5.
Connectives and discourse markers
Describing structural and pragmatical markers in the framework of textology[1]
Iꜱᴛᴠᴀ́ɴ Cꜱ�űʀʏ

This chapter is about elementary components of text/discourse to which no categorisation of universal consent has been applied yet, as it is shown by the terminological diversity in itself (*discourse particles*, *discourse markers*, *connectives*, *connectors*, etc.). Our intent is to outline the achievements of research carried out on their issues in the framework of the *Officina Textologica* (OT) project.

In the first part, publications devoted to connectives and discourse markers are reviewed. The second section discusses the main problems of identifying and classifying such elements and proposes a set of criteria that enables us to tell apart text/discourse structuring element types in a simple yet complete and useful way. Terminological issues are addressed and difficulties of empirical research are highlighted. Part three presents analyses of several text excerpts in order to demonstrate how syntactic, informational and discourse structures interact and how connectives act as an interface between them.

1. Studies related to connectives and discourse markers published in the volumes of Officina Textologica

In the original conception and the first publications of the OT project, no specific attention was paid to lexical items used as discourse markers, connectives or frame markers inasmuch as they were either considered as mere formal, logical-like means of linking sentences or simply judged to be of poor interest in a semiotic-textological approach of (written) discourse. In fact, global theoretical challenges of describing text as a complex sign as well as the large amount of work to be done on coreference and linearization hindered the issues of such, apparently isolated, phenomena. However, owing to the polyglot setting of the research program and the different scientific background of the participants, papers on structural and pragmatic markers of discourse have been published in the OT series as early as 2001.

Having treated various problems of grammar, text linguistics and textology in volume 5, the author of the present study introduces connectives in this framework as a full-fledged research topic. In (Csűry 2001b), he intends to define the category of connectives by outlining a classification of lexical and grammatical means of structuring texts. Given the lack of any comprehensive approach to connectives as well as the vagueness of their definitions in Hungarian textology, the paper mainly refers to French linguists' works, especially in pragmatics[2], manifesting a sustained attention to this category. In spite of some controversial aspects of these sources, they offer useful considerations for determining the essential features of connectives. (Csűry 2001b) points out that the latter serve to construct complex textual units as wholes by setting up semantic/pragmatic relations between explicitly and implicitly formulated contents, and form a functional rather than lexical category. Notwithstanding the cases in which they assume a connexive role as well, the primary function of connectives is to build up coherence. A series of analyses of textual examples is given at the end of the paper in order to illustrate the authors' claims, at the same time showing that phenomena related to the use of connectives are to be examined not on the sentence but on the text/discourse level.

Volume 7 is devoted to issues of linearization in correlation with information structure. In this framework, I address syntactic peculiarities of connective use on the basis of a corpus study of French and Hungarian data. These two languages differ with regard to word order: syntactic function determines the place of constituents in the former whereas in the latter, information structure decides which one is to be chosen among several possible linear arrangements. However, connectives of an adverbial nature are more or less mobile in both, and display some puzzling variations of word order. The principal claim of (Csűry 2002) is that connective position and semantic structure of text are interdependent. As it was formerly stated, connectives have an essential, semantic/pragmatic function in realizing coherence relations, and a potential one of formally linking contiguous pieces of text (or sentence) as connexity markers. While the first is effective regardless of the syntactic position of the connective, its scope varies depending on word order, and sentence- (or clause-) initial position[3] activates the second, connexity marker, function in such a way that the syntactic unit preceding it immediately is interpreted as the bearer of the meaning the given coherence relation is referred to. In other cases, the so-called left term of the connective may spread over text blocks of variable dimensions

---

[2] Adam, Anscombre, Ducrot, Moeschler, Roulet
[3] That means the leftmost possible position, constrained by the presence of other occurring terms (see section 3, page 6).

and structure, and, what is more, the immediate neighbour next to the left of the clause containing the connective may even be an embedded unit that does not directly contribute to the coherence relation it establishes. As for French, corpus research reveals massive evidence for this text and sentence building strategy, formulated by the author as the principle of the primality of connexion, and parallel phenomena were observed on a smaller sample of Hungarian.

In volume 9, which is a collection of papers on thematic progression, (Csűry 2003) discusses the question whether it is possible to draw conclusions, on the basis of the presence of certain connectors, regarding the thematic structure of their co-texts, and, reciprocally, whether the interpretation of semantic relations marked by connectives depends on the thematic structure of the surrounding text block. The answer is to some extent positive insofar as thematic shifts delimit text chunks taken in consideration as bearers of meanings linked together by the connective. Unfortunately, no clean-cut rule seems to apply to mechanically segmenting text from this point of view because of the recursive nature of thematic structures, interwoven thematic networks and the absence of mutually univoquous correspondence between (types of) textual and thematic units. The paper presents analyses of eight excerpts from texts of several types in order to demonstrate these assumptions.

Volume 10 discusses the role of conceptual schemas in constructing text. My contribution consists in examining the use of explanatory and combinatorial lexicology[4] and of hypertextual linking of lexical representations for the sake of investigating conceptual schemas at work in building text meaning. In this perspective, connectives confront lexicologists with special issues but, in a textological approach, their procedural meaning can be successfully analyzed in light of such representations of meanings in context. Reversely, by virtue of their procedural meaning, connectives make emerge latent conceptual schemas in context, as it is showed by (Csűry 2004).

Volume 13 is entirely devoted to connectives: this monograph, intituled *Kis könyv a konnektorokról* (*Small Book on Connectives*) summarizes the results of research carried out till then, marks orientations for future investigations and demonstrates the indispensability of a corpus linguistic approach. (Csűry 2005) is intended to be a contribution to working out definitions and problems concerning connectives by placing all linguistic issues to be raised in the adequate, i. e. textological, framework of investigation.

Chapter 1 deals with the notion of connectives and with the peculiarities of their function and use. After a survey of several approaches, connectives are defined (partly according to relevance theory) as a function in the text structure fulfilled by units with a procedural meaning and used for optimizing information

---

[4] For a detailed presentation of the theory, see (Mel'cuk, Clas, and Polguère 1995).

processing. This category is identified as one of the subsets of complete text constituent marking functors and to be considered in the largest category of text structuring elements. I also point out the anaphoric nature of connectives and argue for a multi-level representation of the relation of conjuncts.

