Using corpus linguistic software in the extraction of news frames: towards a dynamic process of frame analysis in journalistic texts

Abstract
The article seeks to demonstrate the contribution that corpus linguistic software can make in news frame analysis and how it can help address some of the methodological challenges in the extraction of news frames. With the employment of corpus linguistic software WordSmith, the study conducts an inductive frame analysis of the UK press coverage of the Greek financial crisis in which it combines principles of qualitative and quantitative analytical approaches. The findings demonstrate that, by integrating a statistical measuring mechanism in a qualitative analytical approach, corpus linguistic techniques can offer a systematic connection of stylistic and ideological features of news content and a more reliable identification of the loci for frames. Such techniques can also allow a better approximation of the unconscious level of frame construction. This can lead to a more efficient identification of frames that exhibit deeper cultural values and are more likely to shape the receivers' interpretations.

Introduction
Since the early 1990s there has been a steady growth in the use of frame analysis in research about news and journalism, in an effort to offer insight into the forces that shape media interpretations of reality and their potential influence on audiences. The roots of framing as a theory are situated in the field of sociology, where the term has been in use since the mid-1950s (Bateson, 1955). It is this sociological approach to framing that has underlay the study of news frames so far, with frames being examined as social constructs and outcomes of journalistic norms or organizational constraints, as well as sponsored by social and political actors. Goffman's (1974) prominent formative work has been pivotal in this field. Goffman defines frames as 'schemata of interpretation' that enable individuals to understand certain events and 'to locate, perceive, identify and label' occurrences. He calls these schemata primary frameworks because 'they turn what should be a meaningless aspect of a scene into something meaningful' (p. 21-22). An important implication for media and communication research is that the coverage of events in the news media depends on the frameworks employed by journalists (Scheufele, 2000). The potential impact of such frameworks on audiences' cognitive processing of information has triggered extensive investigation of the possible link between news coverage and the individuals' interpretation of events. It has also led to the development of different theoretical and methodological approaches to the analysis of frames. These attempts have opened up an on-going debate regarding notions of 'frame' and the methodological tensions created by its operationalization.

This article presents an experiment that attempts to address some of the challenges in the current methodological approaches, with the use of the corpus linguistic software WordSmith Tools. In response to calls for frame analysis to be a dynamic process that requires the bridging of qualitative and quantitative analytical approaches, this article suggests that computer-assisted qualitative data analysis offers a useful instrument for a more systematic extraction of frames by using software generated content as a guide for the analyst's interpretations. For the purposes of this experiment, the press coverage of the Greek financial crisis is used as case study.

We begin with a review of some of the most prominent studies that have employed frame analysis as a methodology. We identify the main questions that the various approaches present and the ways in which corpus linguistic techniques can help address them. This is followed by the description of our research design and the main findings of our analysis. We conclude with a review of the strengths and limitations of our study and suggestions for further research.

Methodological approaches to framing

Framing or frame analysis has its roots in sociology, psychology and linguistics and is used also in other disciplines, including politics and media studies. Gamson and Modigliani (1989) have offered one of the most widely-cited definitions of framing, describing it as 'a central organizing idea or story line that provides meaning to an unfolding strip of events ... The frame suggests what the controversy is about, the essence of the issue' (p. 143). Moreover, frames represent an emphasis or salience of different aspects of a topic; they are meant to simplify complex issues by lending more weight to certain considerations and activate schemas that encourage audiences to think in a particular way (Chong and Druckman, 2007; Gross and D'Ambrossio, 2004; Iyengar and Simon, 1993).

In the area of media and communications, what makes framing a useful approach, theoretically and methodologically, is its relevance to news construction and [audience] reception. As frames can be located at different levels, including public discourses, news production practices and audiences' information processing, an analysis of news frames can offer insight in the choices and interpretations journalists make when framing a story, which can ultimately define the nature of the debate and suggest to audience members how an issue can be interpreted (Van Gorp, 2010).

Studies that have used frame analysis as methodology have drawn from various theoretical and methodological approaches, leading to working definitions that are often tentative and specific to the purposes of the respective study. This has resulted in relatively little consensus regarding how frames can be operationalized and detected in the news. Moreover, the complexity of frames as 'part of a collective struggle over meaning that takes place through a multiplicity of media and interpersonal communication [and] draws from a range of resources' (Vliegenthart & van Zoonen, 2011:112) has generated questions about the process of using frames as methodology. Nevertheless, frame analysis can be usefully employed in many different ways, without losing its theoretical and methodological value in bridging news production processes with potential news effects on public perceptions (D'Angelo, 2002).

