidimensional structures of meanings and for the linguistic conditions that favor superposing various interpretations. Over time, these particles rarely evolve toward a single semantic-pragmatic function and meaning but rather tend to increase their ability to operate at various pragmatic levels. In this paper on the Spanish expression *en realidad* ([ER] in reality, in fact, actually) we propose a comprehensive bottom-up approach in terms of a conceptual network of related meanings. We seek to understand the ambiguity of ER, first, by accounting for its meanings and functions, and second, by estimating the probabilistic conditions underlying usage decisions. The corpus used in this study comprises 503 ER occurrences in Cuban Spanish variety from the Corpus of the Real Academia Española's databank CREA.

This work introduces a statistical method, called Decision Tree (DT), for disambiguating both lexical and pragmatic uses. While the properties for the DM's syntactic identification have received more attention (e.g. Siegel & McKeown [1994] & Eddington [2010]), the disambiguation of functions and meanings through statistical analysis has been rather overlooked. The method applied in this paper is intended to provide insights as to the type of morpho-syntactic, pragmatic, and semantic information necessary to further improve DM sense disambiguation. It points to correlations among the variables that most significantly explain the DM uses.

Our working hypothesis is that the pragmatic use of ER ranges over six related functions, all of which involve a notion of reformulation. Although the relation between appearance (p) and reality (q) belongs to the semantic field of 'opposition', the manner in which this 'opposite meaning' between p and q is expressed by ER is not uniform. The different interpretations of ER are in essence due the imprecision or vagueness of the dichotomy expressed by this DM. Thus, it is unsatisfactory to classify it merely as an argumentative reinforcement operator (Martín Zorraquino & Portolés Lázaro 1999: 4140).

In this contribution we also question the 'reliability' or adequacy of individual grammatical cues for the pragmatic disambiguation of ER. The syncretism of different values in one expression inevitably reduces the transparency of its interpretation. To compensate for the ambiguity in the ER use, speakers make use of information about the text type, as well as of formal and semantic cues, such as the temporal deixis, the type of predicate and the type of syntactic constituent the DM introduces. In this context, DT offers a new methodological route that bridges grammar and discourse and produces a more holistic approach of the functional and semantic disambiguation of DMs.

### References

- DELBECQUE, N. (2012): "En el fondo: polifuncionalidad y polifonía de la localización interna", in *Cum corde et in nova grammatica: estudios ofrecidos a Guillermo Rojo*. Santiago de Compostela, 247-263.
- EDDINGTON, D. (2010): "A comparison of two tools for analysing linguistic data: logistic regression and decision trees", in *Italian Journal of Linguistics* XXII, II, 265-286.

- FISCHER, K. (2000): *From cognitive semantics to lexical pragmatics: the functional polysemy of discourse particles*. Berlin: Mouton de Gruyter.
- GARCÉS GÓMEZ, M. P. (2010): "Marcadores de corrección y rectificación en los textos escritos", in *Revista de Investigación Lingüística XIII*, 87-105.
- GONZÁLEZ MANZANO, M. (2013): *Gramaticalización de los marcadores epistémicos en español*. Universitat de Barcelona.
- GRIES, S. TH. (2012): "Behavioral Profiles: a fine-grained and quantitative approach in corpus-based lexical semantics", in Gonia Jarema, G./ Westbury, C. (eds.), *Methodological and analytic frontiers in lexical research*. Amsterdam & Philadelphia: J. Benjamins, 57-80.
- HUMMEL, M. (2012): *Polifuncionalidad, polisemia y estrategias retóricas en los signos discursivos con base atributiva – entre oralidad y escritura. Acerca del uso de bueno, claro, total, realmente, etc.* Berlin: de Gruyter.
- LÓPEZ SERENA, A./BORREGUERO ZULOAGA, M. (2010): "Los marcadores del discurso y la variación lengua hablada vs. lengua escrita", in Loureda Lamas, Ó./Acín Villa, E. (eds.), *La investigación sobre marcadores del discurso del español, hoy*. Madrid: Arco/Libros, 415-495.
- MARTÍN ZORRAQUINO, M. A & PORTOLÉS LÁZARO, J. 1999. "Los marcadores del discurso". In Demonte Barreto, V & Bosque, I. (eds.). *Gramática descriptiva de la lengua española Vol. 3: 4051-4214*.
- MORTIER, L./DEGAND, L. (2009): "Adversative discourse markers in contrast: The need for a combined corpus approach", in *International Journal of Corpus Linguistics*, 14, (3), 301- 329.
- SCHWENTER, S. (2000): "Viewpoints and polysemy: Linking adversative and causal meanings of discourse markers", in Couper-Kuhlen, E./ Kortmann, B. (eds.), *Cause–Condition–Concession–Contrast. Cognitive and Discourse Perspectives*. Berlin: Mouton de Gruyter, 257–282
- SIEGEL & MCKEOWN. 1994. "Emergent Linguistic Rules from Inducing Decision Trees: Disambiguating Discourse Clue Words". In *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94)*. Seattle, Washington, July 1994. AAAI Press/MIT Press, Cambridge, MA: 820-826.
- TRAVIS, C. E. 2005. *Discourse markers in Colombian Spanish: a study in polysemy*. Berlin: Mouton de Gruyter.

## Les marqueurs discursifs de la relation d'ajout. Une étude contrastive français-italien

Anna-Maria De Cesare

### Abstract

Cette communication se penche sur les marqueurs de la relation d'ajout (d'adjonction ou d'addition), qui peut être considérée comme une relation discursive basique. L'opération qui consiste à ajouter une information à une ou plusieurs autres informations déjà données auparavant, et ainsi à enrichir et modifier le bagage de connaissances du lecteur (si l'on considère uniquement la réception des textes écrits), se réalise en effet par défaut et n'a généralement pas besoin d'être explicitée par une marque linguistique (cf. (1)). Les langues européennes disposent toutefois d'un paradigme d'expressions relativement différenciées (étudiées notamment par Nøjgaard 1992, 1995, Perrin-Naffakh 1996, Pierrard 2012 pour le français; De Cesare 2004, 2008 pour l'italien ; et Gil 1995, De Cesare & Borreguero 2014 dans une perspective romane), qui vont de la simple conjonction de coordination (2), aux connecteurs d'ajout, comme en (3) (auxquels il faut ajouter au moins les connecteurs issus d'adverbes temporels, comme 'puis' et 'enfin', sur lesquels on verra, parmi de nombreux autres, les travaux de Mosegaard Hansen 1995 et de Saussure & Morency 2013 pour le français et Mingioni 2014 pour l'italien), en passant par les adverbes paradigmatiques additifs, comme 'également' et 'aussi' (4).

(1) Eva est fatiguée. Elle a froid.

(2) Eva est fatiguée. Et elle a froid.

(3) Eva est fatiguée. En plus / De surcroît / En outre / Qui plus est, elle a froid.

(4) Eva est fatiguée. Elle a également / aussi froid.

Dans cette communication, dédiée en particulier à la classe des connecteurs d'ajout (et plus précisément aux 'sériels successifs mixtes' ou 'sériels binaires terminaux' dans la classification fonctionnelle très raffinée de Nøjgaard 1992 §§ 157-163), on répondra aux questions suivantes : (i) comment est-ce que se définit la relation logico-argumentative ('de dicto' ou rhétorique) d'ajout et quels sont ses sous-types (cf. par exemple l'ajout sans planification chez Ricci 2007 ; l'ajout totalisant et réalisant chez Sauerwein Spinola 2013) ? (ii) quelles sont les inventaires de formes qu'il faut poser pour le français (sur la base de Nøjgaard 1995, § 406) et pour l'italien ; (iii) quelles sont les différences les plus importantes entre les systèmes développés par le français et par l'italien ? (iv) à partir de quels critères (syntactiques, sémantiques, informationnels) peut-on dire qu'un adverbe paradigmatique additif (comme les fr. 'également', 'aussi' et l'it. 'anche') fonctionne comme connecteur d'ajout ?

On répondra aux questions (i) à (iv) d'une part en s'appuyant sur la bibliographie existante en matière (relative au français et à l'italien) et d'autre part sur une recherche empirique des connecteurs d'ajout les plus fréquents dans les textes rédigés en français et en italien (tant comme langues sources que comme langues cibles) du corpus européen EUROPARL (cf. Koehn 2005).

## Références

- De Cesare, Anna-Maria 2004, "L'avverbio 'anche' e il rilievo informativo del testo", in Ferrari, Angela (ed.), *La lingua nel testo, il testo nella lingua*, Torino, Istituto dell'Atlante Linguistico Italiano, 191-218.
- De Cesare, Anna-Maria 2008, "Gli averbi paradigmaticizzanti". In Ferrari, Angela et al., *L'interfaccia lingua-testo. Natura e funzioni dell'articolazione informativa dell'enunciato*. Alessandria: Edizioni dell'Orso, § III.8.1, 340-361.
- De Cesare, Anna-Maria & Margarita Borreguero Zuloaga 2014. "The contribution of the Basel Model to the description of polyfunctional discourse markers. The case of It. 'anche', Fr. 'aussi' and Sp. 'también'", In Salvador Pons Bordería (éd.), *Discourse Segmentation in Romance Languages*. Amsterdam-Philadelphia: Benjamins, 55-94.
- de Saussure, Louis & Patrick Morency 2013, « Adverbes temporels et sériels », in De Mulder W. & al. (éds) : *Temps, mode, modalité: du sens aux emplois et inversement*, Amsterdam : Rodopi, 337-353.
- Gil, Alberto 1995, *Textadverbiale in den romanischen Sprachen: eine integrale Studie zu Konnektoren und Modalisatoren im Spanischen, Französischen und Italienischen*. Frankfurt: Lang.
- Koehn, Philipp 2005, "Europarl: A Parallel Corpus for Statistical Machine Translation", In *Machine Translation Summit X*, 79-86.
- Mingioni, Ilaria 2014, *Aspetti formali e semantico-pragmatici dei connettivi di chiusura del discorso in prospettiva storica e sincronica*, Thèse de doctorat, Università di Roma Tre.
- Mosegaard Hansen, Maj-Britt 1995, « 'Puis' in Spoken French: from Time Adjunct to Additive Conjunct? », *Journal of French Language Studies* 5/1, 31-56
- Nøjgaard, Morten 1992, 1995. *Les adverbes français. Essai de description fonctionnelle*. Historisk-filosfiske Meddelelser 66/1, 3. Copenhagen : Munksgaard.
- Perrin-Naffakh, Anne-Marie 1996, "Aussi' adjectif: de la syntaxe à la sémantique". *Le français moderne* 64/2: 136-154.
- Pierrard, Michel 2012, « 'Aussi bien que' et le marquage de l'adjonction », *Travaux de linguistique* 65/2: 85-103.
- Ricci, Claudia 2007, « L'ajout non planifié ou la reconstruction a posteriori d'une relation de discours ». In Corinne Rossari (éd.), *Les moyens détournés d'assurer son dire*, Paris: PUPS, 57-76.
- Sauerwein Spinola, Sibylle 2013, « Le marqueur 'de surcroît' : l'ajout totalisant et l'ajout réalisant en français », Pirazzini, Daniela / Schiemann, Anika (éds.), *Dialogizität in der Argumentation. Eine multidisziplinäre Betrachtung*. Bern, Lang, 251-262.



## Causal Categories in Turkish connectives; First Results From a Corpus Study and a Judgement Task

Derya Cokal, Deniz Zeyrek and Ted Sanders

### Abstract

Connectives are the prototypical linguistic markers of coherence relations in discourse. But connectives cannot express just any coherence relation. In fact, connectives can be grouped according to the type of relation expressed, such as additive, temporal, causal or contrastive relations (3; 4). In addition, recent studies show various languages sometimes have subtle distinctions within the same class of connectives. In Dutch, for instance, the causal connectives *omdat* and *want* have been shown to express the difference between typical consequence-cause and claimargument relations (5; 6; 7). Such relational differences in causality are explained in

different (but related) terms, such as 'objective vs. subjective relation' (7) or 'epistemic vs. content domains' (8). In example (i) below, *because* connects units (a) and (b) in a claim-argument relation, which is inferred by readers (i.e., [b] provides the speaker's subjective reason for [a]). The paraphrase of (i) would be: "The reason I thought the neighbors are not home is their lights are turned off." In example (ii) the same connector *because* is used in a consequence-cause relation. Consequently, the paraphrase of (ii) is "The reason the temperature rose is the shining of the sun."

While Ex. (i) conveys an epistemic, or subjective relation, Ex. (ii) conveys real-world causality (i.e., an objective relation).

(i) (a) The neighbours are not at home because (b) their lights are off.

(ii) (a) The temperature rose because (b) the sun was shining.

Using categorization principles, including 'domains' and 'subjectivity', causal connectives have been studied in Dutch (6, 7), French (9), German (1), and English (5). This has led to conclusions regarding similarities (e.g., distinctions like these are clearly relevant in Dutch, German and French) and differences (In English *because* can express any causal relation.). While these corpus studies' results are interesting, they have at least one serious limitation: they are all concerned with a limited set of closely related European languages. Only recently have typologically different languages like Mandarin been studied seriously. (2) demonstrated different causal connectives signal different degrees of subjectivity. In this paper, we outline our study that involves another typologically different language: Turkish. Coming from an Altaic language family, and with a rich list of causality connectives, borrowed from Persian and Arabic, our study promises to enrich the discussion of types of causality. Currently there is little knowledge about the distribution of Turkish causal connectives (i.e., *çünkü*, *-diği için*, *da* and *zira* roughly all mean "because"). To our knowledge, they have not been described in terms-relation categorizations like domains or subjectivity. To fill this gap, we present the results of our first corpus analysis and acceptability judgment task to categorize various Turkish causal connectives. The main questions of this paper are: (1) Can

Turkish causality markers used in written discourse be categorized in terms of domains or subjectivity? and (2) Does Turkish have specialized connectives, which are really specific for objective/epistemic or subjective/content domains? Our initial logistic regression analysis from an acceptability judgment task shows that while

“çünkü” and “zira” are mostly preferred when expressing a subjective relation, “-diği için” and “da” can be used to express both subjective and objective relations.

The current paper will contribute to the ongoing work on Discourse Relational Devices in three key ways: (a) Carrying out a functional categoration of Turkish causality connectives; (b) Highlighting difficulties experienced while annotating discourse relations marked by Turkish causality markers; and (c) Discussing whether in Turkish there are strong conceptual contrasts between ‘objective’ and ‘subjective’ relations or between ‘epistemic’ and ‘content’ domains that have been found in European languages.

## References

- 1- Keller, R. (1995). The epistemic weil. In Stein, D., Wright, S. (Eds.), *Subjectivity and Subjectification* (pp. 16–30). Cambridge, CUP.
- 2- Li, F., Evers-Vermeul, J. & Sanders, T.J.M. (2013). Subjectivity and result marking in Mandarin - A corpus-based investigation. *Chinese Language and Discourse*, 4 (1), 74 –119.
- 3- Mann, W. C., Thompson, S. A. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8, 244-277.
- 4- Sanders, T.J.M. & Knott, A. (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*, 30, 135-175.
- 5- Sanders, J., Sanders, T. J.M. & Sweetser, E. (2012). Responsible subjects and discourse causality. How mental spaces and perspective help identifying subjectivity in Dutch backward causal connectives. *Journal of Pragmatics*, 44(2), 169-190.
- 6- Sanders, T.J.M. & Spooren, W. (2015). Causality and subjectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text. *Linguistics*, 53 (1), 53-92.
- 7- Sanders, T.J.M. & Spooren, W. (2013). Exceptions to rules: a qualitative analysis of backward causal connectives in Dutch naturalistic discourse. *Text & Talk*, 33(3). 399-420.
- 8- Sweetser, E. S. (1990). *From etymology to pragmatics*. Cambridge: CUP.
- 9- Zufferey, S. (2012). ‘car, parce que,puisque’ revisited: two empirical studies on French causal connectives. *Journal of Pragmatics*, 44, 138-153.

## What does "Connective" Mean Cross-Linguistically?

Margot Colinet, Tatjana Scheffler, Manfred Stede and Laurence Danlos

### Abstract

Discourse connectives are a class of lexical items that can signal discourse relations. Pre-theoretically, the notion 'discourse connective' is semantic and pragmatic, and in many individual studies its definition is not debated, since only a subset is considered anyway. However, large-scale annotation efforts of local discourse structure have led researchers in different languages to explicitly define the class under consideration. In this paper, we present a detailed analysis of the state-of-the-art definitions of 'discourse connectives' in several linguistic traditions, making underlying assumptions explicit. We show that a purely semantic/pragmatic definition of the class of connectives is not possible, and morphosyntactic constraints are necessitated in each language. Due to linguistic idiosyncrasies, these constraints lead to sometimes substantial differences in the discourse connective lexica. In our paper, we present the specific differences in the definitions and lexical classes for connectives in German, French, and English.

More specifically, after extracting (non-)correspondences between two major resources of discourse connectives, the French LexConn (Roze et al., 2012) and the German DiMLex (Stede, 2002), we have been able to detect a number of mismatches. Here, we present only two for space reasons:

One major difference is that LexConn contains many more adverbial PPs than DiMLex. The number such adverbial PPs depends on idiosyncrasies: certain PPs form a lexical unit in one language but not in the other (à seule fin de/mit dem einzigen Zweck), other expressions seem lexicalized cross-linguistically (par exemple/zum Beispiel/for example). Here, it might be possible to determine cross-linguistically valid criteria to decide whether a multiword expression is a lexicalized unit or not (modifiability, compositionality, etc.), though the specific connectives that match these criteria would still differ between languages.

Another major difference relies on the formal definition of which morphosyntactic categories are allowed as arguments of a discourse relation. LexConn contains only prepositions taking an infinitival clause as complement, because a discourse connective in LexConn should minimally link two clausal elements like *pour* in (1) but unlike *pour* or *en raison de* in (2) :

(1) Il a payé une amende pour avoir grillé un feu rouge.

'He paid a fine for having run a red light.'

(2) Il a payé une amende pour/en raison d'un feu rouge grillé.

'He paid a fine for a run red light.'

DiMLex contains both kinds of prepositions, i.e., not only those that can take an infinitival clause as their complement like *ohne* in (3) but also those that can take only a noun (nominalization or eventive noun) like *trotz* in (4). As a result, DiMLex contains many more prepositions than LexConn.

(3) Er kann nichts fertigstellen, ohne Hilfe zu bekommen.

'He can't finish anything without getting help.'

(4) Wir gingen spazieren trotz der Verschlechterung des Wetters.

'We went for a walk despite the worsening of the weather'

These observations, among others, should lead to an overall enhanced understanding of what should go into a connective lexicon, and in what form it should be represented.

#### References

- Berger, D., Reitter, D., and Stede, M. (2002). Xml/xsl in the dictionary: the case of discourse markers. In Proc. of the 2nd Workshop on NLP and XML (NLPXML-2002), Taipei.
- Colinet, M., Danlos, L., and Steinlin, J. (2015). FDTB1: The first step in annotating a French corpus in discourse. In Poster presentation at the COST TextLink Action meeting in Louvain-la-Neuve, Belgium.
- Roze, C., Danlos, L., and Muller, P. (2012). LexConn: a French lexicon of discourse connectives. *Revue Discours* [Online] URL : <http://discours.revues.org/8645>.
- Stede, M. (2002). DiMLex: A lexical approach to discourse markers. *Exploring the Lexicon: Theory and Computation*.
- Stede, M. and Umbach, C. (1998). DiMLex: A lexicon of discourse markers for text generation and understanding. In Proceedings of the 17th ACL, pp 1238-1242.

## Position and context of structuring discourse markers: a corpus-based investigation of scope and cognitive load in French and English

Ludivine Crible

### Abstract

Discourse markers (DMs) are usually broadly defined as structuring and organizing devices that help produce and understand meaning-in-context (e.g. Schiffrin 1987, Fischer 2000). The panel of functions that they can perform ranges from local cohesion (e.g. a causal relation between (sub)clauses) to “global coherence” (Lenk 1998, e.g. topic shift), hence a great variability in the size and type of units they apply to. The notion of scope of DMs has received some attention in grammaticalization studies (Traugott 1995) and more recently in studies on position and segmentation (Pons Bordería & Estelles 2009, Estelles & Pons Bordería 2014). However, scope remains a complex feature to annotate directly in spoken corpora, given the interplay of syntactic and semantic-pragmatic interpretation it involves.

I claim that a scale of DM scope can be obtained by merging the information provided by a multi-layered annotation of formal and functional features of DMs, here applied to a comparable corpus of spoken French and English. These annotations include (but are not limited to): function(s) of the DM; a three-fold positioning system (clause, dependency structure, turn-of-speech); co-occurrence patterns; disfluencies in the immediate co-text of the DM (silent/filled pauses, repetitions etc.). My proposal is that corpus-based annotations – coupled with metadata on the contextual settings (e.g. degree of preparation, number of speakers) – can be interpreted in terms of scopes and their relative cognitive load.