In chapter 2, I examine the place of connectives in the linguistic system. A certain ambiguity was unavoidable at this point inasmuch as I intend to characterise not only the function of connectives themselves but also the classes of linguistic items usually appearing in this function. In this sense, I have to deal with the semantics of connectives, i. e. the problem of relation types and the interpretation instructions conveyed by different lexical items. As an example, one of the basic semantic relations marked by connectives, namely, contrast, is defined and described at some length, in order to present its underlying relationships as well as its surface realisations. Contrast can be defined in an exact way as a binary relation of concepts founded on negation. The way contrast-marking connectives mobilize underlying contents (implications or expectations) is also demonstrated. Since it is often difficult to access these explicitly non-manifest contents, there are several approaches to the interpretation of contrast. Apart from criticizing the argumentativist and the syllogistic ones, I emphasize the role of conceptual schemata.

In chapter 3, I focus on text structure. After an analysis and demonstration of how partial semantic structures of discourse may be jumbled, I take into account the difficulties and possibilities of localizing and delimiting the so-called poles of connectives, i. e. the portions of context which bear the meanings they link together. I describe an XML-based annotation schema the use of which permits to put further research on connectives on an empirical basis by the use of appropriately tagged corpora. I look for semantic and formal criteria in the text structure which are necessary to identify discourse constituents connected by connectives in a consistent and unambiguous way. I also reconsider the lexical characteristics of items in the role of connectives in order to highlight their specific potential for the articulation of a given context.

Volume 14 resumes the discussion on conceptual schemas and focuses on the role of scripts or scenarios in constructing and interpreting texts. In (Csűry 2005), I examine the place and function of connectives in dialogues, emphasizing the double way they may link semantic contents, the one being anaphoric (co-textual) and the other deictic (contextual). This possibility allows connectives not only to bridge units of meaning inside a single turn or belonging to different speakers' turns but to establish coherence relations as well with any element of context, including cognitive elements of the speakers' (supposed) common ground. Thus, reference can be made to scripts that are likely to underlie the current interaction. The paper starts with a review of types of text structuring elements that may be used as cues in dialogue

production/interpretation. In parallel, dialogue phases are distinguished and some basic distinction is made of dialogue types as well in terms of the formal or institutionalized character of their scenarios. On these grounds, (Csűry 2005) claims that dialogic use of deictic connectives as indirect means of structuring scenarios is common mostly in internal sequences of informal dialogues. Their functions consist, on the one hand, in ensuring that the realization of the scenario carry on normally and, on the other hand, in sanctioning and/or correcting any deviation from its expected fulfillment. As for the frequency of this type of connective use, it seems to be rather limited with respect to the overall number of words in dialogue texts, especially in comparison with explicit dialogue structuring means in dialogues that follow formal scenarios. By the same token, (Csűry 2005) suggests that the term of script/scenario should refer to two kinds of analogous structures belonging to different levels: a given type of a communicational event has a scenario as a whole, composed of a limited number of more general partial scripts, such as questioning and answering, or treating a misunderstanding. Deictic connectives may signal scripts/scenarios of both levels.

From the point of view of our concern, Volume 16 is, after Volume 5, another landmark in the evolution of the Officina Textologica project. In fact, two papers of the seven published in this volume address issues of classification and treatment of text structuring items as well as of their relationship with different aspects of cohesion and coherence. (Furkó 2011) leads the reader on the slippery ground of the so-called discourse markers by providing an evaluative overview of their study in the relevant English literature. He concludes that terminology is not unified and there is no generally accepted typology, concurring with (Csűry 2001b), who came to the same conclusion with respect to connectives. Furkó looks at the functions discourse markers have on different planes of discourse as well as the role they play in connectivity. He points out a range of uncertainties and unclarified issues the resolution of which is crucial with a view to formulating a unified approach to cohesion and coherence in general and discourse markers in particular.

The other study in question, (Csűry 2011), recapitulates the main advances of the *Officina Textologica* project in the field of connectives and the principal concerns they present with respect to semiotic textology and research on discourse markers. As for the latter, it is suggested in the introductory part of the paper that the term should be considered as a denomination of the broadest category of items structuring (meaning in) discourse, which allows us to see connectives as a subcategory in this framework, facilitating further discussion. The author then turns to French linguistics/pragmatics as an abundant source of knowledge on connectives that present several clean-cut orientations and may provide Hungarian research in textology with useful points of reference. After

defining nine criteria according to which the different approaches as well as the results they have permitted to obtain can be compared, he gives an overview of argumentation theory (Anscombre, Ducrot and others), conversation analysis (Eddy Roulet and his Geneva School), relevance theory (Moeschler, Reboul and others) and some other grammatical, semantic, text linguistic, corpus-based and diachronic approaches (Adam, Morel, Lamiroy and Charolles, Combettes). Finally, he outlines the major claims of the relevant Hungarian literature and points out the lack of a generally accepted theoretical synthesis founded on solid empirical evidence, and suitable for language description tasks. The paper concludes with programmatic statements concerning the discourse marker/connective branch of the *Officina Textologica* project in view of the fact that the multidisciplinary character of semiotic textology as a theoretical framework seems to fit such research. (Csűry 2011) urges that a comprehensive theoretical revisiting of the field be carried out, combined with empirical investigations, and that research should be extended to oral discourse.

2. Some basic problems of the study of text/discourse structuring element types

As a starting remark, we should stress that only the special category of connectives has made an object of research followed in the OT project, and not all possible kinds of text/discourse structuring elements, although such an extension is, undoubtedly, inevitable, given the global aims of textology. In what follows, let us therefore concentrate principally on connectives, making , at the same time, necessary remarks on the neighboring categories, especially as the latter need to be taken into account for a proper definition of the former.