Methodological approaches to frame analysis can be broadly divided in two categories, namely deductive and inductive. Deductive approaches rely on predetermined framing categories that are brought in contact with the news stories. Such methods typically follow the logic of conventional quantitative content analysis and focus on measuring the frequency with which certain frames occur in a given text. One of the most representative deductive analyses was conducted by Semetko and Valkenburg (2000), who examined the news framing of European politics. Their analysis was based on five frames - namely conflict, human interest, economic consequences, morality, and responsibility - the frequency of which was measured through a series of questions to which the coder had to answer ''yes'' or ''no. A similar logic is found in the conceptual definition of framing offered by Snow and Benford (1988), which several other analysts have sought to replicate. A common strategy employed in these cases is the use of pre-existing theoretical and empirical knowledge based on which the deductive framing categories are devised. The reliance on already established frames may enable a fairly systematic analysis

of news content; however the specific method also suffers from a certain degree of inflexibility and risks missing important and newly emerging frames that may not be included in the predetermined categories (Matthes and Kohring, 2008:262-263). Other deductive approaches to framing analysis rely on coding for manifest indicators, such as key words (Entman, 1991) or framing mechanisms (Tankard, 2001). A more consistent approach to this kind of analysis has been developed by Matthes and Kohring (2008) who measured media frames through the clustering of empirical indicators. However, as Scheufele and Scheufele (2010:122) explain, these clustering techniques do not provide a tool for coding latent content, compromising the validity of the analysis.

Some of the shortcomings of the deductive approach can be tackled by inductive methods. Such methods could vary from those that are open and qualitative and extract frames through thorough interpretations of the text in connection to a given cultural context (Boni, 2002; Coleman & Dysart, 2005; Fair and Astroff, 1991; Tucker, 1998); to the more systematic ones that focus on language and analyse the structure of specific words and sentences in a text (e.g., Entman, 1991; Esser & D'Angelo, 2003; Pan & Kosicki, 1993). A commonly used approach starts with loosely theoretically defined frame categories, which serve as guidance for the extraction of more specific frames, through a grounded analysis that aims to identify all the possible frames (Gamson, 1992). Such inductive framing studies have also used CAQDAS (Computer-assisted qualitative data analysis software) including such programs as ATLAS.ti (Trenz, 2004), MAXqda/winMAX (Downey & Koenig, 2006; van de Steeg et al., 2003), and NUD*IST (Reese et al 2001), for example, to assist with qualitative coding.

In other studies, analysts generate framing categories through a qualitative analysis of a sample of the text, which they then develop into a codebook. The codebook is then applied to a larger population of texts, in line with the principles of quantitative content analysis (Meyer, 1995; Simon and Xenos, 2000). A frequent problem in these cases is the arbitrary and unsystematic nature of the analysis (Tankard, 2001), as the criteria that guide the extraction of the frames remain vague, depending on the interpretations of the researcher. To a certain extent, this subjectivity can be overcome through a systematic identification of linguistic elements and structural dimensions that denote the frame, such as the syntax, script, theme, and rhetoric (Pan and Kosicki, 1993). However, apart from the fact that this approach makes the analysis of large samples difficult, it also focuses on the manifest attributes of the text without answering questions of how these attributes of texts are clustered around an argument or frame (Reese, 2007; Matthes & Kohring, 2008)

Amongst the most systematic approaches are those based on computer assisted analysis, such as the 'frame mapping' conducted by Miller et al. (1997). Their automatic frame identification method employs the VBPro family of computer programs for content analysis and is based on the co-occurrence of key terms, identified with the help of cluster algorithms. In this case, frames are extracted through the links that exist between words and which may be present in some texts but not in others. A key question permeating these approaches is the tendency to reduce frames to either subjective interpretations of text; or clusters of words, attributes and story topics, but without explaining how all these elements are configured and associated with specific issues. If frames are more about the organisation and structuring of these elements into a larger story than the manifest content, frame analysis should be a more dynamic process of negotiating meaning that requires the bridging of qualitative and quantitative approaches as well as empirical and interpretative ones (Reese, 2007). The importance of culture, media genre and other contextual factors in identifying and analysing frames highlights the need for the human

interpretation that will integrate context in the analysis, something that linguistic and computer based approaches are unable to perform. In the case of news, frames emerge from complex processes that include the frames of different actors, shaped by their own ideological principles and institutional roles. They are also shaped by journalists' own ideological predispositions and professional practices as well as their emphasis on causal reasoning (Gamson and Modigliani, 1989; Iyengar and Simon, 1993; Pan and Kosicki, 2001). It is based on such processes that frame analysis in the context of news should be conducted. At the same time, the need for empirical indicators, such as words, should not be neglected if we want to remove some of the subjectivity of frame analysis. In this paper, we aim to take these assumptions further and explore the role of corpus linguistic techniques in establishing the words and categories around which particular frames evolve.