I will focus on “sequential” DMs (Gonzalez 2005, [author]), which form a functional subgroup attending to topic management and dialogue structure, viz. turn opening, turn closing, topic shifting, topic resuming, and listing (see [author] for operational definitions of these values), to examine the different ways in which they relate to co-text and context. According to Roberts & Krisner (2000), sequential DMs should theoretically be more complex to produce, given that they work on both the mental “linearization” (Levelt 1989) of the global order of segments (e.g. macro-structure of topics) and on the actual “linearity” of the articulated output (e.g. local transition between topics or turns). I will therefore look for local and contextual evidence for heavy cognitive load, viz. presence of disfluencies, number and function(s) of co-occurring DMs, formality and complexity of the situation (e.g. transmission of information, professional setting). My hypothesis is that sequential DMs should co-occur with more disfluencies (including other DMs) than DMs from other (ideational, interpersonal) domains. I also expect that those markers which are associated with heavy cognitive load take scope over larger units (several propositions, a whole turn).

This corpus-based in-depth investigation of DMs in a variety of contexts should uncover patterns and clusters of features that provide indirect evidence for the scope of these structuring elements of discourse. This will in turn contribute to better define the notion of speech linearity, and bring us closer to a cognitive-pragmatic model of discourse structure, following the lines of Van Dijk (e.g. 1989), Degand & Simon (2009) or the Val.Es.Co group (e.g. Pons Bordería 2006).

### References

- Degand, L. & Simon, A.-C. 2009. “Mapping prosody and syntax as discourse strategies: How Basic Discourse Units vary across genres”. In A. Wichmann, D. Barth-Weingarten, N. Dehé (eds), *Where prosody meets pragmatics: research at the interface*, Emerald: Bingley: 79-105.
- Estelles, M. & Pons Bordería, S. 2014. “Absolute initial position”. In S. Pons Bordería (ed.), *Discourse Segmentation in Romance Languages*, Amsterdam, John Benjamins: 121-155.

- Fischer, K. 2000. *From cognitive semantics to lexical pragmatics*. Berlin: Mouton de Gruyter.
- Gonzalez, M. 2005. "Pragmatic markers and discourse coherence relations in English and Catalan oral narrative". *Discourse Studies* 77/1: 53-86.
- Lenk, U. 1998. "Discourse markers and global coherence in conversation". *Journal of Pragmatics* 30: 245-257.
- Levelt, W. 1989. *Speaking. From intention to articulation*. Cambridge: MIT Press.
- Pons Bordería, S. 2006. "A functional approach to the study of discourse markers". In K. Fischer (ed.), *Approaches to discourse particles*, Amsterdam, Elsevier: 77-100.
- Pons Bordería, S. & Estelles, M. "Expressing digression linguistically: Do digressive markers exist?". *Journal of Pragmatics* 41: 921-936.
- Roberts, B. & Kirsner, K. 2000. "Temporal cycles in speech production". *Language and cognitive processes* 15/2, 129-157.
- Schiffrin, D. 1987. *Discourse markers*. Cambridge: CUP.
- Traugott, E. 1995. "The role of the development of discourse markers in a theory of grammaticalization". Paper presented at the International Conference on Historical Linguistics (ICHL) XII, Manchester, UK.
- Van Dijk, T. 1989. "Cognitive context models and discourse". In M. Stamenow (ed.), *Language structure, discourse and the access to consciousness*, Amsterdam: John Benjamins: 189-226.

## Processing of expected/unexpected focus in pragmatic scales in Spanish: experimental notes about the Spanish focus particle *incluso*

Adriana Cruz Rubio

### Abstract

Our subject of study is the Spanish focus particle *incluso* ("even"), which imposes a certain process of information retrieval in utterances. Focus particles, as elements with a mainly procedural meaning in discourse, act upon information structure to focalize the hearer's attention on certain constituents and guide him towards the expected inferences (Blakemore 1992; Portolés 2001; Murillo 2010; Escandell/Leonetti 2011).

The present research aims to complement theoretical works on focus particle and information structure with an experimental approach: a series of experiments, designed with the eye tracking methodology (Rayner 1998) which allows to keep track cognitive processes involved in reading activity ("eye-mind hypothesis") (Richardson et al. 2007). Complemented with a comprehension test (Just et al. 1982), it leads to relevant data about the effective information processing that are obtained.

Experimental results presented in this communication, obtained from quantitative and qualitative analysis, concerned the variable of interest: expected/unexpected focus. Different pairs of utterances, which are evoked by the instruction of the focus particle (Portolés 2007), were inserted in two types of context according to the pragmatic features of their constituents.

For example:

(1) Letizia y Paola conocen Sevilla, Granada e *incluso* Málaga

(2) Letizia y Paola conocen Málaga, Granada e *incluso* Sevilla

The focus particle *incluso* codifies a scalar culminative instruction and evokes a mental representation introducing an element Málaga (in 1) and Seville (in 2) as more informative than the other elements of the utterance. Since these utterances have less world knowledge to process, because the world knowledge does not present one city more important than another, they should show the same processing effort.

On the contrary, in utterances like (3) and (4) the common cognitive environment is predominant over the procedural meaning of the focus particle (Pons 2008):

(3) Ana y Marta hablan inglés, francés e *incluso* chino

(4) Ana y Marta hablan chino, francés e *incluso* inglés

The world knowledge allows to create a mental representation, where Chinese (in 3), the most informative element, could be expected as more difficult than the other mentioned languages (English, French). Higher processing efforts will be implied in utterance (4) where the utterance is pragmatically strange: it is less expected that speaking Chinese will be easier than speaking French or English; the focus particle *incluso* focalizes on the strongest element of a word-chain. Therefore, the utterance (3) is contradictory, not so (4), in which elements are classified attending their difficult learning degree.

Our main research questions can be formulated as:

- Does all utterances like in 1 vs 2 (or in 3 vs 4) have the same processing effort?
- How does the presence of focus particles influence the pattern of information distribution in utterances, if the focus is expected or unexpected?
- Do focus particles enhance readers' effective comprehension of utterances?

## References

- BLAKEMORE, D. (1992): *Understanding utterances*. Oxford: Blackwell.
- ESCANDELL-VIDAL, V./LEONETTI, M./AHERN, A. (2011): *Procedural meaning: problems and perspectives*. Bingley: Emerald.
- JUST, M. A./CARPENTER, P. A. (1980): "A theory of reading: From eye fixations to comprehension", *Psychological Review* LXXXVII/4, 329-54.
- MONTOLÍO, E. (1998): "La teoría de la relevancia y el estudio de los marcadores discursivos", in M. A. MARTÍN ZORRAQUINO/E. MONTOLÍO DURÁN (ed.): *Los marcadores del discurso. Teoría y análisis*. Madrid: Arco/Libros, 93-119.
- MURILLO, S. (2010): "Los marcadores y su semántica", in Ó. LOUREDA/E. ACÍN (ed.): *La investigación sobre marcadores del discurso del español, hoy*. Madrid: Arco/Libros, 241-280.
- PONS BORDERÍA, S. (2004): *Conceptos y aplicaciones de la Teoría de la Relevancia*. Madrid: Arco/Libros.
- PORTOLÉS, J. (2001 [1998]): *Marcadores del discurso*. Barcelona: Ariel.
- PORTOLÉS, J. (2007): "Escalas informativas aditivas. Pruebas del español", *Spanish in Context* IV/2, 35-157.
- RAYNER, K. (1998): "Eye movements in reading and information processing: 20 years of research", *Psychological Bulletin* 124/3, 372-422.
- RICHARDSON, D. R. et al. (2007): "Eye movements in language and cognition: A brief introduction." in M. GONZÁLEZ MÁRQUEZ/I. MITTEBERG/S. COULSON/M. J. SPIVEY (ed.): *Methods in Cognitive Linguistics*. Amsterdam/Philadelphia: John Benjamins, 323-344.
- SPERBER, D./WILSON, D. (1995): *Relevance*. Oxford: Blackwell.



## Annotation of multi-word discourse markers in spoken Slovene

Kaja Dobrovoljc

### Abstract

With increasing empirical evidence that a considerable amount of spoken human communication is made of prefabricated lexical chunks, stored and retrieved as a whole (Pawley and Syder 1983, Sinclair 1991, Wray 2002; Brazil 1995, Biber et al. 1999, Erman and Warren 2000, Leech 2000), there is a growing need to move the multi-word discourse relational devices, usually referred to as second-level discourse markers (Siepmann 2005), secondary connectives (Rysová and Rysová 2014) or alternative lexicalizations (Prasad et al. 2010), from the periphery to the center of discourse structuring research. To explore the functional and formal features of multi-word discourse markers in speech, relevant both to their corpus identification and annotation, the proposed paper presents the creation, annotation and initial analysis of multi-word discourse markers in the first manually annotated treebank of spoken Slovene, currently a work in progress.

The corpus has been created by discourse-aware sampling of the reference corpus of spoken Slovene (Verdonik et al. 2013), a balanced and representative collection of transcripts of approximately 120 hours (1 million words) of spontaneous speech in different public and private, formal and informal, everyday situations. The sampled subcorpus currently amounts to 30,000 tokens with a similar text type and speaker demographics distribution, in order to assure its wide usability in linguistic and NLP research alike. In addition to the existing manual segmentation of utterances and turns, tokenization, orthographic and standardized transcription, four new layers of manually verified linguistic annotation have been added: lemmas, morphosyntactic tags, dependency relations and discourse marker annotations.

Within the current scope of our research, only multi-word discourse markers are considered. A multi-word discourse marker is defined as any syntactically optional lexically fixed contiguous string of two or more words with prevailing procedural meaning (Blakemore 2002), including both relational and non-relational discourse structuring devices. The initial functional taxonomy used within the annotation scheme is based on previous research of discourse markers in Slovenian speech (Verdonik 2007, Verdonik 2008) and currently includes five top-level categories: ideational (e.g. *zaradi tega ker* 'due to', *to pomeni da* 'this means that'), interactional (e.g. *veš kaj* 'you know what', *ja ja ja* 'yes yes yes', *to je res* 'that's right'), meta-commenting (e.g. *pa ne vem* 'I don't know', *kaj je že* 'what's it called'), interpretative (e.g. *in tako naprej* 'and so on'), and, given the fuzzy boundaries between discourse and modal marking (Degand et al. 2013), also modal (e.g. *na neki način* 'in a way', *v bistvu* 'in fact'). Given the recently proposed ISO set of core discourse relations (Prasad and Bunt 2015) and other language- or mode-independent functional taxonomies (Zufferey and Degand 2014, Crible and Zufferey 2015), the annotation scheme might be revised and further sub-level categories added.

After describing the creation of the corpus, the annotation process and some of the key issues related to multi-word discourse marker identification and delimitation, the third part of the paper presents the first analysis of the annotated data, namely: the number of multi-word discourse marker types and tokens, the scope of their functional polysemy, their formal characteristics (utterance position, syntactic structure, syntactic function) and their distribution across the different text types.

## Cohesive markers of contrast in English and French: what, where and why?

Maité Dupont

Languages differ in the types of devices that they use to signal discourse relations, as well as in the extent to which they need to signal these relations explicitly by means of cohesive markers. As regards the English-French language pair, the dominant position is that French tends to be more explicitly cohesive than English (see e.g. Vinay & Darbelnet 1958, Hervey & Higgins 2002). In the absence of solid empirical evidence, however, this claim remains largely hypothetical. The first aim of my research project is to test this assertion on an empirical basis, focusing on the meaning relation of contrast. For that purpose I use the powerful methods of corpus linguistics combined with the theoretical framework of Systemic Functional Grammar. On the basis of a comprehensive list of French and English markers of contrast analysed in large comparable and translation corpora, I aim to determine which language uses the larger number of markers of contrast overall and identify the preferences of each language for specific subtypes of markers from among the range of syntactic devices available to them (i.e. adverbial connectives, conjunctions of coordination, conjunctions of subordination, and so-called 'Alternative Lexicalizations', see Prasad et al. 2010). Another major objective of the study is to compare the placement patterns of adverbial connectives expressing contrast (e.g. *cependant*, *however*) in English and French, not only in terms of the possible positions, but also of the preferred positions in each language. A preliminary corpus-based study (Dupont 2015) has demonstrated that, in addition to systemic differences between English and French, the positional patterns of adverbial connectives were also influenced by discourse and lexical factors. The corpus analysis sets out to assess the relative weight of each of these three factors – systemic differences, discourse and lexis – on the placement patterns of adverbial markers of contrast. Finally, the study adopts a variationist approach throughout with a view to assessing the impact of register on the use of cohesive markers in each language. Two distinct registers are compared: quality newspaper editorials and academic articles in the Humanities.

### References

- Dupont M. 2015. Word order in English and French. The position of English and French adverbial connectors of contrast. *English Text Construction* 8(1), 88-124.
- Hervey S. & Higgins I. 1992. *Thinking translation*. London: Routledge.
- Prasad, Rashmi, Aravind Joshi and Bonnie Webber. 2010. Realization of discourse relations by other means: Alternative Lexicalizations. *Coling2010: Poster Volume*. Beijing, August 2010, 1023-1031.
- Vinay J.-P. & Darbelnet J. 1995 [1958]. *Comparative Stylistics of French and English. A Methodology for Translation*. Amsterdam and Philadelphia: John Benjamins.

## El marcador discursivo coloquial bueno: funciones pragmáticas en conversación coloquial y descripción en manuales de E/LE de nivel B2.

Laura Eliodoro-Furió

### Abstract

La presente comunicación parte del contraste entre los datos proporcionados por corpus lingüísticos de hablantes nativos de español con los contenidos gramaticales en los manuales para aprendices extranjeros de español como lengua extranjera (E/LE). En concreto, nos centramos en el uso por parte de dichos hablantes del conector bueno y del tratamiento y la información lingüística, gramatical o pragmática que de él se refleja en manuales de E/LE en el nivel B2 (Consejo de Europa 2001, Instituto Cervantes 2006). A partir del uso de dicho conector en algunas conversaciones del Corpus Valencia Español Coloquial (Val.Es.Co) (Cabedo y Pons 2013) y de las informaciones que nos proporcionan algunos trabajos relacionados con el tema (e.g. Pons Bordería 2003, 2006, Cuenca 2007, 2008, Briz et alii (2008)), analizamos la descripción de los valores que aparecen en los manuales de E/LE. Nuestra hipótesis es que existe una cierta distancia entre los datos de corpus coloquiales reales y la información proporcionada al alumnado y que, por tanto, se desaprovecha el potencial que los corpus lingüísticos pueden desarrollar en ese sentido. Por todo ello, consideramos que se hace necesaria una reflexión sobre el tratamiento de los conectores en los manuales de E/LE y sobre qué tipo de información procedente de muestras reales de habla podría ser incluida o formalizada en una caracterización de los conectores con finalidades aplicadas en el marco de la enseñanza de la pragmática en la clase de E/LE (Pons 2005).

### Referencias

- BRIZ, A., PONS, S. Y J. PORTOLÉS (coords.) (2008). Diccionario de Partículas Discursivas del Español. Disponible en línea en [www.dpde.es](http://www.dpde.es).
- CABEDO, A.; PONS, S. (eds.) (2013). Corpus Val.Es.Co 2.0. Disponible en <http://www.valesco.es>.
- CONSEJO DE EUROPA (2001). Marco común europeo referencia para las lenguas: aprendizaje, enseñanza, evaluación. Madrid: MEC-D-Anaya. Disponible en: [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/marco/](http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/)
- CUENCA, M. J. (2007). La pragmática en la gramática. Barcelona: Institut d'Estudis catalans, pág. 1-42.
- CUENCA, M. J. (2008). "Pragmatic markers in contrast: the case of well." *Journal of Pragmatics* 40 (8): 1373-1391.
- INSTITUTO CERVANTES (2006). Plan curricular del Instituto Cervantes. Madrid: Instituto Cervantes-Biblioteca nueva.
- PONS BORDERÍA, S. (2003). "From agreement to stressing and hedging: Spanish bueno and claro." In *Partikeln und Höflichkeit*, ed. by Gudrun Held, 219-236. Frankfurt: Peter Lang
- PONS BORDERÍA, S. (2005). La enseñanza de la pragmática en la clase de E/LE. Madrid: Arco/Libros.
- PONS BORDERÍA, S. (2006). «A functional approach to the study of discourse markers». Kerstin Fischer (Ed.). *Approaches to Discourse Particles*. Oxford: Elsevier, pág. 77-99.

## Applying visualisation tools to analyse turn-taking systems in informal conversations and political interviews

Guadalupe Espinosa-Guerri and Amparo Garcia-Ramon

### Abstract

The present study explores one of the main problems posed by Conversation Analysis, namely, the “sequence-organizational problem” (Schegloff, 2006:73): “How are successive turns of actions formed up to be coherent with the prior one (or some prior one) and constitute a course of action?”. The turn-taking structure is here taken to be a Discourse Relational Device (DRD) in its own right: it is the overall structure of turns-at-talk, emerging from specific relations between utterances, that participants orient to in the course of particular interactions. The main goal is to explore whether the application of a visualisation system (VS) that allows to represent specific connections between speakers’ utterances can show graphically how participants manage to co-construct textual coherence in spontaneous conversation using turn-taking systems.

A VS is a tool used to obtain a non-verbal graphic representation of the main characteristics of any given conversation. The VS employed here is called sawteeth (Briz, 2013) and it was designed in order to connect each utterance of a conversation with preceding and subsequent talk according to interactional criteria, i.e., linking each utterance with the piece(s) of talk that directly triggered that utterance and with the specific utterance(s) triggered by that piece of talk. This procedure enables the analyst to detect recurring speech patterns (or conversational figures) and other phenomena which might have remained undetected if we consider topic development only (e.g., the existence of utterances which are formally isolated, albeit thematically related).

The corpus consists of three (relatively) informal political interviews and three informal conversations. All of them are two-party interactions of approximately 30 minutes each, which belong to the Val.Es.Co. Corpus or were accessed through TV websites and transcribed by the researchers. Once the sawteeth system was applied to each of the four transcriptions by the two members of the research project individually, visualisation problems were discussed and solved in order to improve the VS. Then, the conversational phenomena signalled by the recurring figures were grouped and organised as a (provisional) typology of conversational structures. Further analysis is expected to show that it is possible to use those figures to find qualitative and quantitative differences between political interviews and ordinary conversation regarding turn-taking systems. Through the analysis of a larger corpus, we would be able to provide a graphic characterisation of a variety of dialogic genres which would help define those genres more precisely.

### References

- Briz Gómez, A. (2013). “El relato coloquial”. Paper presented at the XXVIII Congreso de hispanistas italianos, Pisa, Italy (27-30/11/2013).
- Heritage, J. and Greatbatch, D. (1991). On the Institutional Character of Institutional Talk: The Case of News Interviews. In Boden, D. and Zimmerman, D.H. (Eds.), *Talk and Social Structure. Studies in Ethnomethodology and Conversation Analysis*. Cambridge: Polity Press.
- Levinson, S.C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Schegloff, E.A. (2006). Interaction: The Infrastructure for Social Institutions, the Natural Ecological Niche for Language and the Arena in which Culture is Enacted. In Enfield, N.J. and Levinson, S.C. (Eds.), *Roots of Human Sociality. Culture, Cognition and Interaction* (pp. 70-96). London: Berg.

## The presence and absence of DRDs in Bible translations: the implicit and explicit marking of addition, contrast, and cause

Péter Furkó

### Abstract

The aim of the present paper is to take a discourse-pragmatic approach to the use of *kai*, *de*, *gar*, and *alla* in the Textus Receptus as well as their translation equivalents (or the motivation behind asyndeton, i.e. the lack of a translation equivalent) in the Vulgata and the most frequently used English, French and Hungarian Bible translations.

Karlson argues that in Hebrew and Greek texts these DRDs served as "quasi-punctuation marks" as well as "dependency indicators" between clauses or sentences, "binding them into a whole" (Karlson p1). While omitting such DRDs in the original text might result in "tearing the context apart" (ibid.), this is not necessarily true in the target language, e.g. in English, where sentence-initial *for*, *and*, *but*, etc. are, according to many, simply translation effects that inhibit discourse comprehension.

The study is based on the first 150 verses of John and the complete Galatians (also comprising a little under 150 verses) from the TR as the Language A corpus and the relevant verses in the KJV, HKB, Vulg., ASV, CEV, GNB, etc. as the Language B, C, D ... corpora, which enables us to compare the use of DRDs and asyndeton in two different discourse genres—argumentative and narrative discourse—as well as in a variety of (local) contexts.

For the extraction of the texts, e-Sword Version 9.9.1 has been used, a tool that provides instant verse-by-verse alignment of Bible extracts across the major languages that the Bible has been translated into, thus giving a plethora of examples for the study of discourse relations across languages and contexts. After extracting the relevant KWICs, ELAN has been used for the annotation of the formal and functional properties of individual DRDs as well as for two corpus queries (find overlapping labels, N-gram with annotations). The results of the study show a correlation between the (primary/discourse and secondary/interpersonal) functions of the various tokens of the four DRDs and the frequency with which a translation equivalent occurs in the target language corpora. The interpretation of the quantitative results will be complemented by the qualitative analysis of individual DRDs and the contextual motivation behind the use of asyndeton or a target language DRD.

The paper's primary purpose is to contribute to the theme "Translation studies on DRDs", while a secondary focus will also be on the "Relevant features in DRDs' description and classification".

### References

Karlson, Fred. "Ki/gar and waw/kai are often markers and not words to be translated." paper delivered at NW Regional ETS Meeting. <http://bible.org/article/kigar-and-wawkai-are-often-markers-and-not-words-be-translated> accessed 13th August 2014.