Discourse is organized in hierarchically ordered semantic constructions of variable dimensions and complexity, composed of recursive elementary structures, and textured with lexical and grammatical relations, such as co-reference or tense relations, extending beyond sentence borders. Furthermore, these constructions are usually, but not always, articulated by different kinds of lexical items or even by complex expressions specifically used for indicating the makeup of their meaning. The relevant literature proves a high interest of researchers in this field.[5]

Let us consider first of all the linguistic items to which the most permissive conception of connecting means of text/discourse might apply. It should be indeed reasonable to identify them as a whole set of elements ensuring

---

[5] See, for example, (Charolles 1997), or the thematic issues of the journal *Discours*: *Approches fonctionnelles de la structuration des textes* (Ho-Dac and Bolly 2011), *Multidisciplinary Perspectives on Signalling Text Organisation* (Ho-Dac et al. 2012).

connexity, without reference to the status and/or position of the members of the relation they mark. One could designate this largest category, marked *1* in the figure below, as ***text structuring elements***; it embraces, among others, grammatical forms denoting possession or temporal relations, e.g. In other words, every linguistic item used expressly in order to *ensure connexity* and/or as marks of semantic and/or pragmatic relations in text/discourse on the level of the *significans* (more precisely, in the *notatio*[6]) is a text structuring element.
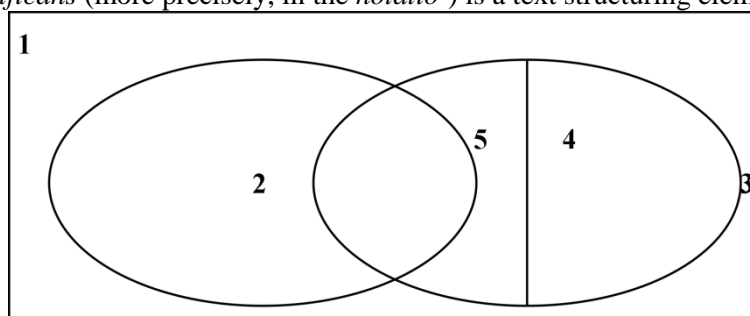

Figure 1

One of the subsets in this category, marked *2*, is constituted of ***connexive elements*** that are *used to ensure linear continuity* of text by means of marking structural and semantic relations of contiguous simple macro-architectonical units. Such a role may be played by a conjunction, a pronoun or an adverb.

The other subset, marked *3*, is that of ***text organizers marking/creating complex textual units as wholes***. Words and syntagms[7] in this function have for common feature a key role in *articulating* the *significans* of the text *according to the semantic relationships* the text has to express.

This subset has to be divided in turn so as to differentiate two further subsets. The first, marked *4* in the figure, contains items ***marking linearly composed arrays of text units***. These mark either the place or the (semantic/pragmatic) role of a macro-architectonical text unit (of the order of sentence, sentence group or sequence) in the overall structure of a given text. Their effect therefore takes place at the level of the representation of the referent (or rather that of *relatum-imago*). Lexical items used in this function have a proper lexical meaning providing them some degree of referential autonomy. The following belong to this category:

- ***linear integration markers*** (e. g. *on the one hand ... on the other hand*; *the first ... the second*, etc.);

---

6 Terms of the Petőfi model of text (passim), especially (Petőfi 1996: 12), (Petőfi and Benkes 1998: 41) and (Petőfi 2004: 27 sqq.).
7 Let us notice as a morphological property of this category that it contains no bound morphemes.

- ***repetition and reformulation markers*** (e. g. *in other words*; *in sum*, etc.);
- ***markers and predicates of universe of discourse*** (examples of the former: in this year, …; according to Steve, …, etc.; examples of the latter: *once upon a time…*; *suppose that…*, etc.).

The second subset, marked *5* in the figure, is that of ***connectives***[8]. A connective signals that the text constituents it connects stand in a specific semantic (semantic-logic, semantic-pragmatic) relationship by activating a certain interpretational-inferential procedure. It follows from this property that, in most cases, connectives relate physically present units of text to explicitly not manifested components of discourse meaning that text needs to be completed with in an explicit form in order to provide exact semantic analysis with a suitable object. The function of connective is held by verbal entities without referential autonomy (conjunctions, certain adverbs, some complex lexical units and syntagms) the presence of which allows for an unambiguous interpretation of the linked text constituents (and, therefore, that of the text).

We need to insist on the fact that the categories having been enumerated are not lexical or syntactical but **text structuring functions**, even if a set of lexical-grammatical units or another has a privileged role in assuming them. That is to say, in conceiving connectives as a collection of lexical items, we must keep in mind that we are dealing with an open set containing some clearly defined core elements as well as (more vague) peripheral ones. The latter are polysemic because they get, in some contexts, a new, procedural meaning while keeping elsewhere their original, referential one. An adverb, for example, generally used as a verbal adjunct, may be uttered as a connective being promoted to the function of a sentence adverb while no specific contextual (syntactic) feature indicates the semantic difference. Let us illustrate this issue with French adverb *maintenant* 'now'.

(1) Luc est devenu riche. Maintenant, il peut s'acheter un yacht. *(Luc became rich. Now, he can afford a yacht. — temporal relation)*

(2) Luc est devenu riche. Maintenant, est-ce qu'il faut en déduire qu'il est malhonnête? (*Luc became rich. Now, may we conclude that he is immoral? — sentence adverb in connective function; concession.*)

---

[8] We might call *connectives* the whole set of all text structuring elements and introduce another term to designate this particular subset. This solution would be more or less consistent with the abovementioned definitions. However, it is especially about this subset and its elements that one can find abundant literature, where the term *connective* generally occurs in this narrower sense.

At the same time, a term should never be called a marker, an organizer or a connective in general, only its utterances considered in (con)text. Furthermore, it should be pointed out that these categories may have a common subset. For instance, an item occurring with a connective function (i. e. having a semantic/pragmatic role) may, in the same time, ensure connexity, as in (3). Obviously, as it is illustrated by (4), the two functions do not necessarily co-occur (in this example, connexity is marked by the underlined part of the sentence).

    (3) For many years we have been building a program to give the farmer a reasonable measure of protection against the special hazards to which he is exposed. That program was improved at the last session of the Congress.
           HOWEVER, our farm legislation is still not adequate. *(Compleat Lexical Tutor v.6.2, corpus "Presidential speeches")*

    (4) The first information in an authentic form from the agent of the United States, appointed under the Administration of my predecessor, was received at the State Department on the 9th of November last.
           This is contained in a letter, dated the 17th of October, addressed by him to one of our citizens then in Mexico with a view of having it communicated to that Department. From this it appears that the agent on the 20th of September, 1844, gave a receipt to the treasury of Mexico for the amount of the April and July installments of the indemnity.
           <u>In the same communication</u>, **HOWEVER**, he asserts that he had not received a single dollar in cash, but that he holds such securities as warranted him at the time in giving the receipt, and entertains no doubt but that he will eventually obtain the money. *(Compleat Lexical Tutor v.6.2, corpus "Presidential speeches")*

One should have also noticed that Figure 1 is somewhat misleading in the light of the last remarks as it suggests that there are preexistent lexical sets of text structuring elements as such. The relationship of these functions might be better represented with graphs, like in Figure 2.
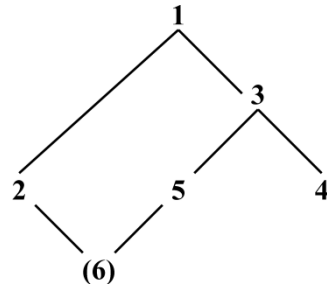
Figure 2