Our analysis focuses on the Greek financial crisis, therefore it falls under conceptualisations of framing that pertain to a specific issue under investigation. Such frames, called issue frames, tend to represent broader key concepts that become more prominent in the text. Issue frames are different to the more generic frames that could apply to various topics, such as the strategic news framing proposed by Cappella and Jamieson (1997) or the episodic/thematic frame binary introduced by Iyengar (1991).

The contribution of corpus linguistic techniques
With the use of software developed for corpus-assisted discourse studies such as Wordmith (Scott, 2011-- used in this study) or Antconc (Anthony, 2005) for example, we aim to develop a methodological tool that reconciles the gap between the interpretative and the attribute-based approaches to frame identification. We follow Entman's suggestion that frames can be detected by searching for the presence or absence of certain keywords, which can function as an indicator of the conscious or unconscious judgement made by the communicator (1993:53). In this case, keywords can be central to the meaning of the text without occurring very often (Hertog and McLeod, 2001: 152) as their salience is established through their association with certain cultural values.

We argue that the use of two corpus linguistic techniques—keywords and concordances-- developed for processing large volumes of data (Sinclair 1991; Baker, 2006) has particular advantages for the process of frame identification and analysis (Koteyko et al, 2008). Keywords are used to compare the *relative* frequency of words in any corpus with reference to another corpus (Scott, 1997; 2011). Here keyness is defined as "a quality words may have in a given text or set of texts, suggesting that they are important, [that] they reflect what the text is really about" (Scott and Tribble 2006: 73). Keyness is measured through log-likelihood, which relies on taking into account the number of instances of a word in the given corpus (compiled for a specific study) and comparing it to the number of instances in a reference corpus, which is representative of some norm (for example, the British National Corpus (BNC) comprising 100 million words used in spoken and written language) This strategy helps identify unusually frequent words that characterise the text as indicators of perspectives by which issues can be discussed and interpreted. In essence, this type of keyness is more likely to enclose the unconscious and culturally driven judgment of the communicator; and offers a useful instrument for identifying words that are central to the meaning of a given text in a more evocative way that conventional frequency of occurrence. Concordances constitute a window of text surrounding a search term and allow the analyst to shift backward and forward between instances of a word use, and in this

case a keyword use, in its immediate textual surrounding and the whole texts themselves, (which, in WordSmith Tools, are accessible through the Viewtext function).

Although keywords will not readily reveal frames, they will direct the analyst to important concepts in a text which may help 'diagnose' and 'nominate' central ideas around which the frame is constructed. By identifying words whose frequency is unusually high in comparison with some norm, the 'keyword' function can elicit words that can work as indicators of emphasis and that concentrate meanings that are drawn from cultural values, political discourses as well as journalistic norms. Examining how such keywords are used through concordances and whole texts and what framing and reasoning devices they may have in common should therefore be revealing.

Moreover, with the use of software, these central ideas or emphasis words will be extracted empirically rather than personally or experientially, removing some of the subjectivity that human judgement entails. Although we will still rely on human interpretation to identify the larger tale and the broader schemas in which these words are woven, the software will enable a more reliable and valid combination of quantitative and qualitative as well as empirical and interpretative examination, through which we seek to develop a more systematic analytical instrument. The value of the software becomes more explicit if we compare this analytical approach to a manual coding and analysis. The keyword function offers a systematic identification of the central meanings contained in the text while the concordance function enables an automated collection of the text extracts where the frame is more likely to be hidden. In this way, the dataset can be compiled much faster than would be possible to do manually. The use of keywords as the starting point also encourages a reflexive process that combines quantitative and qualitative tools in a similar way as grounded theory methods. In essence, as the statistically calculated keyness determines where to go and what to look for, it allows analysis and data collection to inform one another. In this sense, our study approaches frame analysis as a set of rigorous research procedures leading to the emergence of conceptual categories, i.e. the frames. These categories are related to each other as a theoretical explanation of how the news content may have been constructed and with what possible implications for the audiences.