## When "same" means something different: the sign SAME as a discourse relational device in two sign languages

Sílvia Gabarró-López

### Abstract

This paper aims to contribute to the study of discourse relational devices (DRDs) in sign languages (SLs) from a cross-linguistic perspective, namely comparing the functions and the position of the sign glossed as SAME (SS) in French Belgian SL (LSFB) and Catalan SL (LSC). In both SLs, the SS is articulated with the indexes of both hands extended and getting in contact with an inward movement and has the same meaning, i.e. resemblance or similarity. In signed discourse, however, it is very productive but lacks from systematic in-depth research to account for its other uses.

In a first stage of this investigation in LSFB, the SS was involved in bracketing repetitions and introduced additional or more explicit information, the relation between topic and predicate, an approximation, a condition, a consequence, a reformulation and a parenthetical comment. As for LSC, research in progress points out some shared functions with LSFB such as reformulation, explicitness and approximation. Nevertheless, these results pose two problems: (i) they are not 100% comparable because different annotation criteria were used in each SL, and (ii) they are limited to 2 signers for each SL in an argumentative task.

This research will solve the previously mentioned shortcomings by using a common annotation protocol [1] designed to provide operational guidelines for the description of DRDs in oral corpora, which is the modality of our data. Our corpus will be made up of 12 signers, 6 from the LSFB referential corpus [2] and 6 from the LSC pilot corpus [3], balanced in terms of age (2 signers from each SL belonging to one of the following age groups: 18 – 29, 30 – 49 and 50 – 80) and genre (3 men and 3 women per SL). The tasks chosen are a narration of a past memory and an argumentation on "deaf issues". The videos will be annotated using ELAN and discourses will be segmented using the principles of [4] adapted to the specificities of SLs.

For each SL, the goals are (i) to study the distribution of the SS across the two genres, (ii) to analyze the functions of the SS when it behaves as a DRD, (iii) to investigate the position of the SS within signed discourse, and finally (iv) to compare the results for both SLs. Expected intra-language results include a higher frequency in the use of the SS in argumentation rather than in narration because reformulation structures are more likely to occur in the first, so this DRD would be genre-dependent. As for the inter-language differences, the SS seems to be more productive in LSFB. For instance, it is used to add information, while LSC counts on a different sign for this function. Therefore, more functions may be found in LSFB than in LSC and because of this, a higher number of tokens in LSFB. A middle position is predictable for most functions in both SLs except for the condition and adding information that are more likely to be in initial position, at least in LSFB.

### References

- [1] Crible, Ludivine. 2014. Identifying and describing discourse markers in spoken corpora, Annotation protocol v.8. Unpublished working draft.
- [2] Meurant, Laurence. forthc. Corpus LSFB. Corpus informatisé en libre accès de vidéo et d'annotation de langue des signes de Belgique francophone. Laboratoire de langue des signes de Belgique francophone (LSFB Lab), FRS-FNRS, Université de Namur.
- [3] Barberà, Gemma, Josep Quer and Santiago Frigola. In press. "Primers passos cap a la documentació de discurs signat. El projecte pilot de constitució del corpus de la llengua de signes catalana". *Treballs de Sociolingüística Catalana* 25.
- [4] Degand, Liesbeth and Anne Catherine Simon. 2009. "Minimal Discourse Units in Spoken French: On the Role of Syntactic and Prosodic Units in Discourse Segmentation". *Discours* 4. <http://discours.revues.org/5852>. DOI:10.4000/discours.5852

## The Tena Kichwa =mi as a marker of speaker authority

Karolina Grzech

### Abstract

Tena Kichwa (TK) (QI, Quechuan, Ecuador) is an agglutinative, exclusively suffixing language, spoken in the Ecuadorian Amazon. TK has two main grammatical categories: verbs and nominals, each associated with a distinct set of derivational and inflectional markers. TK also has a set of twelve word-final 'free enclitics', which attach to hosts from both grammatical categories.

To date, free enclitics in TK and other Quechuan varieties have mainly been discussed in relation to their evidential semantics. Most Quechuan varieties exhibit an evidential distinction between direct, indirect and reported source of information (cf. Willett 1988; Aikhenvald 2004), marked by a subset of free enclitics. Previous studies (e.g. Parker 1969; Weber 1989) acknowledge that the enclitics, including evidentials, are not syntactically obligatory, but do not discuss their (non-)occurrence in different discourse contexts. Although Quechuan evidential enclitics also function as focus markers (e.g. Muysken 1995), the TK data demonstrate that their focus-marking function alone is insufficient to account for the enclitics' distribution, which seems to rather be motivated by subjective choices of discourse participants.

The syntactic non-obligatoriness of the enclitics, and the fact that native speakers see the ability to use them appropriately as an important part of native speaker's pragmatic competence, lead me to analyse the enclitics as discourse markers. I follow Schiffrin (1988) in defining discourse markers as devices which increase discourse coherence, e.g. by conveying 'rhetorical effects, emphasis, [or] the attitude of the speaker' (Spencer & Luis 2012: 35).

In this presentation, I describe the discourse functions of =mi, the most widely-studied of Quechuan enclitics. In other varieties, =mi has been analysed as validational (Adelaar 1977), marking direct evidence/certainty (Weber 1986; Floyd 1997), or the best possible ground for making a speech act (Faller 2002). The TK data show those analyses are not well-suited for the TK =mi. Therefore, the presentation answers two research questions:

- a. What (if any) is the semantic contribution of =mi to an utterance?
- b. How does =mi contribute to structuring the TK discourse?

In response to the first question, I show that rather than encoding direct source of evidence, the TK =mi indicates the origo's claim to epistemic primacy. Secondly, I demonstrate that =mi organises discourse on the level of 'participant framework' and 'information state' (Schiffrin 1988: ch.10): by using =mi, the speakers choose to present themselves as authorities on the subject under discussion. This limits the interlocutors' possibilities of presenting themselves as such, consequently shaping the further development of the interaction. I also show how the above analysis is compatible with the focus-marking function of =mi.

The analysis presented here is based on the usage-based approach to linguistic data, and draws on elements of sequential discourse analysis. The data come from a corpus comprising a variety of TK oral genres, which includes 2h of elicited discourse (guided conversation, reactions to stimuli) and 12h of naturalistic discourse (interviews, narratives, political discourse etc.). The data was collected during my 10-month-long fieldwork in

Ecuador in 2013 and 2014. The analysis presented here forms part of my doctoral research into TK discourse markers.

## References

- Adelaar, Willem F. H. 1977. Tarma Quechua: grammar, texts, dictionary. The Peter de Ridder Press.
- Aikhenvald, Alexandra Y. 2004. Evidentiality. Oxford University Press.
- Faller, Martina T. 2002. Semantics and Pragmatics of Evidentials in Cuzco Quechua. Stanford University.
- Floyd, Rick. 1997. La estructura categorial de los evidenciales en el quechua wanka. (Serie Linguística Peruana). Lima: SIL International.
- Muysken, Pieter. 1995. Focus in Quechua. In Katalin É. Kiss (ed.), *Discourse Configurational Languages*, 375–393. New York: Oxford University Press.
- Parker, Gary John. 1969. Ayacucho Quechua grammar and dictionary. Mouton.
- Schiffrin, Deborah. 1988. *Discourse Markers*. Cambridge University Press.
- Spencer, Andrew & Ana R. Luis. 2012. *Clitics: An Introduction*. Cambridge ; New York: Cambridge University Press.
- Weber, David. 1986. Information perspective, profile, and patterns in Quechua. In Wallace Chafe & Johanna Nichols (eds.), *Evidentiality: the linguistic coding of epistemology*, 137–55. Ablex Publishing Corporation.
- Weber, David John. 1989. *A Grammar of Huallaga (Huanuco) Quechua*. University of California Press.
- Willett, Thomas. 1988. A Cross-Linguistic Survey of the Grammaticization of Evidentiality. *Studies in Language* 12 (1). 51–97.



## Segmenting discourse units: Incorporating interpretation into decision rules?

Jet Hoek, Jacqueline Evers-Vermeul and Ted Sanders

### Abstract

Discourse relations connect two or more segments. Segmentation is an important step in the process of annotating discourse relations, but often not one extensively discussed in annotation methods or manuals. Ideally, implementing segmentation rules results in text segments that correspond to the units of thought related to each other. However, in many often-used annotation systems this does not always seem to be the case. Most formalized segmentation rules (e.g. Carlson & Marcu, 2001; Mann & Thompson, 1988; Reese, Hunter, Asher, Denis, & Baldridge, 2007; Sanders & van Wijk, 1996) would, for instance, not allow segmenting the conditional relation in (1), either because too many elements in S1 have been elided or because the segment following if would break up a larger unit. Still, the segmentation indicated in (1) seems very plausible and exactly captures the two segments related by the connective if.

(1) (context: The virus harms cold-blooded animals.) It does not replicate at temperatures above 25° centigrade and [would,]S2a if [present in fish for human consumption,]S1 [be inactivated when ingested.]S2b (ep 00-03-01)

In this presentation we present fragments encountered during an annotation effort of (explicit) local discourse relations from the Europarl corpus (Koehn, 2005) that are problematic to segment under most segmentation guidelines. We focus on three specific problems: ellipsis, complement structures, and perspective markers. We propose segmentation options that result in segments that do justice to the interpretation of the discourse relation and use translations (from the Europarl Direct corpus, Cartoni, Zufferey, & Meyer, 2013) as additional support for our analysis. Finally, we explore ways to formulate rules that produce text segments that do justice to interpretation.

We conclude that segmentation is in part dependent on the propositional content of text fragments, and that completely separating segmentation and annotation (i.e. treating it as a two-step process) does not always yield text segments that correspond to the text units between which a conceptual relationship (potentially signaled by a connective) holds (see also Verhagen, 2001). Although relying partly on the content of a text fragment results in better text segmentation, this does in turn raise problems for (semi-) automatically segmenting texts. Identifying specific problems, such as the ones addressed here, and being more explicit in segmentation strategies used in the annotation of discourse relations are important steps toward solving these problems.

### References

- Carlson, L., & Marcu, D. (2001). Discourse Tagging Reference Manual. ISI technical report ISI-TR-545. Retrieved from <http://www.isi.edu/~marcu/discourse/tagging-ref-manual.pdf>.
- Cartoni, B., Zufferey, S. & Meyer, T. (2013). Using the Europarl corpus for linguistic research. *Belgian Journal of Linguistics* 27, 23-42.
- Koehn, P. 2005. Europarl: A parallel corpus for statistical machine translation. *Proceedings of the 10th Machine Translation Summit* (pp. 79-86). Phuket, Thailand.
- Mann, W.C., & Thompson, S.A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8, 243-281.
- Reese, B., Hunter, J., Asher, N., Denis, P., & Baldridge, J. (2007). Reference manual for the analysis of rhetorical structure. Unpublished manuscript. University of Texas at Austin, Austin, TX.
- Sanders, T.J.M., van Wijk, C. (1996). PISA - A procedure for analyzing the structure of explanatory texts. *Text* 16, 91-132.
- Verhagen, A. (2001). Subordination and discourse segmentation revisited, or: Why matrix clauses may be more dependent than complements. In: T.J.M. Sanders, J. Schilperoord, & W.P.M.S. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 337-357). Amsterdam/Philadelphia: John Benjamins.

## Discourse relations and strategic use of extra-clausal constituents in co-constructed argumentative discourse

Carolin Hofmockel, Anita Fetzer and Robert Maier

### Abstract

Discourse comes in with the presumption of being coherent, and discourse coherence is brought about in discourse by various means, for instance discourse connectives signalling the nature of connectedness between discourse units, coherence strands, such as topic continuity, referential continuity, temporal and aspectual continuity, lexical coherence, and discourse relations. We conceive discourse relations (DRs) as cognitive constructs with defining and particularized features (cf. Asher & Lascarides 2003, Dik 1997, Givón 1993), which are cued with extra-clausal (explicit), intraclausal (implicit), and mixed linguistic material. Since DRs are embedded in discourse, their linguistic realization exploits local and not-so-local context. Apart from that, the linguistic realization of DRs is constrained by discourse genre and the corresponding genre-specific constraints and requirements (Fetzer and Speyer, forthcoming). Against this background, DRs may be fully specified by indexical reference to all of their defining condition(s) and all of their particularized features, and they may be under- and overspecified.

This paper reports on overspecification in the linguistic realization of DRs in dyadically co-constructed written discourse produced in an experimental setting, in which participants were asked to “flesh out” a skeleton text. In particular, the focus lies on those DRs that are cued by linguistic material in the left periphery, connecting syntactic positioning with discourse processing, as is reflected in the DR Contrast in excerpt (1):

(1) [In the past], London was a dowdy place of tea-houses and stale rock cakes,  
[but now] it's much more exciting.

Although Contrast is already specified through intra-clausal material, viz. temporal ('was' – 'is') and lexical ('dowdy' – 'exciting') contrast, it is additionally cued by extra-clausal 'in the past' and 'but now', resulting in an overspecification of the DR.

In our data, Contrast and Corrective Elaboration are overspecified consistently, other DRs are fully specified, and still others are underspecified and sometimes supplemented with a discourse connective. Extra-clausal cueing of DRs, we argue, may be understood as a strategic device to ensure activation of relevant defining conditions and particularized features of frequently overlapping DRs such as Comment and Result and Explanation, Elaboration and Background. With respect to Corrective Elaboration and Contrast, which are also consistently indexed through all their defining conditions and particularized features throughout, overspecification seems the default for contrastive or corrective relations.

Unlike underspecification of DRs, which may reflect cognitive economy in that speakers only add as much information as appears 'such as is required' for the linguistic realization of DRs, overspecification may increase the salience of a particular DR and signify its relevance to the hearer.

### References

- Asher, Nicholas & Alex Lascarides. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.  
Dik, Simon. 1997. *The Theory of Functional Grammar* (2 vols). Ed. K. Hengeveld. Amsterdam: Benjamins.  
Fetzer, Anita and Augustin Speyer (forthcoming): *Discourse Relations across Genres and Contexts: A Contrastive Analysis of English and German Discourse*.  
Givón, Talmy. 1993. *English Grammar. A Function-Based Introduction* (2 vols). Amsterdam: Benjamins.

# Different Systems Of Organization Of Spanish Discourse Markers And A New Proposal For Spanish As An L/2

Anais Holgado Lage

## Abstract

Discourse markers, "sequentially dependent elements which bracket units of talks"

(Schiffrin 1987) ranging from mere sounds to complex units, are widely accepted as essential linguistic expressions that every student of a second language should know in order to clearly organize their speech and express their ideas. However, in the case of Spanish, despite being a language studied by a large number of students worldwide, discourse markers and their application to the classroom has not been studied in depth until very recently (e.g. Martí Sánchez and Fernández Gómiz 2013). This study examines if the current systems of classification of these elements are useful and appropriate when teaching discourse markers in a setting of Spanish as an L/2.

In this research I analyze some of the systems of organization of discourse markers in

Spanish – including the best-known functional classification in Spanish by Martín Zorraquino and Portolés (1999), the very similar one adopted by the Instituto Cervantes in the Plan Curricular (2006-07) and some others like Briz (1998), Fraser (1999) or Pons Bordería (2006)–, by comparing them to other systems of organization commonly used in Spanish as a Second Language, in order to determine if applying them to this field could have a positive impact. The results show that these systems, most of them designed to be used by other specialists, would not constitute a useful starting point to teach discourse markers in Spanish due to their intricacy and the fact that there are a high number of elements that do not have a defined place within the organization.

For this reason, I finally propose a new and simpler functional classification of discourse markers, very similar to the one shown in Martí Sánchez (2008), which could be more approachable when teaching these elements. This organization consists of three big groups (linking words, conversational markers and discourse organizers), but ultimately divides all the expressions in 59 groups by their primary function. This study contributes to the understudied field of discourse markers in Spanish as a Second Language and underlines the necessity of creating new materials that are fully appropriate for both students and teachers of this language.

## References

- Briz, A. (1998): *El español coloquial en la conversación. Esbozo de pragmatografía*, Barcelona: Ariel.
- Fraser, B. (1999): «What are discourse markers?», in *Journal of Pragmatics*, 31, 931-952.
- Instituto Cervantes (2007-07): *Plan curricular del Instituto Cervantes, Niveles de referencia para el español (3 volúmenes)*, Madrid: Biblioteca Nueva.
- Martí Sánchez, M. (2008): *Los marcadores en español L/E: conectores discursivos y operadores pragmáticos*, Madrid: Arco Libros.
- Martí Sánchez, M. & S. Fernández Gómiz (2013): *Los marcadores discursivos. Para estudiantes y profesores de español como lengua extranjera*, Madrid: Edinumen.
- Martín Zorraquino, M<sup>a</sup> A. & Portolés, J (1999): «Los marcadores del discurso», in I. Bosque y V. Demonte (eds.), *Gramática descriptiva de la lengua española*, Madrid: Espasa, vol 34051-4214.
- Pons Bordería, S. (2006): «A functional approach to the study of discourse markers», in Fischer, K. (ed.) *Approaches to discourse particles*, Oxford – Amsterdam: Elsevier, 77-99.
- Schiffrin, D. (1987): *Discourse makers*, Cambridge: Cambridge University Press.

## Morphological and syntactic characteristics of discourse connectives in Czech

Pavlina Jinová

### Abstract

On the basis of annotation of discourse relations anchored by surface present connectives in the Prague Dependency Treebank (PDT, corpus of Czech mostly journalistic texts, approx. 50,000 sentences, 19,000 discourse relations in the latest version, PDT 3.0 (Bejček et al., 2013), connective-based F1-measure 0,83 (Jinová, Mírovský and Poláková, 2012)), this contribution presents a basic overview of characteristics of Czech connectives. In our approach, inspired by the PDTB lexical approach (Prasad et. al., 2008) and by Prague linguistic tradition, discourse connectives are subset of discourse markers whose function is to join text spans either within a single sentence or between separate sentences and signal the semantic relation between them.

When all means with discourse connecting function identified in the PDT are taken into consideration, these three groups can be delimited: the most common group is represented by word units, then there are numbers and letters (typically in list structures) and finally, some punctuation marks seem to have text structuring function as well. In this contribution, we concentrate on the word units only and among them only on those with morphologically stable form, so excluding the secondary connectives (e.g. the conditions was, for this reason, cf. Rysová and Rysová, 2014).

Discourse connectives with stable form exhibit following basic features. Typically, they are non-declinable expressions apart from connectives *což* (what, that), *aby* (in order to) and *kdyby* (if) (the first of them is declined, the last two connectives are conjugated).

The annotation revealed that apart from connectives consisting of one or two word units, complex connectives *na na jedné straně na druhé straně však* (on the one hand on the other hand however) can be found as well.

If an origin of connectives is concerned, only one modern connective is opaque, the connective *a* (and). The rest of connectives is analyzable either from the synchronic or from the diachronic perspective.

Another classification criterion is a position of a connective in a sentence (e.g. the connective *však* (but) in contrast meaning is obligatory a clitic).

Finally, the most complex classification is represented by PoS appurtenance, which lacks complete treatise in Czech tradition. Having analyzed studies devoted to particular connectives in Czech (e.g. Bauer, 1972, Pešek, 2011, Bedřichová, 2008), these criteria seem to be useful: conjunctions are expressions with primary connecting function and with obligatory initial position in a sentence, they do not influence a position of clitics and cannot be rhematized, further, they do not combine with the typical conjunction *a* (and). On the other hand, adverbs and particles have other functions than connecting one as well (adverbs are sentence constituents, particles typically signal attitudes). Adverbs can be combined with typical conjunction *a* (and), can be rhematized and often signal some coreference relation with surrounding context. Finally, particles with connecting function often have a presupposition as a part of their meaning. However, for all these categories, there are of course some questionable cases. They will be presented in the full version of this contribution as well as more elaborated description of all criteria.

## References

- Bauer, Jaroslav (1972): Spojky a příslovce. In: J. Bauer, *Syntactica slavica : Vybrané práce ze slovanské skladby*. Brno: Universita J. E. Purkyně, 1972, pp. 351–360.
- Bedřichová, Zuzanna (2008): Částice implikující presupozici jako podstatná složka větného významu. *Čeština doma a ve světě*, 16, pp. 119–126.
- Bejček, Eduard, Hajičová, Eva, Hajič, Jan, Jínová, Pavlína, Kettnerová, Václava, Kolářová, Veronika, Mikulová, Marie, Mírovský, Jiří, Nedoluzhko, Anna, Panevová, Jarmila, Poláková, Lucie, Ševčíková, Magda, Štěpánek, Jan, Zikánová, Šárka (2013): Prague Dependency Treebank 3.0. Praha: Univerzita Karlova v Praze, MFF, ÚFAL, <http://ufal.mff.cuni.cz/pdt3.0/>.
- Jínová, Pavlína, Mírovský, Jiří, Poláková, Lucie (2012): Analyzing the Most Common Errors in the Discourse Annotation of the Prague Dependency Treebank. In: *Proceedings of the 11th International Workshop on Treebanks and Linguistic Theories*, Lisboa, Portugal, pp. 127–132.
- Pešek, Ondřej (2011): *Argumentativní konektory v současné francouzštině a češtině*. České Budějovice: Editio Universitatis Bohemiae Meridionalis.
- Prasad, Rashmi, Dinesh, Nikhil, Lee, Alan et al. (2008): The Penn Discourse Treebank 2.0. In: *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*. Marrakech, Morocco, pp. 2961–2968.
- Rysová, Magdaléna, Rysová, Kateřina (2014): The Centre and Periphery of Discourse Connectives. In: *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computing*, Chulalongkorn University, Bangkok, Thailand, pp. 452–459.

