FIVE TURNS OF THE SCREW: A CADS ANALYSIS OF THE EUROPEAN

PARLIAMENTi

Maria Calzada Perez, Universitat Jaume I, Castellón, Spain

Abstract

The present paper proposes a CADS-based analysis of European Parliament speeches, by

merging (C)DA theoretical constructs (inspired by Laclau & Mouffe 1985) and CL tools.

In this fashion, the European Comparable and Parallel Corpus of Parliamentary Speeches

Archive (ECPC) is examined along synchronic and diachronic, quantitative and

qualitative lines, in an inductive study that commutes from the micro-text to the macro-

context.

Keywords: European Parliament, Corpus-Assisted Discourse Studies (CADS), ECPC.

1. Introduction

Parliaments are institutions of the utmost importance for the world's governance.

However, until recently, they have been notably under-researched within linguistic and

related fields, despite having been defined as "institutions which are dedicated to talk"

(Bayley 2004, 1). At any rate, with the turn of the 21st century, a growing number of

language-related academic publications have shown increasing interest in parliamentary

communicative exchanges (for a thorough bibliographical compilation, see Bayley 2004;

Chilton 2002; Ilie 2010; Wodak & Van Dijk 2000). For all the differences among these

publications, most share a common scholarly goal: to connect (contextual and textual)

1

macro- and micro-levels either directly (as e.g. in Hallidayan approaches) or through a cognitive interface (e.g. Van Dijk 2002, 2010).

Most of these works follow a manual, theory-driven (as opposed to electronic, data-driven) methodology closely associated with (Critical) Discourse Analysis, or (C)DA. Furthermore, they tend to prefer top-down approaches, where macro- or micro-contextual questions are posed and answered before delving into macro- or micro-textual queries. This methodology has produced informative studies (like those cited above), but it also seems logical to argue (as we do in this paper) for a complementary data-driven analysis employing both quantitative and qualitative protocols. Corpus linguistics (CL) offers tools to make this possible.

Across fields, there is plenty of evidence (e.g. Baker 2010; Baker et al. 2008; Baker & McEnery 2005; Garzone & Santulli 2004; Koller & Mautner 2004; Mehan 1997; O'Halloran & Coffin 2004; Orpin 2005; Stubbs 1996) that (C)DA–CL integration—within what is known as Corpus-Assisted Discourse Studies, or CADS (Partington 2013)—produces "impressive results" (Garzone & Santulli 2004, 353). Nevertheless, only a handful of analyses have combined (C)DA premises and corpus-based methods to examine parliamentary communication (see Baker 2006; Baker 2010; Bayley, Bevitori & Zoni 2004; Bayley & San Vicente 2004; Bevitori 2004; Dibattista 2004; Garzone & Santulli 2004; Vasta 2004), and none have focused solely on the communicative dynamics of the European Parliament—even though Europe's most important decision—making occurs in the Euro Chamber, and its agenda influences the ideologies, behaviours, and language of national houses.

The present paper aims to begin filling this research gap. It proposes a CADS-based analysis of European Parliament speeches that, like most CL research, proceeds inductively, but based upon (C)DA premises and goals. Consequently, it first singles out some (C)DA theoretical constructs as particularly illuminating of European Parliament communication (section 1). It then outlines basic CL tools (section 2) applied to the European Comparable and Parallel Corpus of Parliamentary Speeches (ECPC) (presented in section 3) as part of a multi-layered, quantitative—qualitative analysis (section 4). The paper ends by drawing some conclusions (section 5) from this synergic CADS attempt.

2. (C)DA constructs

Out of all the attempts to apply (C)DA techniques to parliamentary communication, we have been particularly inspired by Montesano Montessori's (2014) work on Mexican legislators, where she operationalises Laclau & Mouffe's (1985/2001) Discourse Theory (DT). Notice that, like Montesano Montessori, we do not seek to describe or apply DT in full here, but only to make use of those constructs that we believe are most illuminating for our analysis. Montesano Montessori's study shows this approach is indeed rewarding.

Briefly (due to space constraints), Laclau & Mouffe (1985/2001) build a theory of hegemony and identity stemming from pivotal notions such as those of empty and floating signifiers, elements, moments, myths, and social imaginaries. In a post-Saussurean vein, Laclau & Mouffe argue that communication occurs through signs, which are the result of an interconnection between signifiers (often, words, symbols, etc.) and signifieds (or meanings). However, this interconnection is never fully fixed; signifiers are "empty" until they enter discourse, where they are eventually (and temporarily) linked to certain

meanings. From this emerges a view of discourses as arrangements of favoured signs that, deployed, impose their categorization on the world. Signs with unfixed meanings are called "elements", and Laclau & Mouffe (1985/2001, 113) associate them with floating signifiers, "incapable of being wholly articulated to a discursive chain". One type of element is the myth, aimed at producing an overall construction of society. By contrast, signs whose meanings are (provisionally) established are "moments". When a myth is hegemonically accepted, it becomes social imaginary.

In this view, communication in general (and parliamentary exchanges in particular) are dynamic, never-ending, spiralling processes, floating from emptiness to fixation and back again in a loop, leading to the emergence of myths and the "social imaginary". In order to delve into this communicative spiral, synchronic studies are never enough, since they cannot perceive the floating movements within the system; diachrony is required.

3. CL tools

Out of a broad variety of computer tools employed in Corpus Linguistics to generate data, this study uses Mike Scott's WordSmith Tools (WST) 6.0—one of the most popular, user-friendly, complete, and reliable concordancers that exists at present. With WST 6.0., we have generated keywords, clusters and keywords in context (KWICs). Furthermore, we have created sub-corpora along two axes: diachronic and the political. These constitute our main data.

The keywords result from the comparison of the terms of a given corpus (such as our ECPC corpus, described in section 3) with those of a reference corpus (such as Clear's

2003 Bank of English corpus [BoE] or Anthony's British National Corpus [BNC] word frequency list). Keywords are items of unusual frequency in a given corpus, and relate to its most idiosyncratic features. Hence, they seem to be a source of good, informative, data upon which to identify areas of research interest.

According to Biber et al. (1999, 992), clusters are "sequences of word forms that commonly go together in natural discourse". When dealing with clusters, specialists apply "cut-off points for lexical bundles which 'count'" (Kopaczyk 2012, 86). These are threshold levels below which (quantitative) relevance is seen as diluted. In this paper, we adopt the threshold advocated by Biber, Conrad & Reppen (1998): only clusters with above 40 occurrences per million words are seen as quantitatively reliable.