Research design

For our analysis, we have used the Greek financial crisis as a case study and the coverage it received by the UK press, specifically the Guardian. A frame analysis of the Greek financial crisis, as it was covered by the UK press, was conducted by Touri and Rogers (2013) who employed a deductive approach with the use of pre-defined frame categories. In essence, how the story was told was measured through interpretations of the crisis that were fairly broad and culturally neutral. The analysis focused on the emphasis promoted through the frequency of occurrence of specific interpretations including the problem definition, causes of the crisis, attribution of responsibility, consequences and solution; while each of these frames contained sub-frames that embodied different aspects emphasised within the frame (ibid: 180). The analytical categories used in this codebook were constructed after a careful review of existing framing theories and theories about the media's role in European politics; and aimed to offer insight into how existing theoretical considerations regarding the UK media's role in European politics are understood in the context of the Greek financial crisis. This tight connection with existing theoretical knowledge renders the frames used in this study a particularly useful basis for comparative evaluation.

Our sample was collected through Nexis, a daily updated electronic database for UK and international newspapers, using the search terms 'Europe', 'EU', 'euro', 'Greece', 'economy', 'debt' and 'crisis'. It consisted of 40 news stories that were featured in the Guardian only, from January 2010 when the first signs of the crisis surfaced, to August 2010 three months after the first rescue package was agreed. This is a purposeful sample, selected in line with the aims of this pilot study. The Guardian accounted for the largest share of stories and foreign correspondents covering the crisis in the UK press during the selected period. It also exhibited a wider range of frames and an effort to explain the crisis in its national and European structure even from its early stages (Touri and Rogers, 2013).

Analytical framework

In principle, we followed the procedure of the inductive analysis suggested by Van Gorp (2010; 2012). With the use of the frame matrix, VanGorp's is one of the most systematic qualitative approaches to frame analysis. We therefore, integrated our keywords and concordances into the inductive process suggested by Van Gorp and made them part of the frame matrix construction.

One of the main contributions of a corpus linguistic software such as Wordsmith Tools is that it prepares the ground and 'diagnoses' the data based on which the frames are identified. Hence, our analysis is primarily data driven. As a first step, we produced a list of keywords using the Guardian word list[1] (the time span of the articles included into this reference corpus is between 1994 and 2004) and a wordlist generated specifically for our sample/the corpus in question. The Keywords function of the WordSmith Tools software was used to compare the contents of these two lists (Scott, 2011)[2]. Those words that turned out to be unusually frequent in our corpus compared to the frequency of the same word type in the reference corpus of Guardian articles were displayed as a list. The more statistically significant an item is, the more *key* it is (its keyness value is higher), and the higher it is placed on this list (Partington 2010). After lemmatisation and removal of grammatical words and proper names[3], the keywords were grouped into four broadly defined semantic sets in the order of decreasing salience: 'economy' (including words such as austerity, fiscal, monetary, financial), 'geography' (including words that referred to geographical locations such as Athens, Germany, Berlin, Portugal, Spain, Europe), 'politics' (including words such as government, political, state, sovereignty) and 'other/unsorted'. The keywords were then concordanced to obtain a first glimpse of potential frames. The next step therefore consisted of compiling a datasheet with 316 text extracts based on the concordances generated for the selected keyword groups. These text extracts comprised the units of analysis and constituted the locus of framing and reasoning devices. Table 1 presents a snapshot of the concordances computed for keyword 'austerity' depicting the immediate text

---

[1] The list is provided by Mike Scott on the WordSmith Tools software site http://www.lexically.net/wordsmith. We also used the BNC wordlist, also available on this site, which did not yield significant differences in the output of keywords.

[2] The 'keyness' of a lexical item is computed taking into account the following (a) its frequency in the corpus in question, (b) the number of running words in the corpus in question, (c) its frequency in the reference (comparator) corpus, and (d) the number of running words in the reference corpus. Then the program cross-tabulates these, and applies Log Likelihood test (see Scott 2011: Help Menu).

[3] Researchers classify software-generated keywords into three types: proper nouns, other 'lexical' words, and 'grammatical' words (Scott 1999). The common strategy in analysing the keyword list is to discard proper nouns and grammatical words because these tend to reveal only dramatis personae or obvious content and grammatical features, and focus on remaining 'lexical' items. Such words indicate the 'aboutness' of a text or corpus, and are thus "generally those which are most interesting to analyse" (Baker 2006: 127).

surrounding the keyword. In order to extract meaningful frames, the text extracts used for the analysis comprised the entire sentence containing the keyword as well as the preceding and following sentence.