Let us now take a closer look at terminological issues as the very first problem related to text structuring elements is a terminological one. As for the semantic/pragmatic structures of discourse, Charolles (1997) introduces a fairly comprehensive and widely accepted taxonomy and terminology; however, given its use in cognitive semantics, the key term of *cadre* 'frame' might lead to ambiguities in a multilingual context. Consequently, these structures are rather called *discourse frames* in English. Derivates like *framing adverbials* (and their equivalents in other languages) are less ambiguous and, therefore, could be generally used referring to a class of structural markers. These correspond to our class 4, i. e. the markers of linearly composed text arrays, and establish ***forward-looking relations***. The question of how to refer to the whole category of text structuring elements remains unanswered, elements, which are instantiated, on the one hand, by conjunctions and other, more or less complex, expressions and, on the other hand, by items that do not or not always mark units of meaning one should call "frames". In particular, this is the case of marks/markers of ***backward-looking relations***, such as connectives, the anaphoric nature of which is apparent. The equivalents of the term *connective* are used in several languages (e. g. *connecteur* in French, *Konnektor* in German and in Hungarian) with various implications: it may denote a logical, a grammatical, a textological or a pragmatical approach or some specific combination of these factors. This term is often a part of terminologies that cover a differentiated set of means of structuring or marking text/discourse on the verbal, semantic and pragmatic levels. In this way, *discourse marker* and its translations may be used with reference to all kinds of items adapted to such functions. However, this term seems vague to some extent due to its controversial views in pragmatics (for instance, (Schiffrin 1988; Fraser 1990; Redeker 2006)). Nevertheless, even if considering these terminological problems unresolved, one might refer to the widest category of linguistic means playing a role in (de)marking structural units of text/discourse on the semantic/pragmatic level as *discourse markers* (DMs) and reserve the term *connective* (C) for one of its possible subcategories, distinguished from the others by (at least) their procedural meaning in logical-

like constructions of sense. In fact, it is the very vagueness of the term *discourse marker* that allows such an extended usage, whereas the above mentioned term *text organizers marking/creating complex textual units as wholes* is, despite its descriptive exactness, rather lengthy and inconvenient for current use. At the same time, it is quite clear that confusions might result from this terminological choice as well since expressions that have barely anything to do with our functional categories are often called DMs.

Evidently, there are underlying theoretical problems in addition to this terminological diversity. As a matter of fact, structures of text/discourse are described in various theoretical frameworks that results in heterogeneity of perception, categorization and terminology of the linguistic items involved. Moreover, different research projects focusing on similar or the same subjects often do not make reference to each other, especially if researchers belong to different language areas (English vs. French e. g.). Such discrepancies are also due to the fact that the lexical items under scrutiny do not constitute a homogenous class, their uses and effects are varied and, what is more, they do not necessarily and invariably mark a particular discourse relation. In the relevant literature, either the category of DMs is treated from a fully theoretical point of view, without any substantial empirical support, or empirical analyses follow a restrictively predetermined orientation that excludes in advance the taking into account of phenomena that do not fit the given theoretical framework. In most cases, the (sets of) items under scrutiny are considered separately from the others, and even if attested examples of language use replace simple introspection, their sources are restricted to particular types of discourse. For instance, the description of French *mais* by Luscher (1994) illustrates the application of relevance theory, papers written in the framework of successive versions of Anscombre's and Ducrot's argumentation theory examine individual Cs, [9] Hungarian analyses of text meaning structures (like Békési (1993)) are based on well-styled written discourse while works following the Anglo-American DM tradition focus on spoken interaction. All these are valuable contributions to a better knowledge of how semantic-pragmatic text/discourse relations can be marked, identified and interpreted but, as a consequence of research-methodological choices, the descriptions are, for the most part, partial, incommensurable, suffer from empirical weaknesses or are simply lacking. In our own work[10], we have pointed out that previous research along these lines has yielded partial results; further progress can only be made by following an integrative approach.

---

[9] For instance, (Anscombre 1983; Anscombre and Ducrot 1983).
[10] (Csűry 2001a; Csűry 2005) among others.

Beside the general theoretical problems, several particular issues arise as well. First of all, we need data in order to confirm hypotheses and ground analyses in DM research. As it was highlighted by Péry-Woodley (2005: 185), "Studies on discourse are actually characterized by a qualitative approach of small amounts of data by means of manual, thus subjective, methods […] which creates an obstacle […] to the generalization of their results. […] We are in need of techniques allowing us to apprehend and to articulate often confounded mechanisms at different levels of granularity"[11]. In recent years, indeed, one can observe a growing interest in discourse-level corpus linguistics. However, several problems arise from the very first steps of corpus analysis in this particular domain inasmuch as structures to be observed do not fit a unique and clearly describable pattern. Furthermore, computerized processing of linguistic data is a very complex issue since it is not just *words* we have to look for in corpora but *meaningful units* of variable dimension and structure, which are not only contiguous but may display embedded and overlapping arrangements, and *their relations* in a functional perspective. Thus, formal clues for computerized processing of such textual/discursive data form a matter of research in themselves. What is more, taking account of live communication involves similar efforts in the field of processing multimodal signals.

The other particular issue consists in the systematic description of lexical/phraseological items habitually occurring as DMs. Obviously, dictionaries and descriptive grammars give basic (or, in some cases, even somewhat more sophisticated) information about them and we have case studies as well as analyses of particular sets of items at our disposal. Despite this, given the divergences of orientations and the partial character of research described above, there is a lack of synthesis on DMs considered in their mutual relationship and with respect to every relevant aspect of their use, founded on solid empirical investigations. Not surprisingly, contrastive analyses are far from being systematic in spite of delicate problems of equivalence related to DMs; a comprehensive study like (Rudolph 1996) is a rare exception.