A keyword in context concordance is a line of words extracted from a corpus under analysis revolving around a node and its immediate context (and further linked to the larger context), as in Figure 1.

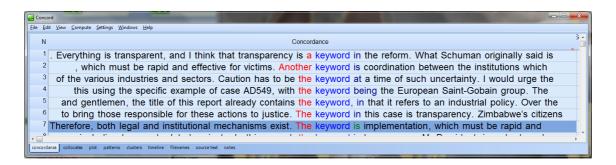


Figure 1: KWIC

Sub-corpus selection allows researchers to investigate phenomena not just in a corpus as a static whole but also in specific parts of it. Comparison between these parts (i.e., sub-corpora) may produce revealing results. Selection may be accomplished either manually,

by choosing those texts that are to be analysed, or automatically, if the overall corpus is (contextually) XML-tagged with the specific parameters required for recovery. With very large corpora (like our ECPC), automatic selection is the wisest alternative. As may be deduced, selection may happen along an ample variety of axes (in the case of automatic selection, as ample as the number of XML tags used as parameters). One possibility is that of creating sub-corpora with particular kinds of texts (in our case, parliamentary speeches) from different years in order to perform what we herein call intra-diachronic comparisons. Another approach (also resorted to in this paper) is to compare and contrast texts from two different political streams (e.g., the Conservatives versus the Social Democrats).

Corpus Linguists base the interpretation of their quantitative results on statistical measures. WST 6.0. generates a set of basic figures, such as

- 1. corpus, text, and sentence (average) word length; and
- 2. standardised type/token ratio (STTR): the ratio of the different words (i.e. [word] types) in the corpus to the its total number of words (tokens); STTR is normally calculated in sets of 1,000 words and then an average is established; it may be used to measure lexical variation.

Corpus Linguists further draw on two main types of statistical measures to establish the relevance of their data: statistical significance and effect size. The more commonly used is statistical significance. According to Gabrielatos (2012), this provides "[t]he p value of the frequency difference, as measured by a statistical test—usually log likelihood or Chisquare". Contrary to an extended use, the p-value does not actually serve to indicate "the

magnitude of an observed finding" (Rosenfeld & Penrod, 2011, 342 in Gabrielatos 2012) but rather the margin of error in the production of results. In other words, the p-value establishes the likelihood of obtaining the same figures if we repeat the experiment with different samples from the same or similar corpora. The p-value ranges from 0 to 1; the lower it is, the smaller the margin of error. Corpus linguists tend to use 0.01 and 0.05 as their top thresholds.

The measure that does weigh the magnitude of results is effect size (often represented as %Diff), based on frequency difference, which establishes "the practical significance of a result, preventing us claiming a statistical significant result that has little consequence" (Ridge & Kudenko, 2010: 272 in Gabrielatos 2012). Unlike statistical significance, effect size does not have a threshold above or below which we can safely claim a given magnitude for our results; we always need to place our results in context and interpret them in relation to other results.

Both statistical significance and effect size may be calculated with the aid of the UCREL Effect Size and Log-Likelihood Calculator (http://ucrel.lancs.ac.uk/llwizard.html), by inserting frequencies for the term under scrutiny in two corpora (or sub-corpora) and the overall sizes of these corpora, as in Figure 2:

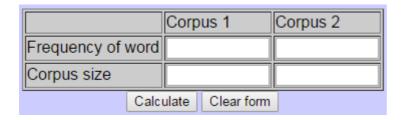


Figure 2: UCREL Calculator

4. The ECPC Archive

The European Comparable and Parallel Corpora of Parliamentary Speeches Archive (ECPC), compiled at the Universitat Jaume I (Spain) (and available to the academic community at http://ecpc.xtrad.uji.es/glossa/html/index.php?corpus=ecpc), is a collection of XML-tagged corporaⁱⁱ containing speeches from three European chambers: the European Parliament (EP), the British House of Commons (HC), and the Spanish Congreso de los Diputados (CD). The EP corpus (used here) consists of day sessions of European Parliament proceedings ranging from 15 April 1996 to 25 June 2011 (the date when the EP stopped producing multi-lingual versions of its debates).

ECPC consists of two main EP sub-corpora: EP_en, the official English version of EP proceedings, and EP_es, its official Spanish counterpart. The former (used in the present paper) has 51,345,208 tokens of (oral and written) speeches only (excluding metatextual information, headings, and comments such as "applause"). EP debates in English are published in the *Official Journal of the European Union* and downloadable from the EP's website (at http://www.europarl.europa.eu/plenary/en/debates-video.html).

Like the other components of ECPC, the EP_en corpus is XML-tagged with information about the structure of day sessions, speakers, and speeches. Among our XML parameters, two are of particular relevance for this study: "date_of_sessions" and "ep_group". Both allow us to create automatic subsets of the overall EP_en corpus along, respectively, the diachronic axis (in the case of "date_of_sessions") and the political axis (for "ep_group").

A cautionary note seems pertinent here. The main goal of this paper is to capture the EP's official representation of speeches' meaning rather than the speeches' meaning as such. The EP_en corpus contains one of the primary official representations of EP debates in English. As such, it is an important source of linguistic and ideological material.

5. Analysis

W. Nelson Francis, a CL pioneer, used to wear a spanner pin in his lapel after a colleague remarked that "anyone who would use a computer on good literature was nothing but a plumber" (Svartvik 2007, 20). In this paper, we want to "apply the spanner" to perform a series of five research "screw turns", in order to depart from the micro-level of texts (consisting of signifiers and signifieds, elements and moments) and approach the macrocontext (the sphere of discourse and society). For our first turn, we propose a quantitative study of EP en clusters and concordances around the key term "economic" (section 5.1), which provides an overall, synchronic snapshot of the whole genre, capturing an important section of the EP's social imaginary. In the second turn, we perform a diachronic study of speeches from 1999, 2005 and 2010 (section 5.2), that serves to unveil the floating "elements" behind the apparently static "moments" integrating the EP's social imaginary. Our third turn incorporates a qualitative stance alongside the previous procedures (section 5.3). The fourth turn revolves around the political axis: we produce a second set of sub-corpora, of Conservative PPE and Social-Democratic PSE speeches delivered in 2005 and 2010 (section 5.4). All these turns, at the micro-textual level, are informed by effect size and log likelihood statistics, which we contrast to the macrocontext in our fifth, final turn of the screw (section 5.5).