[Table 1 near here]

The identification of keywords and the concordances extracted based on these keywords were the two most crucial stages in our analysis and what sets our approach apart from the traditional types of frame analysis. One of the biggest challenges that analysts face is detecting the frame in a given text. The different types of framing devices and other working definitions that have been developed over the years have offered valuable guidance; but are often tentative and specific to the purposes of the respective study. The identification of keywords that reflect the 'aboutness' of the text provides a statistical method to determine where the frame could be located in this text. Moreover, these statistically determined keywords encapsulate perspectives by which issues can be discussed and interpreted. Hence, they function as a connecting link between the extracts of text that were selected for analysis and the cultural notions that are more likely to have informed the construction of frames. In essence, this method enables us to statistically determine the text extracts generated through the concordances; and achieve a more focused detection of the places where the relevant cultural notions and frames were likely to be located.

Each text extract was coded for the framing and reasoning devices it contained, through an open coding process. Traditionally, open coding does not rely on predefined coding categories; rather, the codes are expected to emerge from the content itself. In this process, framing devices represented key terms and manifest elements that came from the material itself and can function as demonstrable indicators of the frame. Reasoning devices are those elements that are not explicitly included in the message but are likely to come up in the interpretation of the message as they reflect a thought process. Both types of devices were deemed capable of invoking a framework of thought at the audiences' cognitive level (VanGorp, 2012: 1275). The open coding was followed by an axial coding during which we ran through the data for a second time in order to identify causal patterns and relationships between and within the coding categories identified in the first stage of analysis. We identified similarities, differences and contrasts between the various devices and reduced them to broader meanings or dimensions. This enabled us to reduce the initial codes to fewer and more significant codes that represented bigger archetypal ideas, which we called frames. This stage led to the creation of the main frame packages that represented the frames. The analysis produced ten main frames, based on which the frame packages were constructed, as shown in table 2.

In the construction of the frame packages, we followed the process adopted by VanGorp (2012) in his frame analysis of dementia. The frame packages were created through the identification of a logical chain of framing and reasoning devices that appeared to convey a coherent overarching idea. The packages are composed of a general definition of the problem or issue; the reasons why this is a problem or else, an explanation of the issue in question; the consequences and solution or action associated with the issue in question; and lastly, the moral values involved. Each row in the matrix represents a frame package, while the columns show the various framing and reasoning devices that make up the frame. The rows are logically integrated clusters of framing and reasoning devices that refer to the same overarching idea that represents

the main frame. In other words, a frame package shows how the idea that the Euro is an unrealistic project (frame 1) is manifested and understood through different dimensions of the frame, i.e. as a consequence of the Greek financial crisis or through the action that needs to be taken in order to address this problem.

Through the concordances we run for these statistically identified keywords, we were able to identify repeated extracts that contained more than one keyword. As each keyword reflected a statistically significant topic, they functioned as useful indicators of the different meanings and interpretations that were hidden in a given text. This enabled us to reveal additional frames that could be overlooked in a standard content analysis procedure. It also helped reveal the repetition and cohesion at work in framing processes. The three stages of coding that led to frame packages was conducted by the two authors and a third coder independently. The three coders would regularly meet and discuss the process so that discrepancies were examined and reconciled.

[Table 2 near here]

Findings
Qualitative analyses of this type do not normally allow for an estimate of the frequency of occurrence of each frame. However, with the use of keywords, our approach allows us to estimate the prominence of frames based on the breadth and density of the networks of keywords they are associated with. More specifically, frames 1 and 2 were the most prominent as they emerged from a rather wide network of keywords from the category 'economy', from the concordances computed for these keywords and from the framing and reasoning devices that were identified from the respective text extracts. Quite often, these two frames were also combined in the same text extract as the journalist attempted to explain the underlying causes of the crisis. From another perspective, frame 5 can also be described as a prominent one due to the heavy concentration of framing and reasoning devices on keywords specifically from the category 'politics', such as the terms 'political' and 'policy'. Along similar lines, it was the keyword 'Germany' that produced the majority of the texts and framing/reasoning devices that made up frame 8. The different ways in which frame prominence can be assessed through the keywords demonstrates different ways in which the frame may make a long term effect. It could either be present throughout the news story/newspaper but also intersecting with other frames that may be present in the same texts; or it could be heavily prominent in specific only texts.

The potential effect of such frames is also assumed through their cultural resonance. The narratives reflected in frames 1 to 4 are in line with the scepticism that is traditionally expressed by politicians, the media and members of the public towards the European Union; and the general hostility toward the European project. This is partly because of the perceived threats posed by other cultures (McLaren, 2002). For instance, evidence provided by the Eurobarometer surveys of public opinion, portrays the British as the most Eurosceptic people in Europe. Moreover, unlike countries such as Italy, there was never an ideological commitment to European integration in the UK. Rather, it was the belief that the country would benefit from the unification that led the UK into the EU, which can explain why the support for the European project has remained low, considering perceptions of limited benefit. At the same time, frame 3 and the emphasis on the battlefield that Europe represents in this sample, is also in line with journalists' professional practices, since emphasis on conflict increases the newsworthiness of a story. In a similar way, conceptions of Germany as an economically strong country but also a

tough actor and villain resonate with historical narratives and memories from the Second World War.