## Cognitive complexity in causal relations in adult L2-writing

Janne Laitinen

### Abstract

The research in progress investigates cognitive complexity in adult L2 writing in Finnish, more precisely, how adult second language users acquire and signal causal, contrastive, conditional and concessive relations (CCCC) in discourse. Previous studies (e.g. Evers-Vermeul & Sanders 2009, 2011) have shown that cognitive competence and linguistic competence develop hand in hand in (child) L1 acquisition. By contrast, an adult L2 user has already acquired the cognitive competence to differentiate CCCC-relations in his/her first language. Yet there seems to be a hierarchy in the order how conjunctions signaling CCCC-relations occur in adult L2 writing. The most evident example of this phenomena is the absence or at least very marginal occurrence of concessive conjunctions before upper intermediate level (B2) texts. In spite of the missing explicit linguistic marker, for instance, the concessive conjunction, adult language users have the need to express these relations. The aim of the study is to explore different explicit and implicit means by which adult L2 writers signal the above mentioned relations in discourse. The variation of semantic and pragmatic meanings/functions of the most frequent CCCC conjunctions will be examined.

In my presentation I will focus on causal relations and explore the overall occurrence of both explicit and implicit causal relations. Furthermore, I will characterize the different categories of causal relations that the most frequently used causal conjunctions convey. The continuum presented in the Common European Linguistics Framework of Reference for Languages (CEFR) proposes that semantic and pragmatic relations are cognitively more complex on higher level of language proficiency. I will examine content, epistemic and speech-act causality in texts produced by L2 writers, and analyze, whether their occurrence is related to the levels of language acquisition in beginner (A1), elementary (A2) and intermediate (B1) levels, which for example define argumentation developing later than reason.

The data has been collected from Finnish adult migrant placement testing and includes three different text types (descriptive, narrative and argumentative) from 2397 different Finnish L2-writers. The writers are from different nationalities, educational backgrounds and ages at language levels from A1 to B1.

The presentation will contribute to the theme of the conference by conjoining formal and functional approaches to the L2-acquisition of DRDs.

### References

- Evers-Vermeul, J., & Sanders, T. (2009). The emergence of Dutch connectives; How cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language*, 36, 829-854.
- Evers-Vermeul, J., & Sanders, T. (2011). Discovering domains – on the acquisition of causal connectives. *Journal of Pragmatics*, 43, 1645-1662.

## The functional equivalents of French *Pourtant* in English and Portuguese

Pierre Lejeune

### Abstract

Even though it seems to be possible to identify discourse relations that are universal (Asher et al. 2001; Prasad et al. 2014), «discourse connectives are often said to be language specific, and therefore not easily paired with a translation equivalent in a target language.» (Zufferey & Cartoni 2012). Classifying DRDs of different languages into categories – fine-grained as they may be – according to the type(s) of relation(s) of which they are markers doesn't provide translators with a toolbox from which they can choose in a target language a functional equivalent of a given source language's marker. We believe that each marker interacts with its context in a specific way, so that one does not get functional equivalents within a language and between languages that are valid in all contexts.

In French *pourtant* is often considered – and taught to translators – as equivalent to *néanmoins*, *cependant* and *toutefois* as a marker of contrast/concession. In order to uncover the nuances between them, the analysis of parallel corpora may prove useful.

We use the Europarl corpus (Koehn, 2005), a multilingual collection of the minutes of debates from the European Parliament. Working with multilingual corpora of international institutions has the advantage of maintaining between source and target texts the same cultural context and communicative function (Scarpa 2010). On the other hand, «texts that originate in the institutions of the European Union can be problematic for researching the activity of translation [...], since it can be difficult to assign the status of 'source text' to one of the language versions» (Olohan 2004; 25). In the case of the Europarl corpus, some marking yet unsystematic of the source language exists, and can be exploited (Cartoni et al. 2013).

We will compare Lebaud 1997's characterization of *pourtant* (In P *pourtant* Q, Q, presented as indisputable, is associated with P', the complementary of P, which brings qualitative instability to P, with associated contextual effects of concession, surprise/anormality, questioning or refutation) with Bell 2010's characterization of *yet* (as a «concessive cancellation discourse marker» which, contrary to *still*, corresponds, in P *yet* Q, to the speaker's point of view located in P).

We will also discuss examples of sentences containing *pourtant* (in NP embedded positions) where the semantic equivalence between the French, English and Portuguese versions is problematic.

We will argue that even recurrent functional equivalences between DRDs of different languages (e.g. in the case of EU documents; the propagation of errors through translation memories; Cartoni et al. 2013) have to be subject to critical scrutiny regarding their acceptability (Lansary & Leroux 2010).

## References

- Anscombre J.-C. (1983). «Pour autant», «pourtant» (et «comment») : à petites causes, grands effets». *Cahiers de Linguistique Française* 5 : 37-84
- Asher N., Busquets J. & Vieu L. (2001). «La SDRT : une approche de la cohérence du discours dans la tradition de la sémantique dynamique». *Verbum*, 23 : 73-101.
- Bell D. L. (2010). «Nevertheless, still and yet: Concessive cancellative discourse markers. *Journal of Pragmatics*, 42 : 1912-1927.
- Cartoni B., Zufferey S. & Meyer, T. (2013). «Using the Europarl corpus for linguistic research.» *Belgian Journal of Linguistics*, 27: 23-42.
- Lansari L. & Leroux A. (2010). «Quel(s) corpus pour l'analyse contrastive ?». In P. Cappeau, H. Chuquet et F. Valetoupolos (éds), *L'exemple et le corpus : quel statut ? Travaux Linguistiques du CERLICO*, 33 : 43-58.
- Koehn, Ph. (2005). «Europarl: A parallel corpus for statistical machine translation». In *Proceedings of MT Summit 2005*: 79-86.
- Lebaud D. (1997). «Pour et pourtant. Un même mode de construction d'une relation». *BULAG* 22 : 155-179.
- Lebaud D., Rhee E. M (2000). «Cependant, Pourtant et altération d'un dire». In *Répétition, altération, reformulation, Colloque International, Besançon, juin 1998. Annales Littéraires de l'Université de Besançon* : 67-90.
- Olohan, M. (2004). *Introducing Corpora in Translations Studies*. London & New York: Routledge.
- Prasad R., Webber B. & Joshi A. (2014). *Reflections on the Penn Discourse TreeBank, Comparable Corpora, and Complementary Annotation*. 40 (4) : 921-950.
- Scarpa, F. (2010). *La traduction spécialisée : une approche professionnelle à l'enseignement de la traduction*, trad. et adapt. Marco A. Fiola. Ottawa : University of Ottawa Press.
- Tricás M. (1990). «L'argumentation concessive française et espagnole : le transfert en espagnol de quelques connecteurs (pourtant, quand même, même si, or...) et d'autres formules concessives». *Meta Translators' Journal*, 35(3) 529-537.
- Zufferey, S. & Cartoni, B. (2012). «English and French causal connectives in contrast». *Languages in Contrast* 12(2): 232-250.



## Negative Emotion Pragmatic Markers – Corpora and Cross-Linguistic Contrasts

Barbara Lewandowska-Tomaszczyk and Paul A. Wilson

### Abstract

The main goal of this work is to present a corpus-based quantitative and qualitative analysis of the properties of negative emotion pragmatic markers and develop a functional cross-linguistic scale of such markers in English and Polish. The last few decades have witnessed a particularly strong interest in the cognitive and linguistic properties of emotions (Kövecses 1990). Corpus linguistics makes it possible to trace regularities and idiosyncrasies in emotion expression and to analyse these from a contrastive, cross-linguistic perspective.

Negation and negativity are a cognitively more salient and less controllable device in discourse than corresponding positive forms (Lewandowska-Tomaszczyk 1996). Therefore the tracing of the elements of negativity in discourse markers can contribute to emotion and negativity research as well as uncover new vistas in the analysis of discourse markers that also include a contrastive element. Pragmatic markers in general contribute to the creation of Event Structures of a varying degree of complexity

The research methodology employed combines discourse and lexical analyses in terms of the Cognitive Corpus Linguistics approach, which involves available corpus tools in the BNC and the National Corpus of Polish (nkjp.pl) to carry out a linguistic and discourse analysis. Parallel corpora are also used as an important source of cross-linguistic data. The qualitative parameters of the analysis explore types of lexical patterns, lexical choices, and discourse behaviour preferences.

Negativity is a complex notion. The prototypical negation, conceptualized in terms of categorial exclusion, puts the unit outside the relevant category named in the utterance, while different forms of negativity imply a potential exclusion of the referent and/or their properties or conditioning from the current discourse domain (Seuren 1985). On the other hand, negativity makes it possible to introduce into the current discourse a number of alternative realities (mental spaces - Facuconnier 1985), expressed in terms of multiple world-creating devices such as some functions of the linking marker *oh no*, activated in use with varying emotionality charge.

As the meaning associated with negative discourse-linking devices is likely to be largely dependent on context, one would also expect the emotions that they convey to also be determined by the situation. We aim to compare the emotions associated with the main meanings of the negative linking device *oh no/no nie* in English and Polish, first elaborated on by Lewandowska-Tomaszczyk (2004). As to our knowledge this area has not thus far been the major focus of scholarly investigation, our work should be seen as an initial attempt at the mapping of emotionality in such forms.

The markers *no nie* in Polish and *oh no* in English are multifunctional mental space-builders. They express inter alia:

(1) emphatic direct truth-conditional negation (highest frequency from among 2245 examples):

Oh no, she definitely wants to go

(2) dialogue with oneself (with audience present or absent):

And a dog oh no watch Graham

The Polish and English functions overlap to a certain extent but the analysis of the parallel corpora displays a considerably more complex set of relationships, both interactional and self-dialogic:

No nie, może nie - Why, no – perhaps not

No nie, o co mu chodzi? – Oh, I say, what's up?

Ale draka, no nie? – Funny, ain't it

The study presents the complexity of the relationships. To conclude, a classificatory system of negativity and emotionality ranking across the two languages is proposed, upon which an annotation system can be designed.

## References

- Fauconnier, G. (1985). *Mental Spaces: Aspects of Meaning Construction in Natural Languages*. Cambridge, Mass., A Bradford Book. The MIT Press, 1985
- Kövecses, Z. (1990). *Emotion Concepts*. NY: Springer-Verlag.
- Lewandowska-Tomaszczyk, B. (2004). "Conceptual blending and discourse functions". *RiL2*. 33-47.
- Lewandowska-Tomaszczyk, B. (1996). *Depth of Negation. A Cognitive Semantic Study*. Lodz: LUP.
- Lewandowska-Tomaszczyk, B. & P. Wilson (2011). "Culture-Based Conceptions of Emotion in Polish and English". In: St. Goźdz-Roszkowski (ed.). *Explorations across Languages and Corpora*. Lang: Frankfurt. 229-239.
- Seuren, Pieter A. M. (1985). *Discourse Semantics*. Oxford: Blackwell.

## Las hipótesis sobre la gramaticalización de por cierto como MD epistémico en español: oralidad, escrituralidad y perspectiva cognitivo-funcional

Araceli López-Serena

### Abstract

El surgimiento de la partícula por cierto (PC) con valor epistémico ('ciertamente') en español ha sido explicado de forma totalmente opuesta por María Estellés (2006, 2009a, 2009b, 2011) y Silvia Iglesias (2015). Para la primera, quien entiende la gramaticalización de esta partícula como un reto frente a la cadena de gramaticalización prototípica de MD propuesta por Traugott (1995), el uso de PC como modificador oracional se habría gestado de manera abrupta en entornos de escrituralidad estrechamente asociados con necesidades de traducción; para la segunda, quien asocia los valores modales de PC con contextos de oralidad, el surgimiento de esta estructura se adecua al itinerario de gramaticalización canónico de MD (adverbial phrase > sentence adverbial > discourse marker). En relación con ambas propuestas, el objetivo de mi comunicación será llevar a cabo una confrontación metateórica de ambas posturas, con el fin de poner de relieve la oposición, en la reconstrucción del deslizamiento semántico subyacente al proceso de gramaticalización de PC, entre explicaciones cognitivistas y patrones explicativos prototípicamente universalistas por una parte, y explicaciones de corte historicista y cultural.

## What does context mean when referring discourse markers?

Ana Llopis Cardona

### Abstract

From the first studies on discourse markers, the influence of context on the determination of functions was underlined, but the concept of context often is explained in a general and imprecise way. Therefore we consider in this paper what does context mean when referring discourse markers, specifically we will focus on the verbal context which contributes to specify the function of discourse markers. We will aim at providing an answer basing on empirical and descriptive works derived from corpus data study.

Analysing examples of different discourse markers, we notice that verbal context could be the discourse operation (a paraphrase, an explanation, a justification, etc.), the type of discourse (narrative, reportative, etc.), the type of discourse unit (intervention, act or subact) or the speech act (assertives, expressives, directives...).

The discourse operation is relevant to delimit the functions of some discourse markers. For example, the core meaning of *en efecto* is confirmation, which gives rise to two pragmatic meanings or functions: first, *en efecto* confirms a content (explicit or implicit) and introduces a reformulation of a previous unit (if the content was explicit), and second, *en efecto* confirms an assertion previously uttered and introduces an evidence to prove it.

Also it is important the type of discourse. The core meaning of *por su parte* consists of continuing the present topic and adding information. When the context where it appears is reported speech (not a narration), *por su parte* has a hybrid function: it adds information and also it points to the source of information. Similarly, the type of speech act plays an important role. When *al menos* is used in a neutral assertive act, it indicates that the information is true and implies that it may be a bigger extent; by contrast, when it is used in an evaluative assertive act, it communicates a positive consideration, since a minimum condition is achieved.

Regarding the discourse markers with an important role in interaction, the most important factor seems to be the type of discourse unit and the position (initial, middle or final) (cfr. Briz & Pons 2010). For instance, *¿verdad?* asks for checking and allows a turn change when it is used at the end of a turn, whereas it points out the uttered information as if this was shared and does not ask for checking when it appears in the middle of the turn.

### References

- Briz, Antonio, Pons, Salvador & Portolés, José (2008). *Diccionario de Partículas discursivas del Español*. Servei de Publicacions de la Universitat de València [en línea]. <<http://www.dpde.es>>
- Fischer, Kerstin (2006). *Approaches to Discourse Particles*. Amsterdam: Elsevier.
- Loureda, Óscar & Esperanza Acín (coords). *Los estudios sobre marcadores del discurso en español, hoy*. Madrid: Arco Libros.
- Llopis Cardona, Ana (2014). *Aproximación funcional a los marcadores discursivos. Análisis y propuesta lexicográfica*. Frankfurt: Peter Lang.
- Schiffirin, Deborah (2001). "Discourse markers: Language meaning and context". In Schiffirin, Deborah; Tannen, Deborah; Hamilton, Heidi (eds.) *Handbook of Discourse analysis*. Oxford: Basil Blackwell, 54-75.

## Presupposition and semantic change: the case of *pero*

Patrícia Matos Amaral and Susana Rodríguez Rosique.

### Abstract

This paper focuses on the role of presupposition in semantic change (Schwenter & Waltereit 2010), by discussing the relation between the origin of the adversative meaning of Spanish *pero* 'but', which is based on an anaphor, as represented in (1), and its synchronic behavior. We analyze the roles of the presuppositional component in that development, from the explicit anaphoric expression to a connective that builds on information that can either be retrieved from context or must be accommodated by the interlocutor on the basis of general expectations.

(1) *per hoc* 'because of that' > *peró* 'despite (that) / nevertheless / however' > *pero* 'but' (Castillo Lluch 1993; Herrero Ruiz 1999; Saldanya & Salvador to appear: 73)

While the presuppositional behavior is originally due to the Latin demonstrative pronoun (*hoc*), it becomes part of the meaning of the connective. Once the sequence has lexicalized, the presupposed element may be explicitly expressed by another demonstrative (2):

(2) *Et el Rey sopo como uenie el Duc. a el pora prenderle. & non sopo al que fazer si non que tomo una saya de un su escudero. & uistiola & fuesse pora la cozina & tomo los capones. & assentosse a assar los. pero esto non sabemos por cierto si fue assi* (CORDE, 1293, Gran Conquista de Ultramar)

*Pero* initially occurs in negative contexts, and is placed at the beginning of the second conjunct, which is why it starts to be associated with contrast; specifically, with a result that contravenes what is expected from the previous information. In this contrastive context, the conjunct introduced by *pero* imposes the conclusion, so it is considered the strong member in argumentative terms (Anscombe & Ducrot 1994). It may appear in structures where the previous (weak) member is introduced by a concessive conjunction (3):

(3) *Et maguer uoluntad es del padre o de la madre non casando que tenga sus hijos si quisiere, pero si el auuelo o auuela... los quisiere tener... sean tollydos al padre o ala madre* (Fuero de Soria, Apud. Vallejo 1925).

At this stage, the presuppositional component is contrastive. From there on, *pero* will be able to introduce both a conclusion being contrary to a previous argument (direct counter-argumentation) and an argument leading to a conclusion contrary to the one triggered by a previous argument (indirect counter-argumentation).

In the next step, as observed in its synchronic behavior (4), *pero* can leave the contrastive contexts and may be used to provide argumentative strength (Ducrot 1980) to an insufficient co-orientated argument (Acín 1993; Portolés 1998; Garrido 2004). Here, *pero* is still presuppositional, now at the level of discourse argumentation.

(4) *Era como una segunda madre de todos, tanto por su autoridad como por sus desvelos, pero además se ocupaba de cualquier extraño que le tocara el corazón* (Noticia de un secuestro. Apud. Portolés 2004).

In conclusion, the different stages analyzed in this paper show that the presuppositional nature of the original anaphoric expression is not lost, but acquires first a contrastive and then an argumentative function.

## References

- Anscombe, J. C. & O. Ducrot (1994): *La argumentación en la lengua*. Madrid: Gredos.
- Acín, E. (1993): *Aspectos de la adversación en español actual*. La Coruña: Universidad de La Coruña.
- Castillo Lluch, M. (1993): "Acercamiento a las partículas adversativas medievales", *Cahiers de linguistique hispanique médiévale*, 18-19, 219-242.
- Ducrot, O. (1980): *Les échelles argumentatives*. Paris: Minuit.
- Garrido, M. C. (2004): *Conectores contraargumentativos en la conversación coloquial*. León: Universidad de León.
- Herrero Ruiz de Loizaga, F. J. (1999): "Sobre la evolución de las oraciones y conjunciones adversativas", *Revista de Filología Española*, 79, 291-328.
- Pérez Saldanya, M. & V. Salvador (forthcoming): "Oraciones Subordinadas Concesivas", in C. Company (dir.): *Sintaxis histórica de la lengua española. Tercera parte: Adverbios, preposiciones y conjunciones. Relaciones interoracionales*. México: Fondo de Cultura Económica-Universidad Nacional Autónoma de México.
- Portolés, J. (1998): "El concepto de suficiencia argumentativa", *Signo y Seña*, 9, 201-224.
- Portolés, J. (2004): *Pragmática para hispanistas*. Madrid: Síntesis.
- Schwenter, S. and R. Waltireit (2010): "Presupposition accommodation and language change", in K. Davidse, L. Vandelanotte & H. Cuyckens (eds.): *Subjectification, Intersubjectification and Grammaticalization*, Berlin / New York: Walter de Gruyter, 75-102.
- Vallejo, J. (1925): "Sobre un aspecto estilístico de Don Juan Manuel", in *Homenaje a Menéndez Pidal, II*, Madrid: Hernando, 63-85.

## THE ROLE OF EXEMPLIFICATION IN DISCOURSE

Caterina Mauri and Alessandra Barotto

### Abstract

The aim of this paper is to analyse the use of exemplifying constructions (e.g. for example, to take an example, etc.) in discourse, based on data from spoken Italian and English. Exemplification plays an important role in human cognition at a cognitive level in categorization processes. However, we will argue that it is also an important linguistic tool that allows speakers to achieve several other functions having to do with the organization of discourse, the structuring of information and the pragmatic dimension of interpersonal relations between speaker and hearer.

Exemplification indeed construes the mentioned exemplar(s) as potential option(s) taken from a hypothetical set, and it is precisely the modal dimension of potentiality that makes exemplification a good candidate to express more pragmatic functions. Although a comprehensive study on the linguistic coding and functions of exemplification is still lacking, the existing literature on specific languages has shown that exemplifying strategies can be used to achieve pragmatic effects in discourse. Specifically, exemplification may be used to be vague and attenuate the speaker's commitment (Channell 1994, Overstreet 1999, Voghera 2013). It may also be employed as a focus marker, especially in requests for information or clarification (Voghera 2013, Ohori 2004 for Japanese). Furthermore, exemplification may convey an additive relation (Halliday and Hasan 1976).