5.1 The first turn of the screw: a (static) snapshot of EP_en

We start our analysis of the EP_en corpus with a brief examination of the key term "economic" (with 56,805 occurrences). The reason for this selection is that when comparing EP_en with the BoE and the BNC (see section 3), this word emerges as the second (vs. BoE) or third (vs. BNC) most idiosyncratic EP_en adjective—with "European" (which is the topic of a forthcoming paper) in first place. We can thus safely claim that this is a frequently exchanged signifier within the European Chamber and a relevant "nodal point" of EP discussion, especially given that since the Lisbon Treaty (signed in 2007), the EP's "economic" role has been an increasingly decisive component of EU governance.

One way of examining the "economic" signifier is through cluster analysis. In short, the concordancer (WST 6.0 here) automatically generates bundles of words that appear together with the term under scrutiny; the researcher focuses on those clusters above a frequency threshold. In our case, we have looked at EP_en's three- and four-word clusters (a standard type of CL analysis) around the term "economic". As stated above, we apply Biber, Conrad & Reppen's (1998) threshold level (of 40 hits per million words) to our EP_en corpus (of over 50 million words). This means we examine those 3 and 4-word groups around "economic" with more than 2000 occurrences in the corpus.

| No. | Cluster | Freq. |
|-----|-------------------------------|-------|
| 1 | ECONOMIC AND SOCIAL | 5086 |
| 2 | ECONOMIC AND MONETARY | 4098 |
| 3 | THE ECONOMIC AND | 2696 |
| 4 | ON ECONOMIC AND | 2543 |
| 5 | ECONOMIC AND MONETARY AFFAIRS | 2430 |
| 6 | COMMITTEE ON ECONOMIC | 2411 |
| 7 | ON ECONOMIC AND MONETARY | 2367 |
| 8 | COMMITTEE ON ECONOMIC AND | 2319 |
| 9 | THE COMMITTEE ON ECONOMIC | 2303 |

Table 1: EP_en reliable clusters around "economic"

As we see in Table 1, the most frequent (quantitatively reliable) cluster in EP proceedings from 1996 to 2011 is "economic and social" (with over 5,000 occurrences, more than double the threshold level). Next comes a cluster ("economic and monetary", with 4,098 occurrences) that, together with the following seven clusters in Table 1, make up the name of an important EP body, the Committee on Economic and Monetary Affairs (ECON). By manually looking at KWIC lines of "economic and monetary", we find that 60% of them actually refer to this Committee. Quantitative reference shows that this is one of the most prominent EP bodies mentioned in the Euro Chamber, together with the Committee on the Environment, Public Health and Food Safety (3,048 occurrences), the Committee on Budgets (2,757 occurrences), and the Committee on Legal Affairs (2,274 occurrences).

According to these data, then, "economic" is an EP_en key signifier that largely fixes its meaning over time in association with "social" matters or as part of a meta-reference—a relevant EP institution—through which the EP points (via metonymy) at itself. These two signifying "moments" of the term "economic" become observable pieces of the EP social imaginary.

5.2 The second turn of the screw: a moving image of EP_en

However, as Laclau & Mouffe argue, "moments" are complex and meaning fixation is a rather dynamic process. This is not appreciated in synchronic operations (like the one above), but is more obvious through diachronic studies. For this reason, we isolated three EP_en sub-corpora with speeches from the turn of the 21st century (1999), the middle of its first decade (2005), and the end of this first decade (2010). The three sub-corpora are: EP_en_99 (3,021,857 tokens), EP_en_05 (3,214,605 tokens), and EP_en_10, (3,106,780 tokens). We then generated clusters for these three periods of time. Our cluster cut-off point should now theoretically be around 120 occurrences; however, we decided to be a bit more flexible for informative purposes, and show below the first 20 clusters in each year.

| EP_en_99¤ | ¤ | EP_en_05¤ | ¤ | EP_en_10¤ | ¤ |
|----------------------|-------------|-------------------------|--------------|-------------------------|--------------|
| ECONOMIC AND SOCIAL | 313¤ | ECONOMIC AND SOCIAL | 244¤ | ECONOMIC AND SOCIAL | 275¤ |
| ECONOMIC AND · | | | | | |
| MONETARY¤ | 269¤ | ECONOMIC AND MONETARY | 146¤ | THE ECONOMIC CRISISE | 249¤ |
| ON ECONOMIC AND | 172¤ | ON ECONOMIC AND | 116¤ | THE ECONOMIC AND | 228¤ |
| COMMITTEE ON- | | ECONOMIC AND MONETARY | | | |
| ECONOMIC | 165¤ | AFFAIRS ¤ | 112¤ | ECONOMIC AND MONETARY | 207¤ |
| ECONOMIC:AND: | | | | | |
| MONETARY AFFAIRS | 163¤ | COMMITTEE ON ECONOMIC | 106¤ | OF THE ECONOMIC | 187¤ |
| ON ECONOMIC AND | | ON ECONOMIC AND | | | |
| MONETARY AFFAIRS | 159¤ | MONETARY | 105¤ | ECONOMIC AND FINANCIAL | 177¤ |
| COMMITTEE ON- | | COMMITTEE ON ECONOMIC | | ECONOMIC AND MONETARY | |
| ECONOMIC:AND¤ | 159¤ | AND¤ | 105¤ | AFFAIRS¤ | 133¤ |
| COMMITTEE ON- | | | | | |
| ECONOMIC:AND: | | THE COMMITTEE ON | | | |
| MONETARY¤ | 159¤ | ECONOMIC | 101¤ | ON ECONOMIC AND | 122¤ |
| ON ECONOMIC AND | | | | | |
| MONETARY¤ | 159¤ | SOCIAL AND ECONOMIC | 91¤ | COMMITTEE ON ECONOMIC | 118¤ |
| THE COMMITTEE ON | | | | ON ECONOMIC AND | |
| ECONOMIC¤ | 151¤ | THE ECONOMIC AND | 83¤ | MONETARY¤ | 117¤ |
| | | | | COMMITTEE ON ECONOMIC | |
| THE ECONOMIC AND | 151¤ | ECONOMIC GROWTH AND | 68 ¤ | AND□ | 117¤ |
| THE COMMITTEE ON | | | | THE COMMITTEE ON | |
| ECONOMIC:AND¤ | 145¤ | POLITICAL AND ECONOMIC | 6 7¤ | ECONOMIC | 114¤ |
| ECONOMIC AND | | | | | |
| MONETARY AFFAIRS | | | | ECONOMIC:AND:FINANCIAL: | |
| AND¤ | 108¤ | ECONOMIC SOCIAL AND | 6 7¤ | CRISIS¤ | 100¤ |
| AND MONETARY | | | | | |
| AFFAIRS AND | 108¤ | THE ECONOMIC AND SOCIAL | 4 6 ¤ | SOCIAL AND ECONOMIC | 93 ¤ |
| THE ECONOMIC AND | | ECONOMIC AND SOCIAL | | THE ECONOMIC AND | |
| SOCIAL¤ | | COHESION | 46 ¤ | FINANCIAL¤ | 93 ¤ |
| OF ECONOMIC AND | 91 ¤ | OF THE ECONOMIC | 45¤ | FINANCIAL AND ECONOMIC | 85¤ |
| SOCIAL AND ECONOMIC¤ | 87 ¤ | ECONOMIC AND POLITICAL | 41¤ | THE CURRENT ECONOMIC | 7 6 ¤ |
| ECONOMIC AND | | | | | |
| MONETARY UNION | 81 ¤ | IN THE ECONOMIC | 38¤ | ECONOMIC CRISIS AND | 75¤ |
| ECONOMIC AND SOCIAL | | ECONOMIC DEVELOPMENT | | | |
| COHESION | 80¤ | AND□ | 34¤ | THE ECONOMIC AND SOCIAL | 75¤ |
| POLITICAL AND | | | | | |
| ECONOMIC¤ | 62 ¤ | POLITICAL ECONOMIC AND | 33¤ | OF THE ECONOMIC CRISIS | 74¤ |