Our basic assumption is that the study of DMs is of the utmost importance as far as they guide the inferential process of interpretation of text/discourse. Furthermore, we are convinced that inquiry into linguistic phenomena related to meaning and interpretation necessitates a global approach, i. e. the study of all kinds of language use (written/spoken, monologic/dialogic, etc.). Our third contention is that the markers of relations in semantic-communicative structure,

---

[11] Our own translation from French.

whatever their nature might be, can only be studied effectively in terms of their interrelations and in context. As previous research suggests, we assume that connectives and other discourse markers do not constitute a lexical class but form a functional category. Finally, we think that some open sets of lexical items, with quasi-paradigmatic inner relationships, could be identified and described on the basis of their regular use as DMs. A corpus-based contextual approach taking into account large pieces of discourse seems to be the most appropriate way to establish a complete and coherent description of these lexical items as well as of discourse relations marked by them.

3. Connectives at the interface of syntactic, informational and discourse structures

The study of connectives is an interdisciplinary one as it is situated in the intersection of lexicology, syntax, text/discourse analysis, semantics and pragmatics. Previously, summarizing the principal claims of (Csűry 2002) and (Csűry 2003), we saw that the position of a connective in the sentence structure is closely related to the thematic structure of its context, and, consequently, influences the way in which the actual discourse sequence will be interpreted. Difficulties arising from embedded relations and recursive structures have been signaled as well. In this section, let us briefly expose, with examples at hand and referring to analyses made in the framework of the OT project, how grammar and meaning interact in the context of connectives and how an explanation of apparently unmotivated word order variations has emerged from textological research.

A fundamental problem the analyst is confronted with is the difference between syntactic and semantic/pragmatic structures. Many difficulties arise from the unpredictable character of the linguistic structures that connectives can connect. While some items are characterized by contextual restrictions, most of them admit as terms[12] words or word groups, phrases, clauses, sentences, and multi-sentence discourse sequences (i. e. macrostructural units) as well. Moreover, the structural status of the connected terms may be different and, of course, nothing allows to preview the amount of discourse making a term. Generally speaking, the discordance of syntactic and semantic relations makes

---

[12] According to (Csűry 2001a), three levels of analysis should be distinguished for the sake of a proper interpretation of coherence relations established by connectives. While they appear at articulation points of *verbal blocks*, semantic entities (their *poles*) to be taken into account for the interpretation of the given coherence relation are frequently carried (explicitly or inferably) by only parts of these verbal blocks. These parts (of texts or sentences) should be called the *terms* of the connective.

rather the rule than the exception on discourse level. Finally, the linear arrangement of discourse constituents does not indicate what belongs to a certain coherence relation and what constitutes an external element in a given context. Several connectives admit distant terms, that is to say, the clause, sentence or sequence on the left side of the connective are not obligatorily its first term since parenthetical discourse constituents may occur at this point. Furthermore, we need to count with overlapping relations as well since text structure is made up of a complex set of imbricate relations or frameworks of different kinds. Syntactic mobility of sentence adverbials acting as connectives as well as semantic effects of their utterance are to be accounted for with respect to all these factors.

In French, both positions of *par contre* and *cependant* (representing here numbers of adverbial connectives that have a similar behavior) shown in examples (5)-(8)[13] are acceptable and no clearly conceivable difference seems to result of this variation as long as analysis does not exceed the boundaries of the sentence.

(5) **PAR CONTRE**, Duroy dînait tous les jeudis dans le ménage et faisait la cour au mari en lui parlant agriculture.

(6) Duroy, **PAR CONTRE**, dînait tous les jeudis dans le ménage et faisait la cour au mari en lui parlant agriculture.

(7) Le nouveau gouvernement a **CEPENDANT** entrepris un programme de grande ampleur.

(8) **CEPENDANT**, le nouveau gouvernement a entrepris un programme de grande ampleur.

However, once considered in a given context, the different word orderings are not equally likely to occur since they prove not only to affect meaning but also to display , at times,differences of acceptability: authors (and speakers) probably have some specific reason to follow a particular pattern when placing connectives in sentence structure. Here is the source of (6) (and of (5) which is derived from the former), an excerpt from a novel of Maupassant:

(9) Il habitait maintenant rue de Constantinople, où il avait transporté sa malle, sa brosse, son rasoir et son savon, ce qui constituait son déménagement. Deux ou trois fois par semaine, la jeune femme arrivait avant qu'il fût levé,

---

[13] French examples, used for having abundant data at our disposal obtained by extensive corpus research, are not translated on purpose at this point.

se déshabillait en une minute et se glissait dans le lit, toute frémissante du froid du dehors.

Duroy, **PAR CONTRE**, dînait tous les jeudis dans le ménage et faisait la cour au mari en lui parlant agriculture ; et comme il aimait lui-même les choses de la terre, ils s'intéressaient parfois tellement tous les deux à la causerie qu'ils oubliaient tout à fait leur femme sommeillant sur le canapé.

Putting *par contre* in sentence initial position would disorient the reader at the moment as he would expect some complementary or contrasting information with regard to that what was given so far about the usual behavior of the young woman when arriving to Duroy's, while this is not the case: it is oddly contrasting peculiarities of their reciprocal visits that we feel emphasized. In English, it is rather difficult to find a perfect equivalent of *par contre*: in some cases, *in contrast* gives a satisfactory solution, in other cases, *on the other hand* is a functional equivalent, but ad hoc translations can only render its actual value in the French source. Unfortunately, the translator of an English version widely spread on internet has considerably abridged the episode by omitting its ironic and erotic elements; however, s/he has maintained the original semantic structure. Reciprocal visits are put in (a weak) contrast while, quite surprisingly, we find a somewhat abusive explicitation of the value of the connective *par contre* to which a final clause corresponds in the translated text:

(10) Duroy moved his effects to the apartments in Rue de Constantinople. Two or three times a week, Mme. de-Marelle paid him visits. Duroy, ***to counterbalance them***, dined at her house every Thursday, and delighted her husband by talking agriculture to him.

(Literally: *Duroy was now living in the apartments in Rue de Constantinople where he had transported his trunk, his brush, his razor and his soap, that was what constituted his moving house. Two or three times a week, Mme. de-Marelle arrived before he would get up, she undressed herself in a minute and slipped in the bed, shivering of outside cold. As for Duroy, he dined at her house every Thursday,* etc.)

Similarly, we understand at first glance why only (7) can correctly depict the actual state of affairs when we replace it in its original context:

(11) La Roumanie doit affronter un défi particulier en matière d'affaires intérieures et de justice. Elle a pour l'instant fait des progrès limités en matière de reprise de l'acquis en ce domaine. Le nouveau gouvernement a **CEPENDANT** entrepris un programme de grande ampleur pour mener à bien les réformes institutionnelles indispensables.