Through a two-stage approach, we will systematically account for the behaviour of exemplifying constructions in discourse, providing empirical evidence for regular associations between specific pragmatic functions and specific distributional and morphosyntactic properties of the strategies employed. In the first stage we will undertake a corpus-based study on Italian and English, using the Santa Barbara Corpus of spoken American English (spoken American English, 249.000 words) and the *Lessico dell'Italiano Parlato* (spoken Italian, 490.000 words). Special attention will be devoted to i) the presence of overt labels for the categories (sets or frames) of which the exemplar(s) are taken to be representative, ii) the position of the exemplifying construction in the sentence (sentence initial, internal or final), iii) the co-occurrence with other discourse relational devices, and iv) the topic continuity of the mentioned exemplar(s) in discourse. In the second stage, we will verify the hypotheses made on the basis of the corpus analysis, through a set of specific tests. The tests will consist of sentences selected from the corpus, associated to a closed set of compatible interpretations regarding the context, the speaker's intentions and assumptions, and the delimitation of the category to which the exemplar(s) belong. 30 Italian native speakers will be asked to assign a plausibility rate to the different interpretations.

We will show that it is possible to distinguish 1) 'purely cognitive' exemplification, aimed at linking exemplars to their category and construing the category in a context-relevant way, from 2) 'cognitive-pragmatic' exemplification, where a pragmatic dimension of vagueness and hedging functions is added on top of the cognitive one, and 3) 'purely pragmatic' exemplification, in which there is no category to be construed and the exemplifying construction works as a focus marker or a discourse relational device.

### References

- Channell, J. (1994). *Vague Language*, Oxford: Oxford University Press.
- Halliday, M. A. K., and Ruqaiya Hasan. (1976). *Cohesion in English*. London: Longman.
- Ohori, T. (2004). "Coordination in Mentalese," in Martin Haspelmath (ed.) *Coordinating Constructions*, 41-66, Amsterdam: John Benjamins.
- Overstreet, M. (1999). *Whales, Candlelight, and Stuff Like That: General Extenders in English Discourse*, New York: Oxford University Press.
- Voghera, M. (2013). "Tipi di tipo nel parlato e nello scritto". In *Di Linguistica e di Sociolinguistica, Studi offerti a Norbert Dittmar* Pag. 185-195 Roma, Bulzoni.

## Processing causality in Spanish: the case of "por tanto" and "por eso"

Laura Nadal and Inés Recio Fernández

### Abstract

The aim of this communication is to illustrate with experimental evidence how the Spanish consecutive connectives "por tanto" and "por eso" (en. 'therefore', 'as a result') constrain information processing.

Connectives are considered a subtype of discourse markers (Portolés 2001 [1998], Portolés & Martín Zorraquino 1999, Montolío 2001, among others). As such, they act as constraints for inferential processes in communication (Blakemore 1987, 1992). Specifically, consecutive connectives connect semantically and pragmatically two or more arguments, which are accessible within discourse or within the cognitive environment of the interlocutors, and – due to their mainly procedural meaning – lead the hearer towards a conclusion derived from all discourse members as a whole (cf. Martín Zorraquino/Portolés 1999: 4093 ss.; Portolés 1993, 2004: 289).

Both "por eso" and "por tanto" are employed to signal cause-consequence relations. However, they display semantic differences, also with regard to the discursive and informative status they convey to the segments they link. While "por eso" introduces "known or presupposed information (information not presented as new)" (Briz et al. 2008) and points towards the cause rather than the consequence, which, in turn, is not expected to be "derived logically" (Montolío 2001: 123) from the premise of the first segment (1), "por tanto" precedes a discourse segment with the status of a conclusion reached "by means of reasoning" (Briz et al. 2008, cf. also Montolío 2001) what is stated in the first segment (2):

(1) Franco quería seguir gobernando hasta su muerte.[explicative cause] Por eso[connective] trajo al Príncipe Juan Carlos, que era más joven. [explicative cause]

'Franco wanted to rule until his death. Por eso, he brought Prince Juan Carlos, who was younger than him.'

(2) Franco quería seguir gobernando hasta su muerte. [basis for the reasoning] Por tanto, [connective] trajo al Príncipe Juan Carlos, que era más joven. [consequence derived from reasoning]

'Franco wanted to rule until his death. Por tanto, he brought Prince Juan Carlos, who was younger than him.'

(extracted and partially adapted from REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [en línea]. Corpus de referencia del español actual. <<http://www.rae.es>> [17-7-2015])

A number of experimental works on causal relations have shown that subtle semantic differences between connectives like the above-mentioned can influence utterance processing (cf. for instance Canestrelli et al. 2012 for Dutch "omdat" and "want" 'because' in terms of subjectivity). In this sense, the cognitive effects of "por eso" and "por tanto" were checked for experimentally with 80 native speakers of Spanish. Participants were confronted with a series of causal utterances that they read silently on a computer screen while their eye movements were tracked. This allowed us to assess the cognitive effort and the reactions (duration of fixations, regressions to certain areas relevant to process the causal relation) needed to process utterances (stimuli) during reading (Rayner 1998). Furthermore, comprehension tests were carried out to check how each connective affects derivation of the implicatures of the utterances (cf. Just et al. 1982). In our talk, the results



of these eyetracking and self-paced reading experiments will be presented and discussed with special reference to theoretical descriptions available so far of "por eso" and "por tanto".

## References

- BLAKEMORE, D. (1987): *Semantic constraints on relevance*. Oxford: Blackwell.
- BLAKEMORE, D. (2002): *Relevance and Linguistic Meaning*, Cambridge: Cambridge University Press.
- BRIZ, A., PONS, S. & J. PORTOLÉS (coord.) (2008): *Diccionario de partículas discursivas del español*. Online, [www.dpde.es](http://www.dpde.es).
- CANESTRELLI, A.R., MAK, W.M. & T.J.M. SANDERS (2012): "Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements", *Language and Cognitive Processes*, 28:9, 1394-1413, DOI: 10.1080/01690965.2012.685885.
- JUST, M., P. CARPENTER & J. WOOLLEY (1982): "Paradigms and processes in reading comprehension", *Journal of Experimental Psychology* 111, 228-238.
- MARTÍN ZORRAQUINO, M. A. & J. PORTOLÉS (1999): "Los marcadores del discurso", in I. BOSQUE & V. DEMONTE (coord.), *Gramática descriptiva de la lengua española*, Madrid, Espasa Calpe, Vol. 3, 4051-4213.
- MONTOLÍO, E. (2001): *Conectores de la lengua escrita*, Barcelona: Ariel.
- PORTOLÉS LÁZARO, J. (1993): "La distinción entre los conectores y otros marcadores del discurso en español", *Verba* 20, 141-170.
- PORTOLÉS LÁZARO, J. (2001 [1998]): *Marcadores del discurso*, Barcelona, Ariel.
- PORTOLÉS LÁZARO, J. (2004): *Pragmática para hispanistas*, Madrid: Síntesis.
- RAYNER, K. (1998): "Eye movements in reading and information processing. 20 years of research", *Psychological Bulletin* 124/3, 372-422.

## Rhetorical Relations Marked by 'You Know' in the Context of Humor

Anna Nagy

The paper's aim is to present the results of ongoing research that focuses on different approaches relating to the role of humor in the (de)construction of coherence. Humor has already been subject to pragmatic considerations out of which its relevance theoretic (Sperber-Wilson 1985) interpretation showed a probable connection between the functional distribution of certain discourse markers and the nature of humor, the latter of which is highly dependent on reference and background information (Nagy 2015).

Though the discourse marker *you know* was involved in the aforementioned research and seemed to behave as part of this larger, hearer oriented strategy necessary for comedy, switching to a discourse analytic point of view poses a problem. Realizing that most relations are not even explicitly marked in a text (Taboada 2006), the actual occurrences of a discourse marker – regardless of function – do not exactly point to a larger discourse organizational scheme. Due to this circumstance the aim of this presentation is to work with RST annotation (Mann-Thompson 1988) and see if the number of tokens of *you know* are relevant to any relations or interconnections found within a larger body of text, with respect to the purpose of humor.

The research will be carried out through the compilation and annotation (identifying co-occurrence patterns, position in turn/utterance, speaker's attitudes) of a corpus that includes around 800.000 words of the sit-com *How I Met Your Mother*, and a control corpus of current, spoken American English. *You know*, described as a discourse marker linked to the accessibility of cognitive environments (Fox Tree-Schrock 2002) and thus the contextual origins of humor, should be more visible in a humorous text, as well as with regard to the relations it exemplifies. Larger deviations from the (non-humorous) reference corpus could serve as an indication of a coherence structure especially suited for the speakers' intentions to transmit traditionally ambiguous messages, and for the hearers' intentions to receive them, for the sake of laughter.

### References

- Fox Tree, J., Schrock J. 2002. Basic meaning of *you know* and *I mean*. *Journal of Pragmatics* 34(6): 727-747
- Mann, W.C., Thompson, S.A. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8 (3). 243-281.
- Nagy, A. 2015. Az angol diskurzusjelölők szerepe a stand-up comedyben. In: Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2015 IX. Alkalmazott Nyelvészeti Doktoranduszkonferencia 30-41.
- Sperber, D., Wilson, D. 1995. *Relevance: Communication and Cognition*. Oxford: Blackwell Publishers
- Taboada, M. 2006. Discourse Markers as Signals (or Not) of Rhetorical Relations. *Journal of Pragmatics* 18(4): 567-592

## Hesitations, repetitions, restarts and self-corrections: Towards a systematization of discourse planning phenomena in Spanish colloquial conversations

Elena Pascual Aliaga

### Abstract

One of the main properties of colloquial discourse that has been highlighted since Beinhauer ([1929]1963: 9) defined Spanish spoken discourse as “speech as it arises naturally and spontaneously in daily conversation” is the spontaneity of the process of constructing this discourse (Narbona, 1991). The spontaneous planning of discourse gives rise to the appearance of particular phenomena like discourse ruptures or fragmentary segments (Briz, 1998). This paper focuses on a group of elements called, within the proposal of the Val.Es.Co. (Valencia Español Coloquial) research group, substructural elements: they comprise hesitations, repetitions, restarts and self-corrections, which are defined as non-intentional segments that are external to the informative and syntactic structure of discourse (Pérez, 2003). Since substructural elements offer an insight into speakers' attempts to structure and organise their speech, they should be regarded as Discourse Relational Devices.

Substructural elements are not considered complete semantic or syntactic units (Herrero, 1995). Therefore, despite the fact that their linguistic study should be envisaged from the theoretical field of oral language syntax (Sornicola, 1981), they have received scarce attention from syntax studies (Pose, 2011) and are considered conversational residues (Briz and Grupo Val.Es.Co., 2014). From an interactional perspective within the framework of Conversational Analysis, some studies have pointed out that fragmentary segments like repetitions (Tannen, 1987) or self-corrections (Schegloff, 1979), even in conjunction with other multimodal resources such as gestures and gazes (Goodwin, 2007) manifest the relation between discourse planning and the interactional structure of conversation. These elements should be studied, according to Schegloff (1979: 381), from the perspective of a “syntax-for-conversation”.

This paper departs from the idea that the linguistic description of substructural elements can contribute to a better understanding of colloquial syntax by shedding light on the processes involved in the linguistic construction and organisation of oral discourse (López Serena, 2012). Given the lack of a systematic description and paradigm of substructural devices that take into account the close relationship between interaction and linguistic structure, this study proposes a definition and classification of substructural segments that contemplates formal and functional linguistic criteria (morphosyntactic, semantic, prosodic and pragmatic) and integrates the interactional basis of Conversational Analysis.

More than 500 interventions of colloquial conversations from the Corpus Val.Es.Co. 2.0. (Cabedo and Pons, 2013) have been analyzed and a database containing quantitative and qualitative variables has been created, according to the structural and interactional parameters studied (e.g. unit type, morphosyntactic characterisation, function). Statistical proofs such as Pearson's Chi-Squared test have been applied to the data with the objective of determining significant relations between functional and formal patterns in the apparition of substructural segments in Spanish colloquial conversation.

The results of this study offer a typology and description of substructural elements and contributes to: (1) shed light on the processes of construction and planning of colloquial discourse; (2) implement the characterisation of colloquial conversation as a discursive genre; and (3) advocate the importance of analysing substructural

elements, which are currently considered mere syntactical residues, and the importance of syntax of spoken discourse as a field of study.

## References

- Beinhauer, W. ([1929]1963): *El Español coloquial*, Madrid, Gredos.
- Briz, A. (1998): *El español coloquial en la conversación. Esbozo de pragmatogramática*, Barcelona, Ariel.
- Briz, A. y Grupo Val.Es.Co. (2014): "Las unidades del discurso oral. La propuesta Val.Es.Co. de segmentación de la conversación (coloquial)", *Estudios de lingüística del español*, 35, 1, 11-17.
- Cabedo, A. y Pons, S. (eds.) (2013): *Corpus Val.Es.Co. 2.0*. Disponible en: <http://valesco.es/?q=es/corpus>.
- Goodwin, C. (2007): "Human Sociality as Mutual Orientation in a Rich Interactive Environment: Multimodal Utterances and Pointing in Aphasia", Enfield, N. y Levinson, S. (eds.), *Roots of Human Sociality. Culture, Cognition and Interaction*, Oxford/New York, Berg, 97-125.
- Herrero, G. (1995): "Sobre construcciones fragmentadas", *Philologica Hispalensis*, 10, 99-114.
- López Serena, A. (2012): "En los márgenes de la estructura informativa: La posición inicial de enunciado como 'base'", *Lingüística Española Actual*, XXXIV, 2, 303-337.
- Narbona, A. (1991): "Sintaxis coloquial y análisis del discurso", *Revista Española de Lingüística*, 21, 2, 187-204.
- Pérez, M. (2003): "Sobre algunas estructuras truncadas en la conversación coloquial", *Interlingüística*, 4, 875-886.
- Pose, F. (2011): *Categorización de los cortes sintácticos en la entrevista semidirigida. Aproximación sociolingüística (Estudio de los materiales del corpus PRESEEA-Granada)*, Granada, Universidad de Granada.
- Schegloff, E. (1979): "The relevance of repair to syntax-for-conversation", en T. Givón (ed.), *Syntax and Semantics: Discourse and syntax*, 12, New York, Academic Press, 261-286.
- Sornicola, R. (1981): *Sul parlato*, Bologna, Il Mulino.
- Tannen, D. (1987): "Repetition in conversation: Toward a Poetics of Talk", *Language*, 63, 3, 574-605.

## Marcadores y partículas discursivas interactivas en Twitter

María Soledad Padilla Herrada

### Abstract

La irrupción de las redes sociales ha dado lugar a un espacio de interactividad que ha supuesto un antes y un después en el modo en que nos comunicamos. Se ha abierto, de este modo, todo un universo de posibilidades comunicativas que nos permite estar en contacto constante con otras personas y con la actualidad diaria. Concretamente, la red social Twitter proporciona a usuarios anónimos la oportunidad de comunicarse con personajes famosos, políticos o periodistas, ya que es un medio de comunicación bidireccional, en el que los usuarios reciben y aportan información.

El objetivo de esta comunicación consiste en ofrecer una clasificación y un análisis de los marcadores discursivos empleados en la red social Twitter. Se prestará una especial atención a las partículas discursivas que favorecen la interactividad en el medio, tales como ¿eh?, ¿no?, ¿verdad?, entre otras. Para ello, tomamos como objeto de estudio los perfiles en Twitter de los periodistas Jordi Évole, Carlos Herrera y Ana Pastor, ya que publican con asiduidad en esta red social y gozan de una gran cantidad de seguidores en la misma. También se analizarán los perfiles de Pablo Iglesias, Mariano Rajoy y Pedro Sánchez, ya que son los dirigentes de las opciones políticas más relevantes en España actualmente. Del mismo modo, tendremos en cuenta los tuits de usuarios anónimos dirigidos a estos personajes. Se pretende así mostrar cómo en los tuits de usuarios anónimos encontramos ciertos elementos que se utilizan para provocar al receptor. En dichos tuits, el grado de descortesía verbal, así como el grado de coloquialidad es mayor, precisamente por el carácter anónimo de los tuiteros. En cambio, se dará cuenta en este trabajo de cómo los políticos y periodistas, para proteger su imagen social, utilizan los marcadores y partículas discursivas con la intención contraria, es decir, para establecer una atmósfera de complicidad y un sentimiento de afiliación con el receptor.

Para la realización de esta investigación parto de trabajos fundamentales sobre marcadores del discurso en español: Pons (2000); Martín Zorraquino- Portolés (1999); Portolés (1998), Fuentes (1996), (2009a); Briz (2008), entre otros. Del mismo modo, me sirvo de importantes trabajos sobre descortesía verbal: Fuentes Rodríguez (2008, 2009b, 2010, 2013...); Bravo - Briz (2004), entre otros.

La metodología seguida para realizar esta investigación es de la lingüística pragmática (Fuentes Rodríguez, 2000).

### Referencias

- Briz Gómez, A. (1998): *El español coloquial en la conversación: esbozo de pragmatología*, Barcelona, Ariel.
- y Bravo, D. (2004): *Pragmática sociocultural: estudios sobre el discurso de cortesía en español*, Barcelona, Ariel.
- y Pons, S. y J. Portolés (coords.) (2008): *Diccionario de partículas discursivas del español*. En línea, [www.dpde.es](http://www.dpde.es).
- Fuentes Rodríguez, C. (1996): *La sintaxis de los relacionantes supraoracionales*, Madrid, Arco/Libros.
- (2008): *(Des)cortesía, agresividad y violencia verbal en la sociedad actual*, Sevilla, Universidad Internacional de Andalucía.
- (2009a): *Diccionario de conectores y operadores del español*, Madrid, Arco Libros.
- (2009b): «Descortesía y agresividad bajo el anonimato: Internet», en Fuentes Rodríguez, C. y Alcaide Lara, E. (eds.): *Manifestaciones textuales de la descortesía y agresividad verbal en diversos ámbitos comunicativos*, Sevilla, Universidad Internacional de Andalucía, pp. 188-208.
- (2000): *Lingüística pragmática y análisis del discurso*, Madrid, Arco/Libros.
- (2013): «Argumentación, (des)cortesía y género en el discurso parlamentario», *Tonos. Revista electrónica de estudios filológicos*, 25. Disponible en: [https://www.um.es/tonosdigital/znum25/secciones/estudios-11c-catalina\\_fuentes\\_\(2013\\_tonos\\_25\).htm](https://www.um.es/tonosdigital/znum25/secciones/estudios-11c-catalina_fuentes_(2013_tonos_25).htm).
- Lara, T. (2012): «Twitter y sus funciones comunicativas», *Blog Tiscar.com*. Disponible en: <http://tiscar.com/2012/03/11/twitter-y-sus-funciones-comunicativas/>.
- Martín Zorraquino, M.A. y J. Portolés Lázaro (1999): «Los marcadores del discurso», en I. Bosque y V. Demonte (dirs.), *Gramática descriptiva de la lengua española*, Madrid, Espasa Calpe, Vol. III, 4051-4244.
- Pons Bordería, S. (2000): «Los conectores» en Briz, A. (2000) *¿Cómo se comenta un texto coloquial?*, Madrid, Ariel, pp. 193-218.
- Portolés, L. (1998): *Marcadores del discurso*, Barcelona, Ariel.
- Yus Ramos, F. (2001): *Ciberpragmática. El uso del lenguaje en internet*, Barcelona, Ariel.
- (2010): *Ciberpragmática 2.0. El uso del lenguaje en Internet*, Barcelona, Ariel.

## From Anaphora to Discourse Relational Devices

Bernardo E. Pérez Alvarez

### Abstract

Discourse Markers have been studied in two main routes: from a semantic to a pragmatic use, what supposes a well known type of grammaticalization; and a syntactic expansion from the sentence to a textual use, well studied in text linguistic theory as a form of cohesion. In these kind of studies less attention has been focused on a discourse semantic form of cohesion that can function as a form of DRD, namely the use of anaphora for establishing discourse relations. It is possible to establish a graded relation from anaphora to discourse marking crossing through discourse deixis as a form of semantic encapsulation, that combined with some kind of prepositions can function as a DRD.

Examples from Spanish demonstratives *esto* and *eso*, and also the relatives *lo que* and *lo cual*, can be used to explain these kind of DRD's, as shown in the following examples:

1) Dentro de ese manejo están previstos los productos químicos, así como el material que utilizan al momento de acolchar los cultivos, ya que el plástico sirve sólo para una siembra y después es tirado en el campo o enterrado, por lo cual en el estado se prevé la instalación de 10 centros de acopio, de los cuales seis operarán en la región, afirmó Ponce Guía. (La Jornada, 12/09/2006)

2) se creó la auditoría superior de Michoacán donde se creó a partir de lo que era la contaduría mayor de hacienda que también dependía del congreso del estado(.) y hasta ahora no nos ha dado resultados satisfactorios no hemos tenido los resultados que quisiéramos no ha habido los resultados que los ciudadanos esperan por eso nosotros nos proponemos en esta materia que se cree un tribunal de rendición de cuentas (02RM17-10-07\_DEBATE).

The examples used in this research come from newspapers and from an oral corpus named Corpus Michoacano del Español, that includes different kind of discourses representative from speech situations that goes from spontaneous discourse to formal one. The results will show how these kind of DRD function in different kinds of speech situations and are of a dynamic nature, not already grammaticalized, what means that they should be studied in their functional nature and not only as specific units. These kind of DRD are good examples of another kind of extension from the syntactic sentence structure to the textual one, operating in both pragmatic and syntactic levels, similar to the comment clauses studied in English and lately Spanish.