Table 2: Reliable "economic" clusters from 1999, 2005, and 2010

A quick look at Table 2 suggests that "economic" clusters are reasonably stable. In fact, some repeat themselves over the years (e.g. "economic and monetary", "economic and monetary" affairs") while others (e.g. "economic and social") even maintain their (first) position over a decade (in 1999, 2005 and 2010). However, underneath this apparent stability, there are varying degrees of (floating) dynamism, which are uncovered by an effect size and log likelihood study, performed with the help of the UCREL Log-Likelihood Calculator.

The first case of dynamism is drawn precisely from a study of Table 2's second most frequent cluster (i.e. "economic and monetary"), which permeates EP en 99 (269

occurrences), EP_en_05 (146 occurrences), and EP_en_10 (207 occurrences). A manual analysis reveals that, in 1999, 163 occurrences of this cluster specifically refer to the Committee on Economic and Monetary Affairs (and its variants, which in 1999 were "the Committee on Economic and Monetary Affairs and Industrial Policy" and "the Economic and Monetary Affairs Committee"). In 2005, in turn, 112 referred to the Committee on Economic and Monetary Affairs (and it sole variant for 2005: "Economic and Monetary Affairs Committee"). In 2010, for its part, 133 referred to the Committee on Economic and Monetary Affairs (and it sole variant for 2010: "Economic and Monetary Affairs Committee").

The UCREL Log-Likelihood and Effect Size Calculator generated Table 3, when comparing, across time, the number of references, out of corpus sizes, to the Committee on Economic and Monetary Affairs:

EP_en_99 vs 05

| Item | 01 | %1 | 02 | %2 | LL | %DIFF |
|------|-----|------|-----|--------|-------|-------|
| Word | 163 | 0.01 | 112 | 0.00 + | 12.93 | 54.82 |

EP_en_05 vs 10

| Item | 01 | %1 | 02 | %2 | LL | %DIFF |
|------|-----|------|-----|--------|------|--------|
| Word | 112 | 0.00 | 133 | 0.00 - | 2.59 | -18.61 |

EP_en_99 vs 10

| Item | 01 | %1 | 02 | %2 | LL | %DIFF |
|------|-----|------|-----|--------|------|-------|
| Word | 163 | 0.01 | 133 | 0.00 + | 3.93 | 26.00 |

Table 3: Statistics measures for reference to the Committee on Economic and Monetary

These data allow us to state that:

1. there is a +54.82 %DIFF difference between figures from EP_en_99 and EP_en

05; this difference is statistically significant (at the level of p < 0.001 and LL of

12.93);

2. there is a -18.61 % DIFF difference between figures from EP_en_05 and

EP_en_10; this difference is not statistically significant.

3. there is a 26.00 % DIFF difference between figures from EP_en_99 and

EP_en_10; this is also statistically significant (at the level of p < 0.05 and a LL

of 3.93).

In sum, this means that there was a statistically significant difference in the reference to

the Committee on Economic and Monetary Affairs between 1999, 2005, and 2010. The

decrease was particularly intense between 1999 and 2005. There is a slight recovery in

the use of the reference between 2005 and 2010. The recovery is not, however,

statistically significant.

5.3 The third turn of the screw: A qualitative cut

We now go back to Table 2 and notice that "economic growth and" is only identified by

WST 6.0. as a cluster for 2005, with 68 occurrences. Since this figure is below our

quantitative (120) threshold level, we put quantification on hold and decide for largely

qualitative protocols. At any rate, using CL tools does not entail the automatic separation

15

of quantitative and qualitative methods, both of which may (and have) proved useful to examine communication.

Even though Table 2 only identifies "economic growth and" as a WST-generated-cluster for 2005, this combination of words was also present in 1999 and 2010. A manual analysis of its KWIC lines proves informative, especially if we compare concordances from 2005 (68 occurrences) and 2010 (50 occurrences). In both years, the "economic growth and" compound is mainly used as part of three different kinds of constructions:

- noun + prep + (adjective) + cluster, where the noun modifies the cluster or vice versa (e.g. "Sustained high oil prices represent a significant risk to global economic growth and are a particularly damaging aspect for poorer countries."). Emphasis here lies on portraying "economic growth" as a noun modifier, hence as part of a state (where agency is not necessarily invoked).
 Furthermore, the logic of the structure implies the prior existence of the reality portrayed by the cluster (in our example, if there is a risk to economic growth, then there must be prior economic growth).
- 2. verb (representing material processes) + cluster as an object (e.g. "They are needed in order to restore confidence [...], promote **economic growth and** increase opportunities for jobs and prosperity"). Emphasis here lies on the portrayal of economic growth as a result of an active process. Agency here is either explicitly determined or invoked. The latter is normally the case in our corpora.

3. Other constructions (i.e. as part of relational processes, juxtaposed phrases, adverbial prepositional phrases or passive voice constructions) upon which we will not dwell due to space constraints.