With (8) in the same context, one would suggest that limited progress has been made in spite of governmental efforts. In reality, it is former progress that is being qualified unsatisfactorily while the new government's program is being opposed to this situation. This is the way we can interpret the English version as well[14]:

(12) Romania faces a particular challenge in justice and home affairs. So far it has made limited progress in taking on the *acquis* in this field. The new government has *<AT THE SAME TIME>* undertaken an ambitious programme *(sic!)* to introduce the essential institutional reforms.

Let us now consider the aforementioned structural issues in the light of which semantic effects of connective position could be better understood. The following English examples not only facilitate the task of demonstration but also show the main cross-linguistic, if not universal, characteristics of connectives from the point of view of syntax and semantics.

Usually, we represent the use of connectives with examples in which they mark the relation of two clauses or sentences. It is indeed one of the basic configurations of structures of meaning in discourse, as shown by (13) (the numbers in brackets serve to identify syntactic units that are semantic blocks of a concessive realization of contrast at the same time):

(13) [1]Although Sam Rayburn affects a gruff exterior in many instances, [2]**NEVERTHELESS** he is fundamentally a man of warm heart and gentle disposition.

---

[14] Typography marks that the official English text contains no connective at this point. Given that the documents of the European Union are published in all official languages of the Union, we cannot establish if the French or the English version (or a third one) is the original. Either the difficulty of reproducing the exact value of a connective in an other language has lead the translator of a French source text to choose not to use any connective at all in the English version, or the intention of clearly expliciting textual relations in the French translation of an English original has made the translator introduce a connective. The spelling mistake allows for supposing at the same time that the English version was not written by a native speaker. According to our experience in language teaching, learners of foreign languages acquire a differentiated use of connectives only at a very advanced stage and often hesitate even as fluent users.

For the sake of comparison, it should be useful to give a formal, graphical representation of this construction (Figure 3).[15]
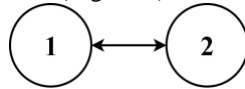

Figure 3

However, as relations marked by connectives may extend beyond sentence boundaries, we have to be conscious of the fact that a sentence containing such an element has a particular position in the meaningful structure of discourse and this must be taken into account even if the configuration is similar to the basic one seen above. In (14), we find the same connective (*nevertheless*) in unit 4 in sentence initial position:

> (14) [1]Presumably, if the reverse is the case and the good effect is more certain than the evil result that may be forthcoming, not only must the good and the evil be prudentially weighed and found proportionate, but also calculation of the probabilities and of the degree of certainty or uncertainty in the good or evil effect must be taken into account. [2]There must not only be greater good than evil objectively in view, but also greater probability of actually doing more good than harm. [3]If an evil which is certain and extensive and immediate may rarely be compensated for by a problematic, speculative, future good, by the same token not every present, certain, and immediate good (or lesser evil) that may have to be done will be outweighed by a problematic, speculative, and future evil. [4]**NEVERTHELESS**, according to the traditional theory, a man begins in the midst of action and he analyzes its nature and immediate consequences before or while putting it forth and causing these consequences. [5]He does not expect to be able to trammel up all the future consequences of his action. [6]Above all, he does not debate mere contingencies, and therefore, if these are possibly dreadful, find himself forced into inaction.

As it is shown in Figure 4, the sentence in question is in a central position in the semantic-functional structure of this text fragment. Although we can interpret the chain of [3] and [4] separately from the rest, the former is tightly connected to [2] and [1] whereas the latter is developed by [5] and [6].

---

[15] For lack of space and for the sake of simplicity, we do not specify in the following analyses the exact nature of semantic-functional relationship of all blocks of meaning. These RST-like relations, marked by arrows and horizontal/vertical arrangement, should be read intuitively.

Nevertheless marks a concessive relation between [3] and [4] and, indirectly, between the two blocks.
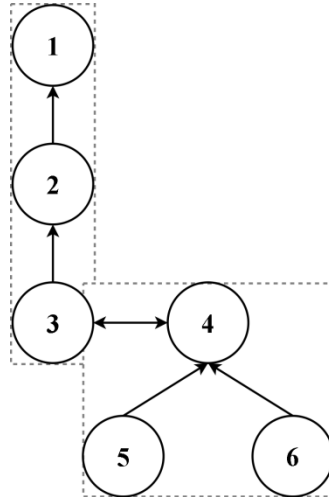

Figure 4

Conversely, connectives are often used in order to mark coherence relations between contents linked to the constituents of the same sentence. In such cases, semantic structure established by the connective does not necessarily map the syntactic one. From the point of view of constructions of meaning of discourse, context should not be neglected, either. In (15), *nevertheless* marks a concessive relation between two adjectives qualifying the same referent (*example*), i. e. puts the attribution of the qualities they denote on the same level of structure of meaning while they occupy quite distant points of syntactic structure, as shown in Figure 5.

(15) This understanding provides a very simple example of the fact that one can eliminate fear without instituting any controls. In fact, although we have dispelled the fear, we have not necessarily assured ourselves that there are no dangers. There is still the remote possibility of planetoid collision. A meteor could fall on San Francisco. Solar activities could presumably bring long periods of flood or drought. Our understanding of the solar system has taught us to replace our former elaborate rituals with the appropriate action which, in this case, amounts to doing nothing. Yet we no longer feel uneasy. This almost trivial example is **NEVERTHELESS** suggestive, for there are some elements in common between the antique fear that the days would get shorter and shorter and our present fear of war.
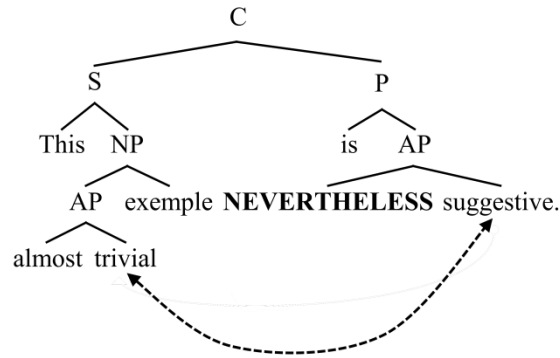
Figure 5

The role of context and of overall semantic structure of discourse can be clearly observed in example (16) which illustrates as well that sentences are not to be automatically considered as basic units of the semantic/pragmatic structure of discourse. It is once more the same connective that marks a concessive relation in this fragment but the sentence in which it occurs, [5], stands in contrast with a clause at some distance backwards, denying or at least restricting expectations that one might draw from it as conclusion due to its argumentative power. This relation is quite obvious given their lexical relatedness (the verb *communicate* of the first unit being replicated in the nominal form *communication* in the second). But (as we can see in Figure 6) there are two other units, [3] and [4], inserted between [2] and [5]. In the former the meaning of [2] is further developed whereas in [4] the author ends quoting directly his source and adds to the block formed by [2] and [3] some information on a previous state of affairs. In other words, the coherence relation marked by *nevertheless* passes over the first two units to the left of [5] implicated in this relationship only in an indirect way.