### References

- Blakemore, Diane (1998): *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.
- Bühler, Karl (1999 [1934]): *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Stuttgart / New York : Fischer (Uni-Taschenbücher 1159).
- Chafe, Wallace (1994): *Discourse, Consciousness and Time. The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: The University of Chicago Press.
- Diessel, Holger (1999) : *Demonstratives. Form, Function and Grammaticalization*. Amsterdam : John Benjamins.
- Givón, Talmy (2001): *Syntax. Volume II*. Amsterdam/ Philadelphia: John Benjamins.
- Fox, B. (2008). "Dynamics of Discourse" en G. Antos, Gerd & E. Ventola (eds.). *Handbook of Interpersonal Communication*, Berlin/ New York, Mouton de Gruyter. 255-284.
- Fuentes Rodríguez, C. (2013). "La gramática discursiva: niveles, unidades y planos de análisis", *Cuadernos AISPI 2* (2013): 15-36.
- Halliday, M.A.K. & Christian M.I.M. MATTHIESSEN (2004): *An Introduction to Functional Grammar*. London: Hodder Arnold.

- Hopper, P. J. (2012). "Emergent Grammar" en Gee & Handford (eds.). The Routledge Handbook of Discourse Analysis, London/ New York, Routledge. 301-314.
- Lambrecht, Knud (1994): Information structure and sentence form. Topic, focus and the mental representations of discourse referents. Cambridge: Cambridge University Press.
- Levinson, Stephen C. (2004): "Deixis". En Laurence R. Horn y Gregory L. Ward, The Handbook of Pragmatics. Oxford: Blackwell, 97-121.
- Portolés, José (2001 [1998]): Marcadores del discurso. (2a ed.) Barcelona: Ariel.
- Raible, Wolfgang (1992): Junktion. Eine Dimension der Sprache und ihre Realisierungsformen zwischen Aggregation und Integration. Heidelberg: Carl Winter.
- Redder, Angelika (2000): "Textdeixis". En Brinker y otros, Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung. Vol. 1. Berlin/ New York: Walter de Gruyter, 283-294.
- Schiffrin, Deborah (1996): Discourse Markers. Cambridge/ New York: Cambridge University Press.

## Connective movement (A demonstration on Czech)

Lucie Poláková

### Abstract

Discourse connectives, the core of which includes conjunctions, particles and adverbs, are typically placed within one of the two discourse segments (discourse arguments) they connect. Sometimes, the first argument in the text order can be non-adjacent (Webber et al. 2003). Less often, the connective can be found outside both the arguments it connects, consider the following example:

Ministerstvo vnitra považuje před dnešním jednáním o charakteristice rozpočtu na příští rok za předčasné poskytovat detailní informace, sdělila tisková mluvčí ministerstva.

Řekla nicméně také, že i ministerstvo vnitra počítá v porovnání s letošním rokem s posílením investic.

[Before today's negotiations on the budget profile for the next year, The Ministry of the Interior considers it premature to provide detailed information, the ministry spokeswoman announced.

She also said, however, that even the Interior Ministry expects strengthening of investments compared to this year.]

[Lit: She\_said however also]

In the second sentence of the example, the main clause with the verb of saying "Řekla nicméně" /"She also said, however" contains two connectives, one of which (také/also) relates the two main clauses with a conjunction relation, whereas the other, the connective nicméně (however) can only be interpreted "lower" in the structure, as a contrast between the two subordinate clauses (reported content), not between the clauses introduced by the two verbs of saying (sdělila – řekla/announced – said).

The situation can be also explained schematically as:

Main clauses plan:

She said A. – She also said B.

\*She said A. – But she said B.

Subordinate clauses plan:

The ministry does not want to reveal any information too early. However, it reveals expectations on strengthening of investments.

We call this phenomenon connective movement: A (usually contrastive) connective is placed outside of both of its arguments, it syntactically belongs to the matrix clause but, from the semantic viewpoint, it is interpreted in the dependent clause.



We study when and why it happens and when not, arriving to the observation, that, at least in Czech, the phenomenon is not only regular, but also more preferred than the option to put the connective into the clause where it syntactically belongs:

Řekla také, že i ministerstvo vnitra nicméně počítá v porovnání s letošním rokem s posílením investic.

[She also said that even the Interior Ministry however expects strengthening of investments compared to this year.]

We document our findings by the evidence from a syntax- and discourse-annotated Prague Dependency Treebank 3.0 (Bejček et al. 2013) and mainly from the much larger Czech National Corpus (2.7 billion tokens, Hnátková et al. 2014).

In the full version of the paper, we name two testing criteria for identification of the scopes (arguments) of the connective in question, we elaborate on the nature of the excluded main clauses (mainly the so-called clauses of attribution) and explain that the connective in these structures moves left towards the border of its first argument in the linear order.

## References

- Bejček, E., Hajičová, E., Hajič et al. (2013). Prague Dependency Treebank 3.0. Data/software, Charles University in Prague, Czech Republic.
- Dinesh, N., Lee, A., Miltsakaki, E., et al. (2005). Attribution and the (Non)-Alignment of Syntactic and Discourse Arguments of Connectives. Proceedings of the ACL Workshop on Frontiers in Corpus Annotation II: Pie in the Sky. Ann Arbor, Michigan.
- Hnátková, M., Křen, M., Procházka, P., Skoumalová, H. (2014). The SYN-series corpora of written Czech. In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), pp. 160–164. Reykjavik, Iceland.
- Prasad et al. 2008. The Penn Discourse Treebank 2.0. In Proceedings of the 6th International Conference on Language Resources and Evaluation, (LREC'08), Marrakech, Morocco.
- Rysová M. 2014: Verbs of Saying with a Textual Connecting Function in the Prague Discourse Treebank. In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), pp. 930–935. Reykjavik, Iceland.
- Webber, B., Joshi, A., Stone M., and Knott A. (2003). Anaphora and Discourse Structure Computational Linguistics 29(4). pp. 545–587.

## The Combination Of Discourse Markers And Discourse Units: Not A Morganatic Marriage

Salvador Pons

### Abstract

The combination of discourse markers is perhaps the only point in which research on DMs has not still produced significant contributions. It is certainly surprising that, despite innumerable studies on this subject, a basic feature of DMs – its combination – is still to be established. One of the reasons for this absence is the lack of a commonly accepted notion of position in spoken language. To know what position means depends, in turn, on defining a set of discourse units where the notion of position can take place. In the absence both notions, it seems impossible to provide an answer to the combination issue:

First, in the research of left-and right peripheries (Rizzi 1997, Dufter and Octavio de Toledo 2014), a DM is at left/right periphery of a sentence-like maximal projection. Therefore, the notion of sentence/utterance is still crucial, even if Dms are unanimously claimed to be sentence-external elements (Schourup 1999, Fischer 2006). Besides utterance, no other units are considered (except for references like sentence-internal and sentence-external positions).

Second, positions are not defined and their meaning is established on non well-established bases. For instance, initial position can be interpreted as “first word in the utterance”, “first syntagmatic projection in the utterance”, “first tone unit”, and so on.

Third, positions and units do not combine and, as a consequence, the possibility that the left periphery of a turn should host functions different from the left periphery of an utterance is not even taken into account.

This paper attempts to overcome the problems above and to establish a basis for the combination of discourse markers. The analysis presented here will be based on Spanish colloquial conversations and will make use of the Val.Es.Co model of discourse units (Briz et al. 2003, Val.Es.Co Research Group 2014). The need of a theory of discourse units will be stressed by comparing it to alternative approaches to this same subject.

### References

- Briz, Antonio et al. (2003). “Un sistema de unidades para el estudio del lenguaje coloquial”. *Oralia* 6: 7-61.
- Cuenca, M.J. and M. J. Marin (2009): “Co-occurrence of discourse markers in Catalan and Spanish oral narrative”. *Journal of Pragmatics* 41, 5, 899-914.
- Dufter, Andreas and Álvaro S. Octavio de Toledo (eds.) (2014). *Left Sentence Peripheries in Spanish: Diachronic, Variationist and Comparative Perspectives*.
- Fischer, Kerstin (2006): *Approaches to Discourse Particles*. *Studies in Pragmatics*, 1. Elsevier, Amsterdam.
- Rizzi, L. (1997) “The Fine Structure of the Left Periphery”, in L. Haegeman (ed.) *Elements of Grammar*, Kluwer, Dordrecht, 281-337.
- Schourup, Lawrence (1999). “Discourse markers”. *Lingua*, 107, 227-265.
- Val.Es.Co. Research Group (2014): *Las unidades del discurso oral. La propuesta Val.Es.Co. de segmentación de la conversación (coloquial)*. *Estudios de lingüística del español* 35, 11-71.
- Verdonik, Darinka; Rojc, Matej and Stabej Marko (2007). “Annotating discourse markers in spontaneous speech corpora on an example for the Slovenian language”. *Language resources and evaluation* 41, 1, 147-180.

## "Razón de más" como inversor argumentativo

José Portolés

Abstract

Del siguiente intercambio:

(1) —La prohibición genera el tráfico de drogas.

—Razón de más para mantenerla. (en [www.elpais.com](http://www.elpais.com), 31/03/2014)

se interpreta que, para el segundo participante, contrariamente a lo que se pudiera concluir a partir de la primera intervención, el hecho de que la prohibición de las drogas genere su tráfico no es una razón para que se levante, sino precisamente para lo contrario, para que se mantenga.

Este será el uso del sintagma "razón de más" que centrará el presente estudio: un turno conversacional reactivo introducido por razón de más que crea e invierte una dinámica argumentativa previa. Con este propósito, se analizan las propiedades de los constituyentes de "razón de más". El sustantivo "razón" constituye una etiqueta discursiva (§ 2), que encapsula lo dicho previamente en un único sustantivo y, por su significado, categoriza el miembro del discurso previo como una razón. En cuanto a de más como cuantificador (§ 3), convoca una escala que se ordena de acuerdo con un criterio, que, como con otros cuantificadores de grado, generalmente se explicita con un complemento con un sintagma preposicional introducido por "para".

Aunque "razón de más" puede utilizarse en el español actual como un sintagma en combinación libre, en los usos que nos ocupan se ha producido un proceso de lexicalización en el que han coadyuvado una posible evolución del español al tiempo que el calco decimonónico de *raison de plus* del francés (§ 4).

(2) Dorotea.- ¿Os sonrojaría?

Doña Florinda.- No a mí, pero quiero que tenga que sonrojarle a él.

Dorotea.- Razón de más para ocultarlo. (Mariano José de Larra, Traducción de Don Juan de Austria o la vocación, de Delavigne, España, CNDHE, 1835)

(3) Dorothee.— Est-ce que vous en rougissez ?

Dona Florinde.— Non assurément ; mais je ne veux pas qu'il en rougisse, lui.

Dorothee.— Raison de plus pour le cacher. (Casimir Delavigne, Œuvres Complètes, II, Didier, París, 1863, pág. 455 gallica.bnf.fr).

Por otra parte, el sintagma introducido por "razón de más" tiene una gramática peculiar y su uso solo se explica como parte de una construcción conversacional que identifican los hablantes (§ 5). Se trata de un sintagma cuantificado que, sin verbo principal, se utiliza como enunciado después de un miembro discursivo previo que se pueda interpretar como aserción. En una de estas construcciones con características particulares, se logra su interpretación como inversor argumentativo (§6). Este fenómeno fue destacado por Oswald Ducrot (Bruxelles y otros 1982, Ducrot 1988) en algunos usos del francés *justement*. En el ejemplo (1) al utilizar razón de más el segundo participante hace inferir que el primero sigue el topos: "Aquello que genere el tráfico de drogas debe

ser eliminado". Este segundo participante, por su parte, invierte este topos y lo convierte en: "Aquello que genere el tráfico de drogas debe ser mantenido" y, por este motivo, habría que persistir en la prohibición —la conclusión—.

## Referencias

- BRUXELLES, SYLVIE Y OTROS (1982) "Justement, l'inversion argumentative", *Lexique*, 1, págs. 151-164.
- DUCROT, OSWALD (1988) *Polifonía y argumentación*, Cali, Universidad del Valle.
- ELVIRA, JAVIER (2015) *Lingüística histórica y cambio gramatical*, Madrid, Síntesis.
- GRAS, PEDRO (2010) *Gramática de construcciones en interacción. Propuesta de un modelo y aplicación al análisis de estructuras independientes con marcas de subordinación en español*, Barcelona, Universitat de Barcelona <<http://goo.gl/BrDwRA>>.
- LÓPEZ SAMANIEGO, ANNA (2014) *Las etiquetas discursivas: cohesión anafórica y categorización de entidades del discurso*, Pamplona, Eunsa.
- MONTOLÍO, ESTRELLA (1999) "¡Si nunca he dicho que estuviera enamorada del él! Sobre construcciones independientes introducidas por si con valor replicativo", *Oralia* 2, págs. 37-70.
- NØLKE, HENNING, KJERSTI FLØTTUM y COCO NØREN (2004): *ScaPoLine. La théorie scandinave de la polyphonie linguistique*. Paris: Kimé.
- PONS, SALVADOR (2008) "Gramaticalización por tradiciones discursivas: el caso de esto es", en Johannes Kabatek (ed.) *Sintaxis histórica del español y cambio lingüístico: Nuevas perspectivas desde las Tradiciones Discursivas*. Madrid/Frankfurt, Iberoamericana/Vervuert, págs. 249-274.
- SÁNCHEZ ILLÁN, JUAN CARLOS (2001) "La edición de periódicos y la empresa periodística", en Jesús A. Martínez Martín (dir.) *Historia de la edición en España (1836-1936)*, Madrid, Marcial Pons, págs. 397-414.

## When Semantics and Pragmatics diverge: the translation of adverbial connectors

Beatriz Sánchez-Cárdenas

### Abstract

In teaching specialized translation, one of the challenges is to help students perceive the text as a whole. This difficulty is directly related to the way that each language-culture tends to structure scientific discourse, as reflected in text types and their information structure. Interlinguistic and intertextual variation directly affects the use of logical connectors, syntax, and semantic prosody. In this sense, general language words can be a challenge for the translator of specialized texts since their behavior in general language texts differs from their behavior in specialized language texts. For example, in bilingual Spanish-English dictionaries, *however/sin embargo*, *currently/ actualmente* and *inadequately/inadecuadamente*, are generally regarded as translation correspondences. Nevertheless, in specialized texts, this equivalence is more apparent than real because of the specific contextual constraints imposed by each of these words, which can vary, largely due to their semantic prosody.

The main objective of this study was to assess the effectiveness of using comparable corpora for teaching specialized translation to undergraduate students with a view to increasing the quality of their text production and their translation competence. For this purpose, we performed an experiment in which undergraduate students were asked to translate a set of English medical text excerpts into Spanish, before and after receiving training sessions during which they learned how to compile, analyze, and exploit comparable specialized corpora. The focus was on the translation of adverbial modification in the form of logical connectors. In the course of this study, students became more sensitive to the problem of translating such words in specialized texts. When the translation produced before and after the training sessions were compared, the results reflected that the students had gained a heightened awareness of the difficulty of translating these apparently simple general language terms and were less apt to choose the conventional dictionary correspondence as a translation solution.

In this sense, corpus analysis reveals that general language words and their interlinguistic equivalents tend to appear in different syntagmatic contexts in each language (Cummins 2002). For instance, the English-Spanish correspondences, *however/sin embargo*, *currently/actualmente* and *inadequately/inadecuadamente* are not always pragmatic equivalents in these two languages, due to different contextual constraints (Charles 2011). Moreover, the behavior of these words in specialized texts is also different from their behavior in general language discourse.

In our opinion, it is crucial for translation students to have a heightened perception of such differences (i.e. language awareness) in order to be able to deal with the translation of specialized texts. Thus, the main objective of this study was to assess the effectiveness of using comparable corpora for teaching specialized translation to undergraduate students with a view to increasing the quality of their text production.

### References

- Charles, Maggie (2011). Corpus evidence for teaching adverbial connectors of Contrast: *however*, *yet*, *rather*, *instead* and *in contrast*. In N. Kübler. (ed) *Corpora, Language, Teaching, and Resources: From Theory to Practice*. Bern: Peter Lang, pp. 113-132.
- Cummins, Sarah (2002). A case study in lexical research for translation. *International Journal of Lexicography*, Vol. 15 No. 2, pp. 139-159.

- Gill, Philip (2011). "...and I dropped my jaw with fear": The role of corpora in teaching phraseology. In N. Kübler. (ed) *Corpora, Language, Teaching, and Resources : From Theory to Practice*. Bern: Peter Lang, pp. 155-164. pp. 49-68.
- Krause, Sylvana (2011). Semantic preference and semantic prosody in the specialist language class. In Nathalie Kübler (ed) *Corpora, Language, Teaching, and Resources : From Theory to Practice*. Bern: Peter Lang, pp. 155-164.
- Kübler, Natalie and Alexandra Volanschi (2012). Semantic prosody and specialised translation, or how a lexico-grammatical theory of language can help with specialised translation. In Alex Boulton, Shirley Carter-Thomas, and Elisabeth Rowley-Jolivet (eds) *Corpus-informed Research and Learning in ESP: Issues and Applications*. Amsterdam and Philadelphia: John Benjamins, pp. 105-135.
- Kübler, Nathalie (2003). Corpora and LSP translation. In Zanettin, F., S. Bernardini & D. Stewart (eds), *Corpora in Translator Education*. Manchester: St Jerome, pp. 47-60.
- López Rodríguez, Clara Inés, and Miriam Buendía (2013). Aplicación de la lingüística de corpus en la didáctica de la traducción científica y técnica. In Emili Casanova Herrero, and Cesáreo Calvo Rigual (eds) *Proceedings of the XXVI Congreso Internacional de Lingüística i Filología Romaniques*, Vol. VIII, Berlin: Walter de Gruyter, pp. 205-216.
- Gill, Philip (2009). Arriving at equivalence: Making a case for comparable general reference corpora in translation studies. In Beeby, Allison, Patricia Rodríguez Inés and Pilar Sánchez-Gijón (eds.), *Corpus Use and Translating: Corpus use for learning to translate and learning corpus use to translate*, pp. 59-63.
- Stewart, Dominic (2009). Safeguarding the lexicogrammatical environment: Translating semantic prosody. In Beeby, Allison, Patricia Rodríguez Inés and Pilar Sánchez-Gijón (eds.), *Corpus Use and Translating: Corpus use for learning to translate and learning corpus use to translate*, Amsterdam and Philadelphia: John Benjamins, pp. 29-46.

## The use of a Spanish Lexicon of Subjectivity for the analysis of Discourse Relational Devices (DRDs) in Spanish academic texts

Andrea Santana Covarrubias

Discourse Relational Devices (DRDs) have been described as explicit processing instructions on how one part of a text is related to another (Britton, 1994; Gernsbacher, 1997; Sanders & Spooren, 2007). Within the Cognitive approach of Coherence Relations (CCR), several studies have demonstrated that specific DRDs express purely objective or subjective meanings (Pander Maat & Sanders, 2001; Degand & Pander Maat, 2003; Li, Evers-Vermeul and Sanders, 2013). However, the intended interpretation of these meanings is not always signaled clearly by the DRDs. There are ambiguous cases, in which some DRDs can be used in both subjective and objective relations (e.g. because). These cases make the analysis of DRDs and coherence relations a complex task.

Moreover, by analyzing relations in context, other factors could influence the interpretation and the predominance of DRDs. Among them, genre could play a relevant role. Sanders (1997) demonstrated that ambiguous cases of subjective and objective causal relations were resolved according to the type of context. Zufferey (2012) also showed that the distribution of some French connectives varies according to different genres. In this investigation our purpose is to explore the relationship between DRDs and genre in Spanish academic discourse in two disciplines, Education and Psychology using automatic analysis. For this, we constructed a corpus of academic texts written in Spanish (essays, 81,640 words; textbooks, 82,319 words; and research articles, 82,734 words) and we identified frequent DRDs for causal relations through concordance analyses. In order to identify if the prototypical DRDs are associated with subjective or objective meanings, we determined degrees of subjectivity in the corpus by comparing it to a Spanish lexicon of subjectivity (Molina-González, Martínez-Cámara, Martín-Valdivia & Perea-Ortega, 2013): the expectation is that if a DRD is used subjectively it occurs in a subjective environment, i.e. a context containing relatively many subjective words.

We hypothesize that there is a continuum of subjectivity in terms of DRDs and coherence relations in academic discourse. On one extreme, the essay is expected to contain the most subjective words since it persuades the reader of the correctness of a central statement (Hyland, 1990), therefore, we could expect that the DRDs in essays are more associated with subjective relations. On the opposite extreme, the textbook is expected to contain the smallest amount of subjective words since it is an introductory genre which describes the accepted knowledge of a discipline as a coherent whole (Swales, 1995), consequently, we could expect that the DRDs are more associated with objective relations. Finally, the research article is in the middle of this continuum since it is a hybrid genre which is persuasive in some sections and descriptive in others. Therefore, we would expect that the DRDs are associated with both subjective and objective relations. This study could give us more information about the usefulness of lexicons of subjectivity to explore the subjectivity in genres, which would constitute an initial strategy to analyze DRDs and coherence relations in languages and contexts few explored from a cognitive perspective, such as the Spanish and the academic discourse.