In 2005, out of the 68 occurrences of "economic growth and", 29 were of type (1), 22 of type (2), and 17 of type (3). Table 4 shows concordances in their co-texts:

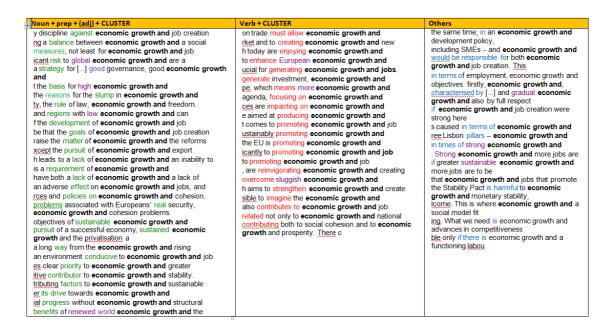


Table 4: Concordances of "economic growth and" in EP_en_05

In 2010, out of a total of 50 occurrences of "economic growth and", 11 of type (1), 29 of type (2), and 10 of type (3). Table 5 shows concordances in their co-texts:

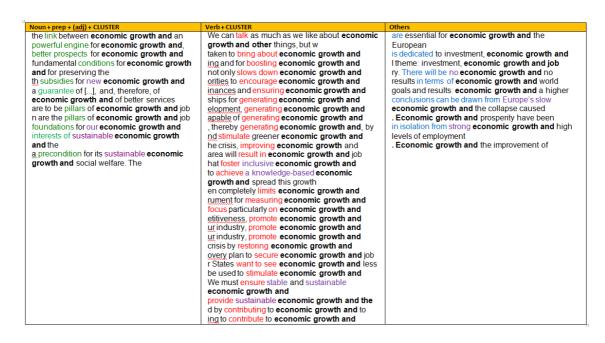


Table 5: Concordances of "economic growth and" in EP en 10

A UCREL comparison between figures for these construction types throughout the periods under analysis is included in Table 6 below. To obtain Table 6, we compared frequencies for each construction type out of corpus sizes for 2005 and 2010.

TYPE (i) CONSTRUCTIONS: 2005 VS 2010

| Item | 01 | %1 | 02 | %2 | LL | %DIFF |
|--------------|----|------|----|--------|------|--------|
| Item Word | 29 | 0.00 | 11 | 0.00 + | 7.80 | 154.79 |

TYPE (ii) CONSTRUCTIONS: 2005 VS 2010

| Item | 01 | %1 | 02 | %2 | LL | %DIFF |
|------|----|------|----|--------|------|--------|
| Word | 22 | 0.00 | 29 | 0.00 - | 1.22 | -26.68 |

TYPE (iii) CONSTRUCTIONS: 2005 VS 2010

| Item Word | 01 | %1 | 02 | %2 | LL | %DIFF |
|--------------|----|------|----|--------|------|-------|
| Word | 17 | 0.00 | 10 | 0.00 + | 1.60 | 64.30 |

Table 6: Statistics measure for "economic growth and" in 2005 and 2010

These data allow us to state that "economic growth and" was used differently (with

dissimilar kinds of meaning emphasis) in 2005 and 2010.

1. In 2005, "economic growth and" was presented as (a presupposed) part of a state

more frequently (+154.79) than in 2010. This %DIFF difference is statistically

significant (at the level of p < 0.01 and LL of 7.8).

2. In 2005, "economic growth and" was presented as the direct recipient of an

active material process less frequently (-26.68) than in 2010. This %DIFF

difference, however, is a tentative conclusion (which requires further research

confirmation) since it is not statistically significant (LL of 1.22).

3. 2005 and 2010 also differed with regard to other constructions. However, these

data will not be discussed due to space constraints.

The patterns of use for "economic growth and" confirm (floating) dynamism.

5.4 The fourth turn of the screw: the political axis

The fourth and final case of dynamism revolves around 2010's cluster "economic and

financial crisis" (see Table 2), with 100 occurrences. As above, this figure is below the

threshold of 120, which seems to suggest that setting aside quantification is a logical

option. At any rate, if we look for instances of the same cluster in 1999 and 2005, there

is not a single occurrence in either of these (reasonably large) corpora. There is no need

to provide log likelihood or effect size data (which we will not do, to save space) to realise

19

that the association of "economic and financial crisis" leads to a (recent) "moment", which merits further exploration. Also, notice the proliferation, in 2010, of different signifiers potentially referring to the crisis ("economic crisis", "economic and financial", "the economic and financial", "financial and economic", the current economic", "of the economic crisis"), as opposed to what happens in 1999 and 2005 (see Table 2).

With ECPC's metatextual XML-tagging, we can further refine our study by creating two sub-corpora (PPE and PSE, respectively) with only those 2010 speeches containing "economic and financial crisis" when uttered by Conservative PPE and Social-Democratic PSE MEPs. Furthermore, we can then identify days with a particularly high concentration of this cluster, and can also isolate specific speeches for analysis. 23 November 2010 was the day that year when PPE and PSE politicians used the "economic and financial crisis" cluster most often, on 14 occasions. On that day, the sixth point on the agenda was "Voting time"; MEPs had to decide on an array of issues, one of which was the "Mobilisation of the European Globalisation Adjustment Fund" (EGF). According to the EP's website, "[t]he EGF provides one-off support to workers losing their jobs as a result of major structural changes in world trade patterns." Point 9 of the agenda consisted of oral and written "Explanations of vote"; it was here that the cluster mostly appeared. Of the 14 times the cluster was used, 11 were by PPE politicians and only 3 by members of the PSE. A total of 9 PPE communications were written explanations of vote regarding the EGF (as opposed to oral speeches); there was only one written explanation of the same issue sent by a PSE MEP (Romanian Silvia-Adriana Ticău). Finally, of the 9 PPE written explanations, 6 belonged to Paulo Rangel, a Portuguese representative who had actually sent two basic templates (instead of 6 different documents), the first of which was repeated 4 times and the second twice, with

the only alteration being the name of the region applying for EGF aid and the sum of money requested. It is Rangel's templates that merit our closest attention, since (we will assume) repetition is bound to have a greater (ideological) impact amongst its receivers. It would then seem plausible to examine Rangel's most productive template (used on four occasions, with very slight alterations) by comparing it with the only written explanation including the "economic and financial crisis" cluster sent by a PSE delegate, Ţicău (see Appendix 1 for both texts).

The two documents are similar. They are both short—184 (PPE Rangel) and 186 (PSE Țicău) words—and basically consist of the same components: the authors' positive vote with regard to a request by two Netherlands regions (Noord-Brabant and Zuid-Holland) for assistance from the EGF, with context sustaining their decision(s).