Although less culturally bound, the Greek profligacy frame could be described as socially constructed. It reflects connotations of a corrupt South and is in line with discourses of a division between the South and North European countries. Such discourses are frequently featured in the media and are seen as a cause for public cynicism about European politics (DeVreese, 2007; Gavin, 2007). In relation to this, it is interesting to see the existence of a counter frame (frame 7) which sees the Greek government as honest and transparent. This was an autonomous frame that was used without reference to Greek profligacy and was also significantly less dominant. It could also be argued that, although it contrasts the Greek profligacy frame, it also complements the bigger story which sees the current Greek government as the one that has to deal with the corruption of its predecessor. A second counterframe refers to the role of Germany, as an actor that also suffers from this crisis (frame 9). The frame was autonomous also in this instance, manifesting a different side to the role of Germany in this crisis. Frame ten is another useful example, promoting an evaluation that goes beyond the simple fact that the crisis and the rescue package are expected to affect the living standards in Greece. Here, the drama reflected in lexical choices such as 'pain' and 'sacrifice' conveys a powerful narrative that could be easier for receivers to notice and remember.

Discussion

In this study, we sought to address some of the challenges in frame analysis. With the employment of corpus linguistics, we tried to remove some of the subjectivity that inductive, qualitative approaches are often associated with, without separating the frames from the cultural and media context that characterises the text. We argue that the use of corpus linguistic techniques has enabled a systematic and meticulous connection of empirical indicators, such as lexical choices, with salient issues identified by the journalists. This paved the way for a more objective and reliable association of manifest content with key meanings that journalists attributed, consciously or unconsciously, to the issues discussed in the selected news stories. The contribution of this methodological approach can be summarised as follows.

Selection of units of analysis: The use of keywords produced significant empirical indicators that were identified in statistical rather than subjective manner. With the use of concordances, these lexical choices enabled the focused and systematic selection of the text extracts for the analysis. This was a crucial stage in the analysis as it ensured that all the possible text extracts where a frame could be located were captured automatically, a process that would prove significantly lengthier if done manually. Moreover, as several keywords occurred together in the same extract, concordances led to a repetition of certain extracts. This repetition validated our selection as it increased the chances that a frame could be extracted from these citations. In effect, the use of keywords and concordances secured a more systematic and reliable identification of the loci for the frames by integrating a statistical measuring mechanism in a qualitative analytical approach.

Open and axial coding: open coding involves the construction of an inventory of framing and reasoning devices that may contribute to the audiences' interpretation of the text. This is the stage where human interpretation becomes important, as emphasis should be paid on how the story is told and what the key elements that could invoke a framework of thought are. Human interpretation is also characterised by a certain degree of subjectivity. In our analysis, this level of subjectivity was reduced through the strategic selection of text extracts guided by the measure

of keyness. The identification of extracts that focused on each keyword allowed for a more in-depth selection of framing and reasoning devices as the keywords enabled us to scrutinize the text more closely. Moreover, when more than one keyword occurred in the same text extract, we were able to approach the given extract from different perspectives and consider different reasoning devices and possible thought processes in line with the respective keyword. Hence, in several cases, the occurrence of two or more keywords in the same extract would induce the same reasoning device and reinforce it; while, in other cases, it would unpack the text into different meanings. This process also allowed for a more comprehensive and accurate extraction of codes, during the axial coding.

Measuring salience: by combining the identification of statistically significant lexical features through keyness with a qualitative concordance-based study, we were able to get closer to a potentially more unconscious level of frame construction and develop frame packages that exhibit deeper cultural values and tell a more complete story. Hence, while a classic content analysis such as the one used by Touri and Rogers (2013) would measure the frequency of occurrence of certain interpretations – i.e. the causes of the crisis and the attribution of responsibility - our frame packages reveal a bigger story constructed through the broader cultural and media context and the inevitable subjectivity of individual journalists. Hence, we see the media's attempt to evaluate rather than just describe situations, in a way that resonates with existing cultural schemata. Our inductive analysis also demonstrates a certain degree of overlap between our empirically extracted frames and the frames that made up the codebook designed by Touri and Rogers (ibid). The deductive categories designed by Touri and Rogers to measure EU policies as a cause of the crisis or the role of the EU and Greek government as responsible actors emerged in our frame packages. So did the possible consequences of the crisis and some treatment suggestions that became obvious through several frame packages. This overlap increases our confidence in the validity of our findings as they are in line with the relevant theoretical knowledge that informed Touri and Rogers' codebook. At the same time, our instrument has also coded the latent content that is shaped by the cultural context within which the news stories were constructed. In the case of the Greek financial crisis, this concerns the role of the UK in the European Union, well-established perceptions and attitudes towards European as well as historical narratives.