### References

- Britton, B. (1994). Understanding expository text: Building mental structures to induce insights. In A. Morton Gernsbacher, (Ed.) *Handbook of Psycholinguistics* (640–674). San Diego: Academic Press.
- Degand, L. & Pander Maat, H. (2003). A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale. In A. Verhagen & J. van de Weijer (Eds.), *Usage based approaches to Dutch* (175-199). Utrecht: LOT.
- Gernsbacher, M. (1997). Coherence cues mapping during comprehension. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships. Studies in the production and comprehension of text* (3-21). New Jersey: Lawrence Erlbaum Associates.
- Hyland, K. (1990). A Genre Description of the Argumentative Essay. *RELC Journal*, 21, 66-78.

- Li, F., Evers-Vermeul J. & Sanders, T. (2013). Subjectivity and result marking in Mandarin: A corpus-based investigation. *Chinese Language and Discourse*, 4(1): 74-119.
- Molina- González, M., Martínez-Cámara, E. Martín-Valdivia, M. & Perea-Ortega, J. (2013). Semantic orientation for polarity classification in Spanish reviews. *Expert System with Applications*, 40, 7250-7257.
- Pander, Maat, H. & Sanders, T. (2001). Subjectivity in causal connectives: An empirical study of language in use. *Cognitive Linguistics*, 12(3): 247-273.
- Sanders, T. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes*, 24, 119-147.
- Sanders, T., & Spooren, W. (2007). Discourse and text structure. In D. Geeraerts & H. Cuykens (Eds.), *Handbook of cognitive linguistics* (37-60). Oxford: Oxford University Press.
- Swales, J. (1995). The role of textbook in EAP writing research. *English for Specific Purposes*, 14, 1, 3-18.
- Zufferey, S. (2012). "Car, parce que, puisque" revisited: Three empirical studies on French causal connectives. *Journal of Pragmatics*, 44 (2), 138-153.



## “On the one hand” as a Cue to Anticipate Upcoming Discourse Structure

Merel Scholman; Hannah Rohde and Vera Demberg

Given evidence of anticipation within sentences for upcoming sounds, words, and syntactic structures (DeLong, et al. 2005; Kamide, et al., 2003; Levy, 2008), an open question is how comprehenders use cross-sentence cues to anticipate relationships between sentences. Within sentences, words combine via syntactic rules to determine what structures are possible. Between sentences, the possible relationships that can hold between pairs of propositions, such as *cause-consequence* and *claim-argument*, create a less constrained discourse structure (Asher & Lascarides, 2003; Hobbs, 1990; Kehler, 2002). Based on evidence of syntactic prediction (e.g., dependencies like *either...or*, Staub, 2006), our goal is to test what comprehenders predict based on the marker *On the one hand*.

Three hypotheses were formulated for this study:

- H1. Readers expect *On the one hand* to be followed by any type of contrast marked with *On the other hand* specifically;
- H2. Readers have contrast-specific expectations, where only content that contrasts with the *On the one hand*-clause will satisfy their expectations (regardless of the connective used);
- H3. Readers have a “flat” prediction of contrast, where any contrastive clause can satisfy the prediction set up by *On the one hand* (regardless of the connective used).

Stories containing one of three types of intervening sentences (non-contrastive (example 1a), globally contrastive (1b), or locally contrastive (1c)) between *On the one hand* and *On the other hand* were tested in a story acceptability study, story continuation study, and eye-tracking study.

The results show that *On the one hand* does not have to be followed by *On the other hand*; the connective *but* can also satisfy the anticipation for a contrast, especially if it is used in a globally contrastive sentence. This is not compatible with H1. In particular, reading times on *On the other hand* were longer in stories with a globally contrastive sentence than in stories with a locally contrastive sentence. The results support H2 and disconfirm H3: readers can build structure-specific expectations based on *On the one hand*. We conclude that comprehenders use discourse connectors to predict a specific discourse structure and can maintain such predictions across clauses.

## Example and illustration

- (1) Sentence A: Joseph is pondering whether he should take a job offer from the Edinburgh Zoo.
- Sentence B: On the one hand, he needs the money, because he should start paying off his student loans this year.
- Sentence C:
- Also, his car needs to be serviced by the end of the month. [no contrast]
  - But he could keep looking for a nicer, better-paying job. [global contrast]
  - But the loans could be deferred for a few more months. [local contrast]
- Sentence D: On the other hand, he hates the idea of cleaning out panda cages every day.

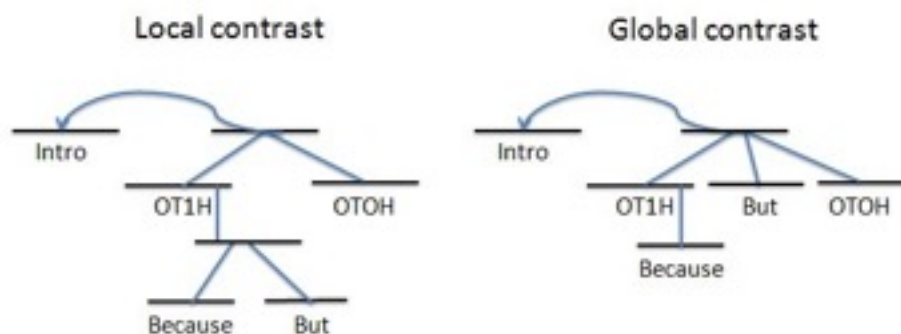


Figure 1. A

## References

- Asher, N., & Lascarides, A. (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.
- DeLong, K. A., Urbach, T. A., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8, 1117-1121.
- Hobbs, J. R. (1990). *Literature and cognition*. Stanford, CA: CSLI. Lecture Notes 21.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49, 33-156.
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126-1177.
- Staub, A., & Clifton Jr, C. (2006). Syntactic prediction in language comprehension: Evidence from either... or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 425.

## Las conjunciones ilativas así que y conque como marcadores de modalidad evidencial

Eugenia Sainz

### Abstract

No resulta fácil dar cuenta de la diferencia de significado y de uso existente entre las conjunciones ilativas así que y conque. Pese a tratarse de un tema recurrente en la gramática del español, no se ha llegado todavía a una descripción contrastiva clara desde la que justificar las diferencias en la codificación semántica y, por ende, en la intención que mueve la enunciación en cada caso: por qué los hablantes escogen a veces la primera y por qué, otras, en cambio, prefieren la segunda. De hecho, es significativo advertir que la Nueva Gramática de la lengua española (2009) se limita a aludir, como es habitual, a los valores modales vinculados con los enunciados introducidos por conque (§46.12a, b, l) y a reconocer la proximidad de ambas unidades con “las unidades características de la gramática del discurso o del texto” (§46.11h).

Pues bien, con el objetivo de responder al interrogante planteado, se analizarán las dos unidades en su empleo como marcadores discursivos, es decir, cuando introducen directamente un acto de habla en ausencia de un primer miembro expreso, y en relación con las modalidades epistémica y evidencial. La intervención se estructurará en tres partes: en primer lugar, se ponen las bases para el análisis siguiendo de cerca las investigaciones sobre evidencialidad (v.g., Aikhenvald 2004, Corneille 2009, Cuenca en prensa, Ifantidou 2001, Nuyts 2001, Plugian 2001, Wachtmeister 2005). En segundo lugar, se focaliza la atención en el funcionamiento discursivo de así que. Para dar cuenta del modo en que el significado procedimental de la partícula interactúa con el contexto, se compara así que con el que enunciativo y con las demás formas de cita indirecta (Demonte y Fernández Soriano 2007, 2013a, 2013b, Etxepare 2010, Gras 2013, Pons Bordería 2003, 2013, Porroche 2000, 2003, Reyes 1994, Rodríguez Ramalle 2008a, 2008b, 2013, 2014, 2015). Por último, se afronta el análisis discursivo de conque desde una hipótesis de significado procedimental que creemos puede justificar el contraste con así que y la carga modal añadida (Porroche 2009, Portolés 1998, Santos 2003). Se tomarán las muestras de lengua del banco de datos CREA.

### Referencias

- Aikhenvald, Alexandra Y. (2004): *Evidentiality*. Oxford: Oxford University Press.
- Cuenca, María Josep (en prensa): “Evidentiality (and epistemicity) in Catalan parliamentary debate”, *Scripta*
- Corneille, Bert (2009): “Evidentiality and epistemic modality: On the close relationship between two different categories”. *Functions of Language*, 16/1, 44-62.
- Demonte, Violeta y Fernández Soriano, Olga (2013a): “Evidentials ‘dizque’ and ‘que’ in Spanish. Grammaticalization, parameters and the (fine) structure of Comp”, *Linguística. Revista de Estudos linguísticos da Universidade do Porto*, vol. 8 - Universidade do Porto.
- Demonte, Violeta y Olga Fernández-Soriano (2013b): “El ‘que’ citativo, otros ‘que’ de la periferia izquierda oracional y la recomplementación”, en D. Jacob y K. Ploog (eds.): *Autour de ‘que’*. El entorno de que, vol. 0, Peter Lang.
- Demonte, Violeta y Olga Fernández-Soriano (2007): “La periferia izquierda oracional y los complementantes del español”, en Juan Cuartero y Martine Emsel (eds.): *Vernetzungen: Kognition, Bedeutung (kontrastive) Pragmatik*, Frankfurt: Peter Lang.
- Etxepare (2010): “From hearsay evidentiality to samesaying relations”, *Lingua* 120, 604-627.
- Gras, Pedro (2013): “Entre la gramática y el discurso: valores conectivos de que inicial átono en español”. In Daniel Jacob & Katja Ploog (eds.). *Autour de que*. El entorno de que. *Studia Romanica et Linguistica*, vol. 37, 89-112. Frankfurt am Main: Peter Lang.
- Ifantidou, Elly (2001) *Evidentials and Relevance*. Amsterdam/Philadelphia: John Benjamins.
- Nuyts, Jan (2001): “Subjectivity as an evidential dimension in epistemic modal expressions”, *Journal of Pragmatics* 33: 383-400.
- Plugian, V. (2001). “The place of evidentiality within the universal grammatical space.” *Journal of Pragmatics* 33, 349-357.
- Pons Bordería, Salvador (2003): “Que inicial átono como marca de modalidad”, *ELUA*, 17: págs. 531-545.
- Pons Bordería, salvador (2013): “Un solo tipo de reformulación”, *Cuadernos AISPI* 2, 151-170.
- Porroche Ballesteros, Margarita (2000), “Algunos aspectos del uso de que en el español conversacional (que como introductor de oraciones “independientes”)”, *CLAC*, 3, 100-116.
- Porroche Ballesteros, Margarita (2003), “Aspectos discursivos en el uso de algunas conjunciones españolas”, *Oralia*, 6, 259-282.

Portolés Lázaro, José (1998): "El concepto de suficiencia argumentativa", en *Signo y seña*, 9, 201-224.

Reyes, Graciela (1996 [1994]): *Los procedimientos de cita: citas encubiertas y ecos*, Madrid, Arco/Libros.

Rodríguez Ramalle, María Teresa (2013): "Notas para un estudio comparado de la evidencialidad", *RSEL*, 43/1, págs. 169-192.

Rodríguez Ramalle, María Teresa (2014): "Sobre marcadores y su relación con la modalidad evidencial", en Marta García Negroni (ed.): *Marcadores del discurso. Perspectivas y contrastes*, Buenos Aires, Santiago Arcos Editor. Colección Instrumentos.

Rodríguez Ramalle, María Teresa (2015): "Información contextual, estructura discursiva y modalidad evidencial en la selección de las locuciones consecutivas del español", *Spanish in Context* 12:1, págs. 80-101.

Wachtmeister Bermúdez, Fernando (2005): *Evidencialidad: La codificación lingüística del punto de vista / Evidentiality: The linguistic coding of point-of-view*. Doctoral dissertation

Department of Spanish, Portuguese and Latin American Studies Stockholm University

## Discourse markers in the Val.Es.Co. segmentation model: functions and uses of "pues eso" in Spanish colloquial conversations

Shima Salameh Jiménez

### Abstract

Discourse markers (henceforth, DMs) allow to formulate, organize and manage contents within discourse structure. As discourse relational devices (DRDs) these units (studied along the last thirty years in several languages and applying different perspectives) show different pragmatic relationships: between linguistic contents (connection), building and structuring textual-oral productions (Pons Bordería 2006); between the speaker and his own message (subjectivity), by making explicit processes of formulation or modalization; and, at last, between the speaker and his hearer (intersubjectivity), by highlighting their interactive relationship through appellative devices, turn-taking/giving mechanisms, etc.

Specifically, our study object is the Spanish conversational construction "pues eso": a polyfunctional DM that works in both monological and dialogical levels, that is, in enunciation and interaction contexts, respectively. It has an abstract lexical basis (formed by a comment adverb/conjunction—"pues"— and a demonstrative neutral pronoun—"eso"—) that encourages uses in diverse discourse positions with variable scope. It could be possible to think that "pues eso" is only a derived from the form "pues" and that, consequently, this study is an easy task to develop. However, the impossibility of separate both units in the construction, the fact that some functions are directly linked to the whole structure (and not only with pues), and the high frequencies of use that it presents in the corpora analysed (Corpus Val.Es.Co.; Corpus Val.Es.Co. 2.0; Crea oral; Corpes XXI), lead us to defend that "pues eso" has experimented a lexicalization process.

With the picture presented for this DM, and a complex description of its functions, this communication aims to systematize discursive features of "pues eso" in each type of interactive concrete context (contexto interactivo concreto, CIC), with a view to solve common treatment difficulties in DRDs. To avoid a non-objective perspective of analysis, based in descriptions without systemic criteria, we apply the tools offered by the Val.Es.Co. model (VAM) for discourse segmentation (Briz y grupo Val.Es.Co. 2003; Grupo Val.Es.Co. 2014). In its terminology, "pues eso" can be labelled as "intervention", "act", or adjacent modal subact" (AMS), "adjacent interpersonal subact" (AIS) or "adjacent textual subact" (ATS): structural units that depending of the unit over which they have scope (other subact, an act or an intervention) and the position occupied, simplify the treatment of each function of "pues eso" (topic regressive, topic conclusive, support-formulation device, agreement mark).

A description based in the Val.Es.Co. discourse units gives objectivity to the qualitative analysis of a DM whose characterization is complicated. Furthermore, permits work with quantitative results: accounting the total of acts and subacts correlated with one or more pragmatic functions, makes possible the following conclusions:

- 1) Corroboration of the DM "pues eso" polyfunctionality;
- 2) Confirmation that In Spanish is necessary to consider "pues eso" as an independent DM in respect with DM "pues".

3) Conveys a complete overall of results obtained, that can be organized in the VAM units and position grill (Briz y Pons 2010)

## Referencias

- Briz, Antonio y Grupo Val. Es. Co., (2003): "Un sistema de unidades para el estudio del lenguaje coloquial", *Oralia* 6, 7-61.
- Briz, Antonio y Pons, Salvador (2010): "Unidades, marcadores discursivos y posición". *Los estudios sobre marcadores discursivos, hoy*, Madrid, Arco/Libros.
- Briz, Antonio, Portolés, José, Pons Bordería, Salvador (Eds.), 2008. *Diccionario de partículas discursivas del español*. Cuenca, M.J. (2006): *La connexió i els connectors: perspectiva oracional i textual*. Vic, Universitat de Vic.
- Diewald, Gabriele (2006): "Discourse particles and modal particles as grammatical elements". *Approaches to Discourse Particles*. Amsterdam, Elsevier.
- Fraser, Bruce (1999): "What are discourse markers?" *Journal of Pragmatics* 31: 931-952.
- Fuentes Rodríguez, Catalina (2009): *Diccionario de conectores y operadores del español*. Madrid, Arco/Libros.
- Grupo Val.Es.Co. (2014): "Las unidades del discurso oral. La propuesta Val.Es.Co. de segmentación de la conversación (coloquial)". *Estudios de lingüística del español* 35, pp. 13-73.
- Loureda, Óscar y Esperanza Acín (coords.), (2010): *Los estudios sobre marcadores del discurso, hoy*. Madrid, Arco/Libros.
- Pons Bordería, Salvador (2006b): "A functional approach to the study of discourse markers" in Kerstin Fischer (ed.), *Approaches to Discourse Particles*.
- Pons Bordería, Salvador (ed.) (2014): *Discourse Segmentation in Romance Languages. Pragmatics & beyond new series*. John Benjamins Publishing Company.

## Unifying Dimensions in Discourse Relations: how various annotation schemes are related

Ted Sanders, Vera Demberg, Jacqueline Evers-Vermeul, Jet Hoek, Merel Scholman, Fatemeh Torabi Asr and Sandrine Zufferey

### Abstract

Over the last fifteen years, annotating discourse relations has gained increasing interest of the linguistics research community. Indeed, it is a promising and challenging research area, which allows for systematic cross-linguistic comparison at the discourse level. A lot of progress has been achieved through large discourse-annotated corpora, leading examples are the Penn Discourse Treebank (Prasad et al., 2008), the Rhetorical Structure Theory (RST) Treebank (Carlson et al., 2001) and SDRT. In fact, the annotation of such relations is at the heart of the COST-Textlink project.

However, existing discourse annotation guidelines differ in other important aspects, such as the type of relations that are distinguished. Some proposals present sets of approximately 20 relations (such as the one developed by Mann and Thompson, 1988), others of only two relations (Grosz & Sidner, 1986). The PDTB contains a three-tiered hierarchical classification of 43 sense tags (Prasad et al., 2008), and the annotation scheme used for the RST Treebank distinguishes 78 relations that can be partitioned in 16 classes (Carlson et al., 2001). Hence, it is not clear which and how many categories or classes (for example, contingency, causal, or informational) and end labels (for example, result, volitional cause, and cause-consequence are all labels for causal relations) are needed to adequately describe and distinguish coherence relations.

One thing that is clear is that annotation has proven to be a difficult task, which is regularly reflected in low inter-annotator agreement scores. Because leading proposals differ in the relations they distinguish, comparing outcomes of corpus-based studies working in different frameworks is hard; often, different labels are used for the same concepts, and the other way around: sometimes the same labels are used for different concepts. This is a confusing and non-optimal situation.

In this paper, our goal is to suggest how the discourse relation annotations used by the different schemes can be mapped onto one another. An important consideration is to be able to represent all of the annotations that the different schemes have considered relevant for discourse relation annotation. Our suggestion is to describe discourse relations in terms of their properties. A range of basic properties, based on a Cognitive approach to Coherence Relations (CCR; going back to Sanders et al. 1992), which capture the most basic distinctions made by all sets of discourse relations, will be extended by additional criteria, which capture additional distinctions made in various discourse annotation schemes. We describe the relations in terms of this limited set of dimensions and additional criteria, and show how the various existing proposals can be related to each other. That way, we develop a unifying proposal, which allows us to 'translate' outcomes from one framework to the terminology of another. We want to make optimal use of existing corpora and facilitate discussion among researchers working in different paradigmata. The ultimate goal is to obtain more systematic discourse relation schemes.

More specifically, we will compare PDTB, RST and SDRT in terms of our limited set of dimensions, and show how they map onto each other. We illustrate our proposal with annotation of corpus examples from various studies: on cross-linguistic comparison and translation, on language acquisition, from an experiment with

automatic discourse relation labelling using dimensions, and from an annotation experiment with naïve annotators using a minimal set of relation categories.

We will argue that our approach of unifying dimensions will lead to more systematicity and is likely to improve existing relation definitions.

## References

- Carlson, Lynn, Daniel Marcu and Mary E. Okurowski (2001). Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure Theory. Proceedings of the 2nd SIGDIAL Workshop on Discourse and Dialogue, Eurospeech 2001, Denmark.
- Mann, William C. and Sandra A. Thompson (1988). Rhetorical Structure Theory: Towards a functional theory of text organization. *Text*, 8(3): 243-281.
- Reese, B., Hunter, J., Asher, N., Denis, P., & Baldridge, J. (2007). Reference manual for the analysis of rhetorical structure. Unpublished manuscript. University of Texas at Austin, Austin, TX.
- Prasad, Rashmi, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi and Bonnie Webber (2008). The Penn Discourse Treebank 2.0. Proceedings of the 6th International Conference of Language Resources and Evaluation (LREC 2008), Marrakech.
- Sanders, Ted J.M., Wilbert P.M.S. Spooren and Leo G.M. Noordman (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, 15: 1-35.