A brief look at the documents shows some interesting differences that merit attention. Rangel presents a deductive explanation, in which the (EP-institutional) context comes first and the yes-vote is revealed at the very end, as the logical consequence of the context ("I therefore voted in favour of this resolution"). In contrast, Țicău chooses an inductive structure, where the vote comes at the very beginning and there are no explicit connectors with the context ("I voted for the European Parliament resolution on the mobilisation of the EGF for granting aid to the redundant workers"). Rangel's intervention is more condensed than that by Țicău, both syntactically and lexically. Syntactically, Rangel's document only has 3 sentences (with a mean of 57.33 words per sentence), whereas Ticău's consists of 7 (25.29 word mean per sentence). Lexically, Rangel's STTR is 62.79 while Țicău's is 56.50. This implies that Rangel's text is more (terminologically) varied than that of Țicău. There are further differences that may be taken into consideration.

Rangel's text foregrounds the document's (EP, institutional) context. In order words, it discusses the EU/EP procedures ("The request [...] fulfils all the legally established eligibility criteria"; "[...] the scope of the EGF was [...]"), while relegating specific data from the social setting ("821 redundancies from 70 companies"; "regions of Nord Brabant and Zuid Holland") to a secondary (syntactically subordinate) position. By contrast, in her 7 sentences, Ticău mainly focuses on the social (rather than EP institutional) background. She gradually presents the request by the Netherlands ("In December 2009, the Netherlands submitted a request for assistance...") and the social situation underlying it ("The application concerns 821 redundancies made in 70 enterprises..."; "The redundancies were made in the period..."; "The economic and financial crisis has also caused a drop in demand in the printing and publishing sector..."; "The printing and publishing industry in the Netherlands went through a major restructuring process"). Agency is not explicitly revealed in either of the two texts, but it is at least hinted at by Ticău through, for instance, the passive voice ("The redundancies were made...") and the agentification precisely of "the economic and financial crisis" cluster, to which the responsibility is ascribed ("The economic and financial crisis has also caused a drop in demand"). Finally, Rangel's document does not express any hint of criticism, whereas Ticău ends her explanation with a request to simplify bureaucracy and ease EU control: "I believe that the procedure for allocating these funds must be simplified to facilitate the affected enterprises' access to the EGF."

Hence, it seems reasonable to argue that, although the overall results of these documents were the same (a vote in favour of granting EGF aid to two regions in the Netherlands), the signifying means to achieve these results were different, in argumentative structure, syntactic and lexical density, institutional or social context foregrounding, agency, and

final (more or less critical) position. In short, the (floating) signifiers used by both MEPs show some degree of confluence (after all, both speakers voted yes) but contending "elements" fight to impose their preferred discourse.

5.5. Into the macro-context

Our CADS-based study of EP_en's key signifier "economic" (and some of its preferred clusters) reveals some data that may, in turn, be connected to some interesting macrocontextual facts.

By analysing (in our first turn of the screw) the EP_en corpus synchronically, as a whole, we not only confirm that the "economic" arena is one of the EP's most idiosyncratic nodal points (together with the "European" site) but are also able to spot two of the signifier's most consolidated "moments" in the Euro Chamber: its meaning fixation in association with the "social" sphere and its envelopment within its institutional (meta-referential) structure.

Looking at the macro-context, this is hardly surprising. This double "fixation" is deeply entrenched in the EU, has been verbally primed (as per Hoey 2005) by EU institutions, and has been pushed into the centre of the EU and EP "social imaginary". In other words, the EU/EP has traditionally portrayed itself in an institution-building light, seeking "economic" (understood as indivisible from social) progress. There is plenty of trace of this imaginary, from the creation of the EEC in 1957 to today's EU. For instance, the "Preamble" to the Treaty of Rome (1957) already starts with affirmation of this institutional-cum-economic/social conglomerate:

HIS MAJESTY THE KING OF THE BELGIANS.

[...]

DETERMINED to <u>lay the foundations</u> of an ever-closer union among the peoples of Europe,

RESOLVED to ensure the economic and social progress of their countries by common action to eliminate the barriers which divide Europe,

Underneath this social imaginary, there is a constant (floating) dynamism, whose (ever-changing) "elements" may be accessed with quantitative and qualitative, diachronic and comparative turns of the screw. This dynamism reflects (and reinforces) macrocontextual events.

In this way, 1999's EP discourse contains a statistically significant institutional metareference to the Committee on Economic and Monetary Affairs (ECON). This coincides with the coming-into-force of the Treaty of Amsterdam, right before the signing of the Treaty of Nice (in 2001). Both treaties reorganised EU institutions and assigned greater monitoring power to the EP, especially on economic matters (the area of competence of ECON).

By the year 2005, the EU/EP's social imaginary comes to draw on (and reflect) clusters such as "economic growth and". In 2005, institutional reorganisation is well under way, but suddenly experiences a halt when the European Constitution (an attempt to intensify the EU's institutional powers and competences) is abandoned. The EU in general (and the EP, as one of its governing arms, in particular) turns its sights on economic growth. Both the *Commission Work Programme* for 2005 and *Eurostat Yearbook 2005* testify to this fact. With the commencement of the year, the Commission sets its main priority in the following way: "An upturn in economic growth is the central policy objective of the

Commission" (European Commission 2005, 5). In retrospect, *Eurostat Yearbook 2005* singles out "the following strategic goal for the next decade: to become the most competitive and dynamic knowledge-based economy in the world capable of sustainable economic growth with more and better jobs and greater social cohesion." (Eurostat 2005, 15). The EU economy is growing (at 2%) but "[a] new impetus to the European dimension of competitiveness will be given through the mid-term review of the Lisbon strategy" (European Commission 2005, 5). This growth comes to an end in 2007, as may be seen from Figure 3:

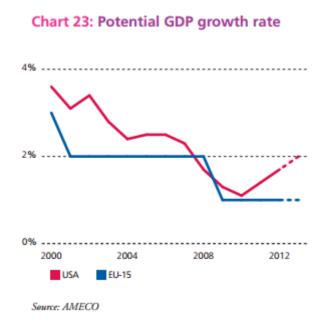


Figure 3: EU's economy chart–Extracted from Balcerowicz et al (2013)

Our data show that 2005 and 2010 diverge with regard to the extent to which the "economic growth and" cluster is used as part of a noun construction or as the result of an active process, with an explicit or implicit agent. It seems interesting to point out that divergence in use coincides with a change in macro-context. Within 2005's bonanza, growth is often depicted (in the EP) as an assumed status quo. Amidst 2010's recession,

MEPs present growth as the result of (desirable) active measures (whose agency is assigned implicitly).