In terms of the broader contribution that this approach could make to frame analysis we could argue that although the main value of keyness lies in systematising the inductive analysis, this method can be useful for deductive approaches too. When the analyst uses a qualitative analysis to generate framing categories that will then develop into a codebook to be applied to a larger sample, tools like Wordsmith could increase the reliability of the process by adding the statistical element to the first stage of the coding. It should also be mentioned that keyness can drive the analysis towards themes as well as stylistic features. From this perspective, our approach is suited for emphasis frames that convey themes or ideas, as well as generic frames that refer to the way a story is told. In either case, keywords can function as the starting point for the identification of the unit of analysis. Following that, the analyst will need to design a coding method that reveals emphasis or generic frames. The method could also apply to texts that discuss more than one issue, since the keyness function will capture all those keywords that indicate what the text is about. Once the keywords are compiled, it is then up to the researcher to identify the themes and ideas that correspond to each issue.

As the analysis was conducted on a small sample of news stories from a pro-European newspaper, the findings are inevitably non-generalizable. Moreover, by following the norms and logic of a qualitative method, the analysis entails a certain degree of subjectivity, in the way the news content was interpreted. Further work is needed that utilises the software's capacity to apply to large datasets. Moreover, a comparison between the results we obtained in this analysis with the results from manual coding of the same material would enable a better evaluation of the effectiveness of this tool and the quality of the analysis. Along with different case studies, this will pave the way for more unbiased configuration of the different elements in the news content and their association with specific meanings.

References

Baker, P. (2006), *Using Corpora in Discourse Analysis*, London: Continuum

Bateson G (1955), 'A theory of play and fantasy', *AP: Psychiatric Research Reports,* 2: 39–51

Boni, F. (2002), 'Framing Media Masculinities Men's Lifestyle Magazines and the Biopolitics of the Male Body', *European Journal of Communication* 17(4): 465-478

Cappella, J. N., and Jamieson, K. H. (1997), *Spiral of cynicism. The press and the public good*, New York: Oxford University Press

Chong, D. and Druckman, J. N. (2007), 'Framing Theory', *Annual Review of Political Science*, 10 (1):103-126

Coleman, C.L. and Dysart, E.V. (2005), 'Framing of Kennewick Man Against the Backdrop of a Scientific and Cultural Controversy', *Science Communication* 27 (1): 1-24

D'Angelo, P. (2002), 'News Framing as a Multiparadigmatic Research Programme: A Response to Entman', *Journal of Communication*, 52(4): 870-88

de Vreese, C. H. (2007), 'A spiral of Euro-skepticism: The media's fault?', *Acta Politica. International Journal of Political Science*, 42 (2/3): 271-286

Entman, R. B. (1991) 'Framing US coverage of international news: Contrasts in narratives of the KAL and Iran air incidents' *Journal of Communication,* 41 (4):6–27.

Entman, R. M. (1993), 'Framing: Toward clarification of a fractured paradigm', *Journal of Communication*, 43 (4): 51–58

Entman, R. M. (2003), 'Cascading Activation: Contesting the White House's Frame after 9/11', *Political Communication,* 20 (4):415-32.

Esser, F. E., and D'Angelo, P. (2003), 'Framing the press and the publicity process: A content analysis of meta-coverage in Campaign 2000 network news', *American Behavioral Scientist* 46 (5): 617-641

Fair, E. and Astroff, R. (1991) `Constructing Race and Violence: US News Coverage and the Signifying Practices of Apartheid', *Journal of Communication* 41(4): 58-74.

Gamson, W, A., and Modigliani, A. (1989), 'Media discourse and public opinion on nuclear power: A constructionist approach' *American Journal of Sociology*, 95(1): 1-37.

Gamson, W. (1992), *Talking politics*, New York: Cambridge.

Gavin, N.T. (2007) *Media, Economy and Democracy: Press and Television in British Politics*, London: Palgrave

Goffman, E. (1974) *Frame Analysis: An Essay of the Organization of Experience,* New York: Harper and Row.