## Contributos para a análise do conector "sim" em Português Europeu Contemporâneo

Sara Sousa

### Abstract

Neste trabalho, propõe-se que, em Português Europeu Contemporâneo, a unidade "sim", tipicamente integrada na categoria dos advérbios, pode igualmente ocorrer como um conector ou marcador discursivo em construções refutativo-retificativas de justaposição do tipo "não p || sim q". Esta consideração tem por base a análise de corpora, nomeadamente o Corpus de Referência do Português Contemporâneo (CRPC), CETEMPúblico e Diários da Assembleia da República (DAR), onde se verifica a ocorrência de sequências como as seguintes:

(i) Falar do primado da educação e da formação não constitui uma moda, constitui, sim, uma exigência do Estado social e uma resposta às ilusões neoliberais [...]. (CRPC, COD\_1015452)

(ii) Não pretendemos um tratamento diferenciado, pretendemos, sim, ser tratados, acompanhados com respeito, dedicação e afecto [...]. (CETEMPúblico, ext876712-pol-92b-2)

(iii) Em primeiro lugar, não estamos perante propostas que apenas concretizam os cortes salariais impostos pelo Orçamento do Estado, estamos, sim, perante propostas que impõem novas e acrescidas reduções remuneratórias aos magistrados, de forma permanente e definitiva [...] (q). (DAR, I Série, XI Legislatura, N.º 38, p. 62)

A presença de "sim" em exemplos como os anteriores permite realçar o contraste de polaridade existente entre os segmentos "não p" e "q", evidenciando que o segmento (q) em que ocorre é aquele que deve ser retido. Mais exactamente, o locutor responsável por uma sequência do tipo ilustrado em (i), (ii) ou (iii) assinala que o que é dito ou implicado em "q" se constitui como uma retificação/correção do enunciado, tipicamente da responsabilidade de uma outra voz ou instância enunciativa, alvo de refutação em "p". Este último segmento contém tipicamente uma negação de carácter metalinguístico (Ducrot, 1972; Horn, 1989).

Em construções deste tipo, "sim" parece comportar-se como um conector/marcador discursivo (de origem adverbial) que, não contribuindo para o conteúdo proposicional das sequências em que ocorre nem sendo necessário à sua boa formação, fornece instruções sobre o seu valor, facilitando o seu processamento e contribuindo, assim, para a sua coerência relacional (neste caso, na sua vertente mais estritamente pragmático-funcional). Este comportamento é evidenciado pelo facto de, nos corpora analisados, esta unidade surgir tipicamente demarcada por pausas, configurando-se como um constituinte prosódico e dando origem a produtos textuais e não fráscicos.

Procurando descrever o funcionamento discursivo desta unidade e justificar a sua inclusão no seio dos conectores/marcadores discursivos disponíveis em Português Europeu Contemporâneo, este trabalho pretende contribuir para uma melhor compreensão deste tipo de expressões conectivas, particularmente ao nível da sua classificação formal e funcional.

### Referências

- Ducrot, Oswald (1972) *Dire et ne pas dire. Principes de sémantique linguistique*. Paris: Hermann.  
Horn, Laurence R. (1989) *A natural history of negation*. Chicago: The University of Chicago Press.  
Lopes, Ana C. M. (2005) *Texto e Coerência*. *Revista Portuguesa de Humanidades* 19(1/2), pp. 13-33.

Moeschler, Jacques (1982) *Dire et contredire. Pragmatique de la négation et acte de réfutation dans la conversation*. Berne: Peter Lang.

Raposo, Eduardo P. et al. (eds.) (2013) *Gramática do Português*, vols. I e II. Lisboa: Fundação Calouste Gulbenkian.

Rudolph, Elisabeth (1996) *Contrast: adversative and concessive expressions and their expression in English, German, Spanish, Portuguese on sentence and text level*. Berlin: Walter de Gruyter.

Sanders, Ted (1997) *Semantic and pragmatic sources of coherence: on the categorization of coherence relations in context*. *Discourse Processes* 24, pp. 119-147.

Sanders, Ted et al. (2001) *Text Representation. Linguistic and Psycholinguistic Aspects*. Amsterdam: John Benjamins Publishing Company.

van der Sandt, Rob (1991) *Denial*. In *Papers from CLS 27(2): the parasession on negation*. Chicago: CLS, pp. 331-344.

#### Fontes

CRPC (Corpus de Referência do Português Contemporâneo). Base de dados disponível em <URL: <http://www.clul.ul.pt/pt/recursos/183-crpc#cqp/>>.

CETEMPúblico (Corpus de Extratos de Textos Electrónicos MCT/Público). Base de dados disponível em <URL: <http://www.linguateca.pt/CETEMPúblico/>>.

DAR (Diários da Assembleia da República). Base de dados disponível em <URL: <http://www.parlamento.pt/DAR/Paginas/default.aspx/>>.

## Concessive marking revisited: 'but' in courtroom talk

Magdalena Szczyrbak

### Abstract

Concessivity, which appears to be a universal feature of human interaction, has been investigated by numerous linguists, most notably as an interclausal connection or as a rhetorical relation. The dialogic approach, in turn, defines Concession as a discourse-pragmatic relation which speakers realise interactionally as (cardinal or reversed) sequences of claims, acknowledgments and counterclaims (Couper-Kuhlen and Thompson 1999, Couper-Kuhlen 2000, Barth-Weingarten 2003). Among the markers associated with the Concessive relation, 'but' is especially frequent, as evidenced by naturally-occurring data representing, for instance, private talk (Barth-Weingarten 2003) and academic discourse (Łyda 2007). To date, however, no analyses of 'but' in courtroom talk have been carried out with a view to determining its role in signalling Concession. Therefore, combining insights from corpus-assisted discourse studies and the dialogic approach to the Concessive relation, this paper aims to describe the use and distribution of 'but' as well as its co-occurrence patterns in an attempt to define the role this marker plays in structuring interactional patterns. To this end, drawing on data from adversarial proceedings – exemplifying highly antagonistic communication between opposing parties – the analysis will demonstrate the most frequent Concessive schemata involving 'but' as well as the deployment of this marker by litigant parties. What is more, accounting for the context of adversarial proceedings including the participants' status and their role in interaction, the findings will be compared with previous research to see how speakers use 'but' strategically to position themselves in various communicative settings and to determine if any of the uses of 'but' are typical of trial discourse. Seen more broadly, the analysis will reveal how 'but' helps speakers to organise interaction during unequal encounters representing legal-lay communication.

## Phatic markers as conversation regulators in interviews and informal conversation

Gloria Uclés Ramada

### Abstract

The main objective of this study is to compare the behaviour of the Spanish phatic markers (PMs) ¿eh?, ¿no?, ¿sabes? and ¿vale? when they are used as mitigating or reinforcement devices (Albelda, 2007, 2015; Albelda et alii, 2014). The analysis shows differences or preferences that help characterise the selected markers. To reach this aim, a corpus has been compiled where PMs can be found in different interactive situations that have an incidence on their value, namely the docureality Gandía Shore. Fourteen episodes (i.e. circa 14 hours/100,000 words) have been transcribed and analysed. All PMs have been extracted and an individual analysis of each of these forms has been carried out.

In Gandía Shore the same speakers feature in two different discourse genres: informal conversations (in their everyday interaction, regularly recorded) and interview (when they are inquired about their feelings and impressions about the events recently occurred or about to happen). This provides a corpus that allows the comparison of PMs in two oral genres, namely (i) unplanned, spontaneous discourse where interlocutors have a relationship of proximity (informal conversations), and b) a more rigid discourse genre where difference in hierarchy and a less familiar (or even unfamiliar) relationship between interlocutors apply (interviews).

The analysis reveals that

PMs tend to be more frequent in conversation than in interviews.

Although PMs exhibit a high mitigating and reinforcement value, there are differences in their distribution across discourse genres: mitigation is more frequent in interviews whereas reinforcement has a higher rate in conversations.

Some PMs tend to be more frequent in conversation, while other PMs are preferred in interviews. These preferences might point to a difference in the kind of phatic meaning conveyed by the PM.

## Adversative Conjunctions in Lithuanian and English

Giedre Valūnaite-Oleskeviciene; Prof. Dr. Jolita Sliogeriene; Vilma Asijaviciute.

### Abstract

This paper investigates semantically adversative conjunctions (according to Halliday and Hasan's classification, 1976) and their Lithuanian counterparts with the aim to determine the functions of these connectors in both languages and to find out if the use of certain conjunctions depends on discourse (spoken or written). The approach of the relevance theory (Sperber & Wilson, 1995; Carston, 2005) was employed to establish the semantic and pragmatic behaviour of adversative conjunctions in spoken and written discourse. While carrying out this research, certain methods were used. Literature analysis helped to identify what research in the field of discourse relational devices has been done so far, which aspects of conjunctions and discourse markers are of interest to Lithuanian and foreign scholars and to provide the taxonomy of Lithuanian conjunctions. Corpus-based qualitative method was employed to determine certain tendencies in the use of adversative conjunctions in both languages. The data was drawn from the following corpora: The Corpus of the Contemporary Lithuanian Language compiled by Vytautas Magnus University and The British National Corpus by Brigham Young University. Content analysis was used to determine the presence of certain conjunctions within texts. The results of the study revealed the multi-functionality of adversative conjunctions. In some cases the semantic meaning and pragmatic function are closely related while in other cases the relations are quite loose. Adversative conjunctions are influenced by discourse in both languages and tend to be more frequent in spoken discourse. Some translation problems may occur because Lithuanian as a Baltic language has a wider variety of adversative conjunctions than the English language, which comes from the Germanic language family. The comparison of the characteristic use of English and Lithuanian conjunctions is an important question in discourse analysis studies as the results of the research could determine possible translation issues and provide some recommendations for translators, interpreters and learners of the both languages.

## Pragmatic Functions of Discourse Relational Devices in Online Customer Reviews

Tuija Virtanen

### Abstract

While electronic word-of-mouth advertising has been subject to study in marketing, psychology, library sciences and many other fields, the language of online customer reviews has been given little attention so far (but see Vásquez 2014). Yet, studies of user appreciation show that, rather than the rating metrics, the review itself is of primary importance, and the longer the text, the better (see e.g. Chevalier & Mayzlin 2006; Lin et al. 2005; Wulff & Hardt 2014). Online customer reviews constitute a mode of computer-mediated communication (CMC; Herring et al. 2013), where users are 'channelled' (Myers 2000) by the genres they employ to share information, opinion and attitude in a commodifying environment. This raises the issue of the pragmatic functions of 'prosumers' (i.e., a blend of producer and consumer)' choice of discourse relational devices (DRD). When readers turn into writers, how do they construct the reviewer persona and write audience conceptions into the text? How does the type of goods reviewed affect the choice of DRDs in such texts?

This study focuses on the DRDs identified in a corpus of written online customer reviews of books on a major retail website, with the aim of exploring how they help prosumers to create credentials for the review and socially authenticate the reviewer persona. Some 250 reviews are studied in depth in order to single out the DRDs used, their textual roles and pragmatic functions in a body of data where the type of goods is controlled. Attention is also paid to the absence of possible DRDs in these data. Findings suggest that prosumers employ DRDs that are prototypically linked to particular styles, registers and genres as these serve to forward the credibility and accountability of the reviewer persona. The choice of DRDs also appears to function in the service of audience design.

The proposed full paper has implications for the understanding of the pragmatics of DRDs in short new media texts that are part of the daily life of an increasing number of people.

### References

- Chevalier, Judith A. & Dina Mayzlin. 2006. The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research* 43(3): 345-354.
- Herring, Susan C, Stein, Dieter & Tuija Virtanen (eds.) 2013. *Pragmatics of Computer-Mediated Communication (Handbooks of Pragmatics 9)*. Berlin & New York: de Gruyter Mouton.
- Lin, Tom M.Y., Luarn, Pin & Yun Kuei Huang. 2005. Effect of Internet book reviews on purchase intention: A focus group study. *The Journal of Academic Librarianship* 31(5):461-468.
- Myers, Greg. 2000. Powerpoints: Technology, lectures, and changing genres. In: Anna Trosborg (ed.) *Analysing Professional Genres*, 177-192. Amsterdam: John Benjamins.
- Vásquez, Camilla. 2014. *The Discourse of Online Consumer Reviews*. London: Bloomberg.
- Wulff, Julie & Daniel Hardt. 2014. Can you trust online ratings: Evidence of systematic differences in user populations. *Proceedings of the 22nd European Conference on Information Systems, Tel Aviv 9-11 June 2014*. <http://ecis2014.eu/E-poster/files/0770-file1.pdf> (accessed 26 September 2014)

## Present Day Italian anzi in a cross-linguistic perspective

Jacqueline Visconti

### Abstract

Italian discourse relational device (DRD) *anzi*, roughly corresponding to English “on the contrary” (but see below), displays wide variation both cross-linguistically and across the spoken and written mode. In the parallel corpus extracted from the proceedings of the European Parliament EurParl (Koehn 2005; Cartoni, Zufferey & Meyer 2013), *anzi* corresponds to a set of very different DRDs in other languages: English on the contrary, in fact, instead, indeed, in addition; French *à l'inverse*, *au contraire*, *ou plutôt*, *même*, *plus précisément*; Spanish *por el contrario*, *mejor dicho*, *más bien*, *incluso*; German *im Gegenteil*, *vielmehr*, *sogar*, etc.

A preliminary inspection of the EurParl data allows us to distinguish at least two groups of examples.

In a first group, two predicates or two states of affairs are contrasted, one of which has not taken place, e.g. ‘vetoing the construction’ vs ‘continuing the construction’, as in (1) and (2):

(1) I membri del Parlamento, che hanno ricevuto un mandato democratico dai propri elettori, non sono stati consultati né hanno avuto l'opportunità di mettere il veto alla costruzione di questa follia, anzi la costruzione del palazzo di vetro è proseguita senza impedimenti.

(2) MEPs, who have a democratic mandate from the electorate, were not consulted on the building of this folly or allowed any opportunity to veto its construction. Instead, the building of this glass palace has continued unimpeded.

In a second type of examples, *anzi* is used to correct a preceding formulation, by introducing a more appropriate wording (e.g. ‘could prove’ vs ‘certainly will prove’), as in (3) and (4):

(3) Quindi, signor Presidente, ripeto che è importante un piano di prevenzione primaria, un piano per tentare di ridurre la domanda di stupefacenti nella nostra Comunità ma, nel contempo, ciò potrebbe essere, anzi è sicuramente insufficiente se poi non si tenta anche la strada del recupero, della cura, della riabilitazione e del reinserimento del tossicodipendente nella società.

(4) Therefore, Mr President, I repeat that we need a programme of primary prevention to try to reduce the demand for drugs in our Community but, at the same time, this could prove - indeed, it certainly will prove - inadequate if we do not set about treating, rehabilitating and the reintegrating drug addicts into society.

This paper aims to refine the characterization of the cognitive-communicative function realized by *anzi* in Present Day Italian across both the written and the spoken mode (cf. Bazzanella 2003; Visconti 2015 for a diachronic analysis).

The investigation will focus on the following two parameters in the construction ‘p anzi q’:

(i) The presence vs absence of an explicit negation on p, as in ‘non era felice, anzi era triste’ (‘she was not happy, anzi she was sad’) – examples (1) and (2) above - vs ‘stava bene, anzi molto bene’ (‘she was well, anzi very well’) – examples (3) and (4) above;

(ii) The scope of anzi, that is whether the relationship bears on phrases, clauses, or wider textual segments.

The corpora used include the Letteratura Italiana Zanichelli and the web corpus Risorse Dinamiche dell'Italiano (RIDIRE) for written data, the Lessico dell'Italiano Parlato (LIP) and the Integrated reference corpora for spoken romance languages (C-ORAL-ROM) for the spoken data, as well as a corpus of texts currently in preparation for the Italian project "Il corpus dei corpora per il nuovo vocabolario dell'italiano post-unitario" (PRIN 2012).

The achieved characterization will be used as the basis for refining the cross-linguistic analysis of the connective, looking, in particular, at the English, French, German and Spanish equivalents found in the EurParl corpus.

#### Data

C-ORAL-ROM = Integrated Reference Corpora for Spoken Romance Languages. 2005. Ed. by Emanuela Cresti, and Massimo Moneglia. Amsterdam/Philadelphia, PA: John Benjamins.

EURPARL = European Parliament Proceedings Parallel Corpus 1996-2011. Ed. by Philipp Koehn et al. (<http://www.statmt.org/europarl/>).

LIP = Lessico di frequenza dell'italiano parlato. 1993. Ed. by Tullio De Mauro, Federico Mancini, Massimo Vedovelli, and Miriam Voghera. Milano: ETASLibri.

LIZ = Letteratura Italiana Zanichelli, version 4.0. 2001. CD-ROM, ed. by Pasquale Stoppelli, and Eugenio Picchi. Bologna: Zanichelli.

RIDIRE = Risorse Dinamiche dell'Italiano (<http://lablita.dit.unifi.it/projects/RIDIRE>).



## Perspective marking and subjectivity in coherence relations: A collocation analysis of Chinese connectives

Yipu Wei and Jacqueline Evers-Vermeul

### Background & Research questions

Connectives such as *because* and *so* are frequently considered processing instructors (Britton, 1994; Mak & Sanders, 2013): they provide information on the type of coherence relation involved (e.g. temporal, adversative or causal), and in several languages they also code information on subjectivity, i.e. the involvement of a locutionary agent (Finegan, 1995). For example, the Dutch connectives *want* 'because' and *dus* 'so' (Spooren et al., 2010) and Mandarin Chinese *kejian* 'so' (Li et al., 2013) prototypically express subjective coherence relations.

On-line processing studies suggest that the processing effects of connectives are interfered by the presence of perspective markers such as *John thinks*, *perhaps* and *according to Peter* (Canestrelli et al., 2013). These perspective markers all relate to the evidentiality/certainty of the utterance, indicating – just like subjective connectives – that someone's mind is involved in the construction of the coherence relation. However, perspective marking is not restricted to evidentiality; four other evaluative dimensions can be distinguished: emotivity, expectedness, importance, and necessity (Bednarek, 2006). If subjective connectives and the perspective markers in Canestrelli et al.'s study overlap in the information of evidentiality they convey, it could be expected that language users try to avoid such repetition. However, for other types of perspective markers, this tendency is not expected.

Collocation analysis can advance our knowledge on the properties of a discourse marker on the basis of its contextual features. Therefore, this corpus-based study investigates the following research questions:

1. Do connectives of different subjectivity degrees differ in the types of collocates?
2. More specifically, do connectives differ in the types of perspective markers they co-occur with?

### Method

We focused on two Mandarin Chinese causal connectives: the specific subjective *kejian* 'so', and a connective that is underspecified in terms of subjectivity and can be used in both objective and subjective relations: *suoyi* 'so' (Li et al., 2013). A distinctive collocates analysis was performed by measuring the association strength between these connectives and other discourse elements. We retrieved data from the CCL corpus, a large, balanced Modern Chinese written corpus. Association scores (G2 and Delta-P) were calculated based on contingency tables of observed and expected frequencies (Evert, 2008; Gries, 2013).

### Results & Conclusion

The collocation analysis generated a list of words/expressions that collocated more often with either *suoyi* or *kejian*. The top 200 of these collocates were further categorized into different semantic types. The results suggest that subjectivity information coded in the connectives does pattern with the use of perspective markers, which take a large proportion of the collocates list. The general connective *suoyi* prefers contexts with perspective markers expressing the dimension of evidentiality: cognition verbs (e.g. *think*, *know*), communication verbs (e.g. *say*, *look*), modal verbs (e.g. *want*, *may*) and adverbials expressing (un)certainly. *Kejian* co-occurred more often with perspective markers related to the dimension of expectedness

(surprisingly, unexpectedly). No differences were found in the connectives' preference for co-occurrence with other perspective markers. We will discuss possible implications of these results for the interpretation of the processing results in Canestrelli et al.'s study.

## References

- Bednarek, M. (2006). *Evaluation in media discourse: Analysis of a newspaper corpus*. London: Continuum.
- Canestrelli, A., Mak, W.M., & Sanders, T.J.M. (2013). Causal connectives in discourse processing: How differences in subjectivity are reflected in eye-movements. *Language and Cognitive Processes*, 28 (9): 1394-1413.
- Evert, S. (2008). Corpora and collocations. In A. Lüdeling and M. Kytö (eds.), *Corpus linguistics: An international handbook* (pp.1212-1248). Berlin: Mouton de Gruyter.
- Finegan, E. (1995). Subjectivity and subjectivisation: An introduction. In D. Stein & S. Wright (eds.), *Subjectivity and subjectivisation* (pp. 1-15). Cambridge: Cambridge University Press.
- Gries, St. Th. (2013). 50-something years of work on collocations: what is or should be next. *International Journal of Corpus Linguistics* 18(1), 137-165.
- Gries, St. Th. & Stefanowitsch, A. (2004). Extending collocation analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1): 97-129.
- Li, F. (2014). *Subjectivity in Mandarin Chinese: The meaning and use of causal connectives in written discourse*. Ph.D. dissertation Utrecht University. Utrecht: LOT. Available online: [http://www.lotpublications.nl/Documents/365\\_fulltext.pdf](http://www.lotpublications.nl/Documents/365_fulltext.pdf).
- Li, F., Evers-Vermeul, J., & Sanders, T.J.M. (2013). Subjectivity and result marking in Mandarin. *Chinese Language & Discourse* 4(1): 74 -119.
- Sanders, T.J.M. & Sweetser, E (2009). Introduction: Causality in language and cognition, what causal connectives and causal verbs reveal about the way we think. In T. Sanders & E. Sweetser (eds.), *Causal categories in discourse and cognition* (pp.1-18). Berlin: Mouton de Gruyter.
- Spooren, W.P.M.S., Sanders, T.J.M., Huiskes, M., & Degand, L. (2010). Subjectivity and causality: A corpus study of spoken language. In S. Rice & J. Newman (eds.), *Empirical and experimental methods in cognitive/functional research* (pp. 241-255). Stanford: CSLI Publications.