So then, in 2007, the global crisis impacts the EU for the first time in decades. "The crisis hit hard", as the European Commission (2010, 3) acknowledges; and, as Copeland and James argue, (this kind of) "discourse [is] central in shaping [and reflecting] the EU's economic governance and driving its economic reform agenda" (Copeland & James 2011, 13). A series of complex phenomena (Constantinescu & Constantinescu 2013, 172), each with its floating signifiers, its elements and moments, is unleashed stage by stage. That is, to cut a long story short, according to Copeland & James (2011), the crisis starts by being "financial", goes through a "financial and economic" period, and eventually turns "economic" or rather "economic and financial". By 2010, the EU is already at an "economic and financial crisis" "moment", and this year sees the innovative use of this very same cluster in the EP.

To face this "moment", on 26 March 2010, the European Council adopts the Europe 2020 Strategy. The EGF, discussed in separate (PPE, PSE) written explanations (see section 5.4), is an old initiative (first approved in 2006) that now comes to reinforce the Strategy. The EU describes Europe 2020 as a mechanism to outline "how the EU should chart its way out of the crisis while building a new economic model" (see http://europa.eu/about-eu/eu-history/2010-today/2010/index_en.htm). This wording is of relevance. Referring to 2008's crisis, David Kotz (2009, 315) insists that "[t]he evidence suggests that we are seeing more than just a severe financial crisis and a severe recession. We are witnessing a crisis of the neoliberal form of capitalism." Neoliberalism is undoubtedly (and more than ever before) at the centre of today's global "social imaginary", to the extent that

other "social myths" seem to have vanished from the system. Main Neoliberal premises (see Kotz 2009, 307) are shared as common sense throughout the world. In Europe, they have heavily permeated the institutional-cum-economic/social imaginary. Now the EU is joining the list of places witnessing some of the (first?) signs of Neoliberal exhaustion and exhaustion with Neoliberalism; hence, the need for "building a new economic model".

However, this exhaustion is far from definite. The two written documents analysed above may show some sort of (PPE and PSE) confrontation between signifiers—apart from other differences (see section 5.4), notice that the PPE MEP poses greater emphasis on the institutional component of the EU imaginary, whereas the PSE speaker directs listeners' attention to the economic/social pairing. However, both politicians show full coincidence in the main voting result. Whether signifying confrontation reveals the beginning of the fall of Neoliberalism or whether confrontation is just a way for the latter to readjust and grow stronger (while achieving its intended result [a yes vote] at each stage) is for the future to tell.

6. Conclusion

The present paper shows that a CADS synergy of (C)DA premises and CL tools provides useful to unveil the signifying dynamism within the EP and to take the research from the micro-levels of text to the macro-context. Nevertheless, the analyses we offer here are only a very modest contribution, performing the first turns of the screw of a much more laborious task. Further research and triangulation—from, notably, CDA realms—is required to confirm, enhance, or refute the results discussed.

Finally, we would like to state that the open-access ECPC Archive may be a useful tool for future studies. Not only does it cover a wide timespan of (European Parliament, Congreso de los Diputados, and House of Commons) parliamentary sessions, it also contains various XML-tagged sub-corpora that allow more detailed analysis of finergrained (sociolinguistic) aspects of parliamentary realities.

References

- Baker, Paul. 2006. Using Corpora in Discourse Analysis. London: Continuum.
- ———. 2010. *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Baker, Paul, Costas Gabrielatos, Majid Khosravinik, Michal Krzyżanowski, Tony McEnery, and Ruth Wodak. 2008. "A Useful Methodological Synergy? Combining Critical Discourse Analysis and Corpus Linguistics to Examine Discourses of Refugees and Asylum Seekers in the UK Press." *Discourse & Society* 19 (3): 273–305.
- Baker, Paul, and Anthony McEnery. 2005. "A Corpus-Based Approach to Discourses of Refugees and Asylum Seekers in UN and Newspaper Texts." *Journal of Language and Politics* 4 (2): 197–226.
- Bayley, Paul, ed. 2004. *Cross-Cultural Perspectives on Parliamentary Discourse*. Amsterdam: Benjamins.
- Bayley, Paul, Cinzia Bevitori, and Elisabetta Zoni. 2004. "Threat and Fear in Parliamentary Debates in Britain, Germany and Italy." In *Cross-Cultural Perspectives on Parliamentary Discourse*, ed. by Paul Bayley, 185–236. Amsterdam: Benjamins.
- Bayley, Paul, and Félix San Vicente. 2004. "Ways of Talking about Work in Parliamentary Discourse in Britain and Spain." In *Cross-Cultural Perspectives on Parliamentary Discourse*, ed. by Paul Bayley, 237–69. Amsterdam: Benjamins.
- Balcerowicz, Leszek, Andrzej Rzonca, Lech Kalina, and Aleksander Łaszek. 2013. *Economic Growth in the European Union*. Brussels: The Lisbon Council.
- Bevitori, Cinzia. 2004. "Negotiating Conflict: Interruptions in British and Italian Parliamentary Debates." In *Cross-Cultural Perspectives on Parliamentary Discourse*, ed. by Paul Bayley, 87–109. Amsterdam: Benjamins.
- Biber, Douglas, Conrad, Susan, and Reppen, Randi. 1998. *Corpus Linguistics: Investigating Language Structure and Use.* Cambridge: Cambridge University Press.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *The Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.
- Chilton, Paul, ed. 2002. Journal of Language and Politics, 2:1.
- "Consolidated Version of the Treaty on European Union." 2008. OJ C115/13.