Gross, K. and D'Ambrosio, L. (2004), 'Framing emotional response' *Political Psychology*, 25 (1): 1–29

Hertog, J. K., and McLeod, D. M. (2001) 'A multiperspectival approach to framing analysis: A field guide', in S. D. Reese, O.H. Gandy, & A. E. Grant (Eds.), *Framing public life,* Mahwah, NJ: Lawrence Erlbaum, pp. 139–162

Iyengar, S. (1991) *Is anyone responsible*? Chicago: University of Chicago Press.

Iyengar, S. and Simon, A. (1993), 'News coverage of the Gulf crisis and public opinion. A study of agenda-setting, priming and framing', *Communication Research*, 20 (3):365-383

Koteyko, N., Nerlich, B., Crawford, P. and N. Wright (2008), "'Not rocket science' or 'no silver bullet'? Media and government discourses about MRSA and cleanliness", *Applied Linguistics*, 29 (2): 223-243

Matthes, J., and Kohring, M. (2008), 'The content analysis of media frames: Toward improving reliability and validity', *Journal of Communication*, 58 (2): 258-279

McLaren, L. (2002) 'Public support for the European Union: cost/benefit analysis or perceived cultural threat?' *The Journal of Politics,* 64 (2): 551–566

Meyer, D.S (1995), 'Framing national security: Elite Public Discourse on Nuclear Weapons During the Cold War', *Political Communication* 12 (2): 173-192.

Pan, Z. and Kosicki, G. M. (2001) 'Framing as a strategic action in public deliberation', in S. D. Reese, O. H. Gandy, Jr. & A. E. Grant. (Eds) *Framing Public Life. Perspectives on Media and Our Understanding of the Social World* , NJ: Lawrence Erlbaum Associates, pp. 35-66

Pan, Z., and Kosicki, G. (1993), 'Framing analysis: An approach to news discourse', *Political Communication*, 10 (1): 59–79.

Partington. A. (2010) 'Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) on UK newspapers: an overview of the project', *Corpora* 5(2): 83-108.

Reese, S. D. (2007) 'The Framing Project: A Bridging Model for Media Research Revisited', *Journal of Communication* 57(1): 148–54.

Scheufele, D. T. and Scheufele, D. A. (2010), 'Of spreading activation, applicability and schemas: Conceptual distinctions and their operational implications for measuring frames and framing effects' in P. D'Angelo, & J. Kuypers (Eds.), *Doing news framing analysis: Empirical and theoretical perspectives*, New York: Routledge, pp. 110-134

Scott, M. (1999) *WordSmith Tools 3*, Oxford: Oxford University Press.

Scott, M. (2011) *WordSmith Tools 6*, Oxford: Oxford University Press.

Semetko, H. A., and Valkenburg, P. M. (2000), 'Framing European politics. A content analysis of press and television news' *Journal of Communication*, 50(2): 93-109.

Simon, A., and Xenos, M. (2000) 'Media framing and effective public deliberation' *Political Communication,* 17 (4):363–376.

Sinclair, J. (1991) *Corpus, Concordance, Collocation*, Oxford: Oxford University Press.

Snow, D. and Benford, R. (1988) 'Ideology, frame resonance, and participant mobilization', *International social movement research*, 1(1):197-217.

Stubbs, M. (1996) *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*, Oxford: Blackwell.

Tankard, J. W. (2001) 'The empirical approach to the study of media framing, in S. D. Reese, O. H. Gandy & A. E. Grant (Eds.), *Framing public life,* Mahwah, NJ: Lawrence Erlbaum, pp. 95–106

Touri, M. and Rogers, S.L. (2013), 'Europe's Communication Deficit and the UK Press: Framing the Greek Financial Crisis', *Journal of Contemporary European Studies*, 21(2):175-189

Tucker, L. (1998) 'The framing of Calvin Klein: A frame analysis of media discourse about the August 1995 Calvin Klein jeans advertising campaign', *Critical Studies in Mass Communication* 15 (2): 141-157.

Van Gorp, B. (2010) 'Strategies to take subjectivity out of framing analysis', in P. D'Angelo, & J. Kuypers (Eds.), *Doing news framing analysis: Empirical and theoretical perspectives*, New York: Routledge, pp. 84- 109

Van Gorp, B. and Vercruysse, T. (2012), 'Frames and counter-frames giving meaning to dementia: A framing analysis of media content' *Social Science & Medicine*, 74: 1274- 1281

Vliegenthart, R. and van Zoonen, L. (2011), 'Power to the frame: Bringing sociology back to frame analysis', *European Journal of Communication*, 26(2): 101-116