> (16) [1]Hildreth states that, "[2]As an interactive system the online catalog can dynamically communicate with its user, [3]it can be responsive and informative at a given time to a given need" [4]all of which was not possible in previous catalogues. [5]In present systems the level of communication is, **NEVERTHELESS**, limited and superficial.
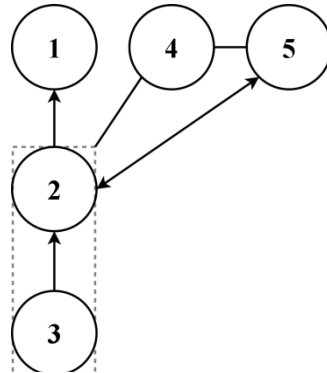
Figure 6

It is precisely in such intricate constructions of meaning, containing embedded units between the poles of a relation marked by a connective, that one finds adverbial connectives in non-initial position. In order to avoid misleading local connections between neighboring text units, writers (and probably speakers) put a thematic marker or a framing adverbial in sentence initial position and relegate the connective to specific adverbial positions inside or at the end of the sentence. The last two examples of the series (again the same text in two languages) demonstrate this phenomenon with a rather complex text structure represented in Figure 7, containing coordinate and subordinate units as well and three whole sentences embedded between the poles of the relation marked by the connective. This time we find a connective in both versions and, what is more, in the same position. In (17)and (18), *toutefois* and *however* mark the denial of the expectation that prices are the main factor of competitiveness, suggested by [3].

(17) [1]Pour permettre à l'agriculture européenne de profiter de l'évolution a priori positive du marché mondial, [2]une nouvelle réforme de la PAC doit améliorer la compétitivité de l'agriculture européenne sur les marchés tant intérieurs qu'extérieurs. [3]L'abaissement des prix profitera aux consommateurs et laissera une plus grande marge pour une différenciation des prix en faveur des produits de qualité supérieure. [4]Une orientation accrue des activités en fonction des impératifs du marché facilitera l'intégration progressive des nouveaux États membres [5]et contribuera à la préparation de l'Union aux prochaines négociations dans le cadre de l'OMC. [6]Elle aidera aussi l'Union à consolider sa position de grande puissance exportatrice mondiale.

[7]Les prix ne représentent TOUTEFOIS qu'un aspect de la compétitivité.

(18) [1]In order to help European agriculture take advantage of the expected positive world market developments, [2]further reform of the CAP must improve the competitiveness of Union agriculture on both domestic and external markets. [3]Lower prices will benefit consumers and leave more room for price differentiation in favour of high quality speciality products). [4]Greater market orientation will facilitate the progressive integration of new Member States [5]and will help prepare the Union for the next WTO Round. [6]It will also help the Union to reinforce its position as a major world exporter.

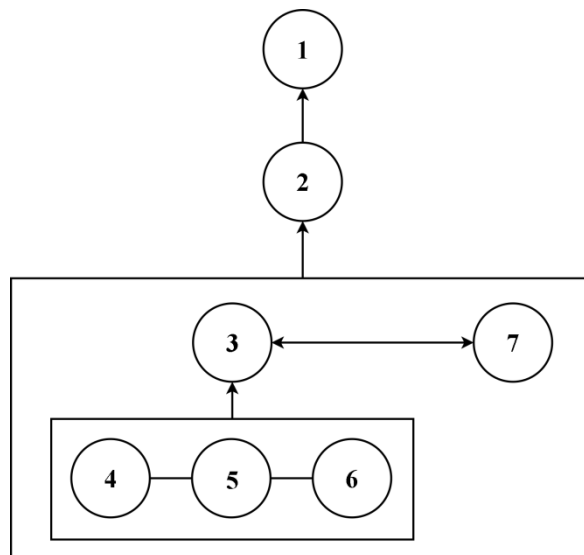[7]Prices are, HOWEVER, only one aspect of competitiveness.

Figure 7

4. Perspectives and directions for future research

Although connectives are small elements of discourse, their study is of the greatest importance owing to their key role in text meaning and coherence relations. However, there is still plenty of work to do in their research. First, as it was mentioned above, there is no universally acknowledged definition of what should be meant by *connective,* and one can find mostly partial approaches whenever phenomena related to this category become the subject of any research. Findings and descriptions of the function of connectives and lexical elements in this role contain many inaccuracies and unfounded generalizations.

Future research on connectives and discourse markers in the framework of the OT project has three mutually dependent aims. The first aim is to answer some general theoretical questions with regard to linguistic items marking the semantic-pragmatic structures of a variety of texts/types of discourse. Critical syntheses and discussion of theoretical approaches to DMs and relevant literature that were carried out so far are an obvious starting point of the research in this field. There is still a lot of work to do in order to elaborate generally admitted principles of the systematization, typology, uniform categorization and the corresponding Hungarian terminology of the linguistic items under scrutiny, which might serve as a common frame of reference for research not only in textology but also in the fields of linguistics and pragmatics in general.

In addition to the integration and synthesis of theoretical principles, our research is also aimed at solving corpus linguistic problems pertaining to the empirical basis of the study of the functional category under scrutiny. Since it is imperative that our theoretical framework should be substantiated by empirical data and that we provide the necessary sources and tools for further (especially descriptive) research based on the same principles, we also have to aim at finding solutions to the theoretical and practical problems of corpus-based connective and discourse marker research, with special reference to the development of corpus analysis tools and methodologies. According to the results of our research obtained so far, we expect that lexical features of items used as DMs as well as contextual clues will prove useful for (semi-) automatic segmentation of textual units of meaning in pre-processed (POS-tagged, syntactically annotated) corpora. In processing spoken discourse, intonation and non-linguistic modalities are likely to have a similar role that has also to be studied.

References

Anscombre, Jean-Claude. 1983. „Pour autant, pourtant (et comment): à petites causes, grands effets." *Cahiers de linguistique française* (5): 37-84.

Anscombre, Jean-Claude and Oswald Ducrot. 1983. *L'argumentation dans la langue*. Liège - Bruxelles: Pierre Mardaga Éditeur.