- Constantinescu, Andrei, and Liliana-Aurora Constantinescu. 2013. "The Global Crisis Financial and Economic Crisis and the Main Priorities of the European Union for 2020." *Bulletin of the Transilvania University of Brasov* 6 (55).1: 171–174.
- Copeland, Paul, and Scott James. 2011. "From Lisbon Strategy to Europe 2020: The Impact of the Financial Crisis on the Governance and Discourse of Economic Reform." *UACES 41st Annual Conference*: 1–16.
- Dibattista, Denise. 2004. "Legitimising and Informative Discourse in the Kosovo Debates in the British House of Commons and the Italian Chamber of Deputies." In *Cross-Cultural Perspectives on Parliamentary Discourse*, ed. by Paul Bayley, 151–84. Amsterdam: Benjamins.
- European Commission. 2005. Commission Work Programme for 2005. Accessed September 8, 2015. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2005:0015:FIN:EN:PDF.
- European Parliament. 2015. "European Globalisation Adjustment Fund." Accessed September 8, 2015.
 - $http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_A TA(2015)549016..$
- Eurostat. 2005. *Eurostat Yearbook 2005*. Accessed September 8, 2015. http://ec.europa.eu/eurostat/web/products-statistical-books/-/KS-CD-05-001.
- Garzone, Giuliana, and Francesca Santulli. 2004. "What Can Corpus Do for Critical Discourse Analysis?" In *Corpora and Discourse*, ed. by Alan Partington, John Morley, and Louann Haarman, 351–68. Bern: Peter Lang.
- Hoey, Michael. 2005. *Lexical Priming: A New Theory of Words and Language*. London: Routledge.
- Ilie, Cornelia. 2010. European Parliaments under Scrutiny. Amsterdam: Benjamins.
- Koller, Veronika, and Gerlinde Mautner. 2004. "Computer Applications in Critical Discourse Analysis." In *Applying English Grammar: Functional and Corpus Approaches*, ed. by Caroline Coffin, Ann Hewings, and Kieran O'Halloran, 216–28. London: Arnold.
- Kopaczyk, Joanna. 2012. "Applications of the Lexical Bundles Method in Historical Corpus Research." In *Corpus Data across Languages and Disciplines*, ed. by Piotr Cap, 83–95. Frankfurt: Peter Lang.
- Kotz, David M. "The Financial and Economic Crisis of 2008: A Systemic Crisis of Neoliberal Capitalism." *Review of Radical Political Economics* 41.3 (2009): 305–317.
- Laclau, Ernesto and Mouffe, Chantal. 1985/2001. *Hegemony and Socialist Strategy. Towards a Radical Democratic Politics*. London: Verso.
- Mehan, Hugh. 1997. "The Discourse of the Illegal Immigration Debate: A Case Study in the Politics of Representation." *Discourse & Society* 8 (2): 249–70.
- Montesano Montessori, Nicolina. 2014. "The Potential of Narrative Strategies in the Discursive Construction of Hegemonic Positions and Social Change." In *From Text to Political Positions: Text Analysis across Disciplines*, ed. by Bertie Kaal, 171–87. Amsterdam: Benjamins.
- O'Halloran, Kieran, and Caroline Coffin. 2004. "Checking Overinterpretation and Underinterpretation: Help from Corpora in Critical Linguistics." In *Applying English Grammar: Functional and Corpus Approaches*, ed. by Caroline Coffin, Ann Hewings, and Kieran O'Halloran, 257–97. London: Arnold.

- Orpin, Debora. 2005. "Corpus Linguistics and Critical Discourse Analysis: Examining the Ideology of Sleaze." *International Journal of Corpus Linguistics* 10: 1: 37–61.
- Partington, Alan, ed. 2013. Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS). Amsterdam: Benjamins.
- Stubbs, Michael. 1996. Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture. Oxford, Blackwell.
- Svartvik, Jan. 2007. "Corpus Linguistics 25+ Year On." *In Corpus Linguistics 25 Years On*, ed. by Roberta Facchinetti, 11–26. Amsterdam: Rodopi.
- Van Dijk, Teun A. 2000. "On the Analysis of Parliamentary Debates on Immigration." In *The Semiotics of Racism. Approaches to Critical Discourse Analysis*, ed. by Martin Reisigl and Ruth Wodak, 85–103. Vienna: Passagen Verlag.
- Van Dijk, Teun A. 2002. "Political Discourse and Political Cognition." In *Politics as Text and Talk: Analytic Approaches to Political Discourse*, ed. by Paul A. Chilton and Christina Schäffner, 203–38. Amsterdam: Benjamins.
- Van Dijk, Teun A. 2010. *Discourse and Context A Sociocognitive Approach*. Cambridge: Cambridge University Press.
- Vasta, Nicoleta. 2004. "Consent and Dissent in British and Italian Parliamentary Debates on the 1998 Gulf Crisis." In *Cross-Cultural Perspectives on Parliamentary Discourse*, ed. by Paul Bayley, 111–49. Amsterdam: Benjamins.
- Wodak, Ruth, and Teun A. van Dijk (eds.). 2000. Racism at the Top Parliamentary Discourses on Ethnic Issues in Six European States. Klagenfurt: Drava Verlag.

Appendix 1

WRITTEN EXPLANATIONS OF VOTE

Paulo Rangel (PPE)—The request submitted by the Netherlands for assistance under the European Globalisation Adjustment Fund (EGF) in relation to 821 redundancies from 70 companies operating in the NACE Revision 2 Division 18 (printing and reproduction of recorded media) in the two contiguous NUTS II regions of Nord Brabant and Zuid Holland fulfils all the legally established eligibility criteria

In effect, under Regulation (EC) No 546/2009 of the European Parliament and of the Council of 18 June 2009 amending Regulation (EC) No 1927/2006 on establishing the European Globalisation Adjustment Fund, the scope of the EGF was temporarily widened to include its intervention in situations like this, in which, as a direct result of the global economic and financial crisis, there are 'at least 500 redundancies over a period of nine months, particularly in small or medium-sized enterprises, in a NACE 2 division in one region or two contiguous regions at NUTS II level'. I therefore voted in favour of this resolution, and I hope that the mobilisation of the EGF will contribute to the successful integration of these workers into the labour market

Silvia-Adriana Ţicău (PSE)—I voted for the European Parliament resolution on the mobilisation of the EGF for granting aid to the redundant workers. In December 2009, the Netherlands submitted a request for assistance to use the European Globalisation Adjustment Fund in connection with the redundancies which were made in eight regions, in enterprises operating in the graphics sector. The application concerns 821 redundancies made in 70 enterprises involved in printing and the reproduction of recorded media. The

redundancies were made in the period between 1 April and 29 December 2009 in the two contiguous regions of Nord Brabant and Zuid Holland.

The economic and financial crisis has also caused a drop in demand in the printing and publishing sector of about 32% for printed advertising material and of between 7.5% and 18.2% for magazines and newspapers. The printing and publishing industry in the Netherlands went through a major restructuring process in order to be able to continue to compete with similar sectors in Turkey, China, and India. I believe that the procedure for allocating these funds must be simplified to facilitate the affected enterprises' access to the EGF.

Author's address

María Calzada Pérez Departamento Traducción y Comunicación Universitat Jaume I calzada@uji.es

About the author

María Calzada-Pérez is Full Professor of Translation Studies at the Universitat Jaume I (Castellón, Spain). Her research focuses on corpus-based translation studies, institutional translation (especially translation at the European Parliament), ideology, and advertising. She has published extensively on these topics. She is Principal Coordinator of the ECPC (European Comparable and Parallel Corpora of Parliamentary Speeches) research group.

_

ⁱ This paper is funded by the Spanish Ministry for the Economy and Competitiveness (FFI2012-39289) and the Universitat Jaume I (P1.1B2012-64).

ii Since this paper resorts to no tagged data, the XML format is not described. Suffice it to say that metatextual tagging encodes speakers' profiles, enabling detailed linguistic research.