Békési, Imre. 1993. *Jelentésszerkezetek interpretációs megközelítése*. Szeged: JGYTF Kiadó.

Csűry, István. 2001a. *Le champ lexical de* mais *: étude lexiko-grammaticale des termes d'opposition du français contemporain dans un cadre textologique*. Studia Romanica de Debrecen, Series Linguistica. Vol. 7. Debrecen: Kossuth Egyetemi Kiadó.

Csűry, István. 2001b. „Hogyan osztályozzuk a szöveggrammatika kötőelemeit? (A konnektorok mibenlétéről és hovatartozásáról)." In *Grammatika -*

szövegnyelvészet - szövegtan [Grammar - text linguistics - textology], edited by János S. Petőfi and Irma Szikszainé Nagy, 27-33. Debrecen: Kossuth Egyetemi Kiadó.

Csűry, István. 2004. „A kérdés: (hogyan) készítsünk furcsa szótárakat?" Chap. 8., In *A szövegorganizáció elemzésének aspektusai. (Fogalmi sémák)* [Aspects of analysis of text organization. Conceptual schemata], edited by János S. Petőfi and Irma Szikszainé Nagy, 93-106. Debrecen: Kossuth Egyetemi Kiadó.

Csűry, István. 2005. *Kis könyv a konnektorokról.* Officina Textologica Vol. 13. [Small book on connectives]. Debrecen: Debreceni Egyetem, Magyar Nyelvtudományi Tanszék.

Csűry, István. 2011. „A konnektor mint diskurzusjelölő-osztály tárgyalásmódjai." In *A szövegösszefüggés elméleti és gyakorlati megközelítési módjai : poliglott terminológiai és fogalmi áttekintés: Petőfi S. János 80. születésnapjára* [Theoretical and empirical approaches to coherence. A polyglott overview of terminology and concepts. In honour of the eighty-year-old János S. Petőfi], edited by Edit Dobi, 95-107. Debrecen: Debreceni Egyetemi Kiadó.

Csűry, István. 2003. „Konnektorok és tematikus progresszió." In *A kontrasztív szövegnyelvészet aspektusai : linearizáció: tematikus progresszió* [Aspects of contrastive text linguistics. Linearization: thematic progression], edited by János S. Petőfi and Irma Szikszainé Nagy, 65-80. Debrecen: Kossuth Egyetemi Kiadó.

Csűry, István. 2002. „A tematikus struktúra és a konnektorok szintaxisa a franciában és a magyarban." In *A kontrasztív szövegnyelvészet aspektusai: linearizáció: téma-réma szerkezet* [Aspects of contrastive text linguistics. Linearization: theme/rheme structure], edited by János S. Petőfi and Irma Szikszainé Nagy, 75-92. Debrecen: Kossuth Egyetemi Kiadó.

Fraser, Bruce. 1990. „An Approach to Discourse Markers." *Journal of Pragmatics* 14 (3): 383-398.

Furkó, Bálint Péter. 2011. „Diskurzusjelölők és szövegösszefüggés: a kohézió vagy a koherencia eszközei? Az implikáció mint a szövegösszefüggés eszköze." In *A szövegösszefüggés elméleti és gyakorlati megközelítési módjai: poliglott terminológiai és fogalmi áttekintés : Petőfi S. János 80. születésnapjára* [Theoretical and empirical approaches to coherence. A polyglott overview of terminology and concepts. In honour of the eighty-year-old János S. Petőfi], edited by Edit Dobi, 37-55. Debrecen: Debreceni Egyetemi Kiadó.

Ho-Dac, Lydia-Mai and Catherine Bolly. 2011. „Approches fonctionnelles de la structuration des textes." *Discours* (8).

Ho-Dac, Lydia-Mai, Julie Lemarié, Marie-Paule Péry-Woodley, and Marianne Vergez-Couret. 2012. „Multidisciplinary Perspectives on Signalling Text Organisation." *Discours* (10).

Luscher, Jean-Marc. 1994. „Marques de connexion : procédures de traitement et guidage inférentiel." In *Langage et pertinence. Anaphores, connecteurs, temporalité, lexique*, edited by Jacques Moeschler, Anne Reboul, Jean-Marc Luscher and Jacques Jayez, 175-227. Nancy: Presses Universitaires de Nancy.

Mel'cuk, Igor A., André Clas, and Alain Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Duculot.

Petőfi, János S.. 1996. *Az explicitség biztosításának feltételei és lehetőségei természetes nyelvi szövegek interpretációjában* [Conditions and possibilities of ensuring expliciteness in the interpretation of natural language texts]. Budapest: MTA Nyelvtudományi Intézet.

Petőfi, S.János. 2004. *A szöveg mint komplex jel*. Budapest: Akadémiai Kiadó.

Petőfi, János S. and Zsuzsa Benkes. 1998. *A szöveg megközelítései*. Iskolakultúra Vol. 2. [Approaches to text]. Budapest: Iskolakultúra.

Redeker, Gisela. 2006. „Discourse Markers as Attentional Cues at Discourse Transitions." In *Approaches to Discourse Particles*, edited by Fischer Kerstin, 339-358. Amsterdam: Elsevier.

Rudolph, E. 1996. *Contrast: Adversative and Concessive Relations and their Expressions in English, German, Spanish, Portuguese on Sentence and Text Level*. Berlin - New York: Walter de Gruyter.

Schiffrin, D. 1988. *Discourse Markers*. Cambridge: Cambridge University Press.

Corpus resources

Compleat Lexical Tutor v.6.2, Cobb, Tom, Université du Québec à Montréal, http://www.lextutor.ca/

Agenda 2000 - Volume I - Communication : FOR A STRONGER AND WIDER UNION. DOC/97/6, Strasbourg, 15 July 1997, http://europa.eu/rapid/press-release_DOC-97-6_en.htm?locale=en

Agenda 2000 - Volume I - Communication :POUR UNE UNION PLUS FORTE ET PLUS LARGE. DOC/97/6, Strasbourg/Bruxelles, le 15 juillet 1997, http://europa.eu/rapid/press-release_DOC-97-6_fr.htm?locale=en

Maupassant, Guy de: Bel Ami. Association de Bibliophiles Universels, http://abu.cnam.fr/

Maupassant, Guy de: Bel Ami. eBooks@Adelaide, The University of Adelaide Library, http://ebooks.adelaide.edu.au/m/maupassant/guy/m45bl/index.html