Patterns of discourse semantics
A corpus-assisted study of financial crisis in British newspaper discourse in 2009

Melani Schröter and Petra Storjohann
University of Reading, UK / Institut für Deutsche Sprache, Mannheim

Corpus-assisted analyses of public discourse often focus on the level of the lexicon. This article argues in favour of corpus-assisted analyses of discourse, but also in favour of conceptualising salient lexical items in public discourse in a more determined way. It draws partly on non-Anglophone academic traditions in order to promote a conceptualisation of discourse keywords, thereby highlighting how their meaning is determined by their use in discourse contexts. It also argues in favour of emphasising the cognitive and epistemic dimensions of discourse-determined semantic structures. These points will be exemplified by means of a corpus-assisted, as well as a frame-based analysis of the discourse keyword financial crisis in British newspaper articles from 2009. Collocations of financial crisis are assigned to a generic matrix frame for ‘event’ which contains slots that specify possible statements about events. By looking at which slots are more, respectively less filled with collocates of financial crisis, we will trace semantic presence as well as absence, and thereby highlight the pragmatic dimensions of lexical semantics in public discourse. The article also advocates the suitability of discourse keyword analyses for systematic contrastive analyses of public/political discourse and for lexicographical projects that could serve to extend the insights drawn from corpus-guided approaches to discourse analysis.

Keywords: discourse keywords (DKW), discourse semantics, semantic presence/absence, frame-based contrastive analysis, corpus linguistics, public/political discourse

ISSN 1878-9714 / E-ISSN 1878-9722 © John Benjamins Publishing Company
1. Corpus-assisted analyses of discourse keywords

1.1 Critical Discourse Analysis and corpus linguistic approaches

(Critical) analyses of public/media discourse use a variety of linguistic concepts and methodological approaches, including concepts of pragmatics. It is a benefit of a ‘movement’ such as Critical Discourse Analysis (CDA; for the relation between Pragmatics and Critical Discourse Analysis, cf. Reisigl 2011) that it provides scope and methodological flexibility to borrow from different conceptual and methodological traditions, so as to test various combinations of these traditions and/or make possible adjustments (cf. Baker et al. 2008: 297; Baker 2010). Both the methodological approaches utilised in CDA frameworks and the analyses drawing on pragmatic concepts have leaned more towards a qualitative textual analysis and hermeneutic paradigms.

The present article attempts to further explore the potential for mutual benefits for CDA and corpus linguistics and a pragmatic exploration of lexical semantics in discourse contexts. It presents the results of a data-driven analysis of the English discourse keyword *financial crisis* in newspaper discourse. We believe that corpus linguistic approaches allow for successful combination with qualitative approaches and that the boundary between quantitative and qualitative approaches is blurred (as argued by Baker et al. 2008: 295ff.; Mautner 2012; Bubenhofer 2013). Using a corpus-assisted approach, we will employ the concept of frame (cf. Fillmore 1976; especially Konerding 1993) in line with suggestions by Ziem (2008a, b), to make sense of patterns of concordances and to relate linguistic structures to meaning and knowledge in discourses. In doing so, we realise that we are dealing primarily with the lexico-semantic side of discourse; compare:

> At the current state of play, and considering the limitations of those tools that are sufficiently widely available, there is a very strong bias in favour of the individual lexical item and clusters thereof. Put simply, ‘the word’ is the peg that everything else is hung on. (Mautner 2009: 124)

If Mautner is right, then there is a need for a pragmatic understanding of lexical semantics and in particular for a clearer conceptualisation of salient lexical items, insofar as their meaning is determined by their use in discourse. And even if ‘the word’ can help us unravel broader discourse semantic structures, we often seem to be dealing with one particular type of word, i.e. with lexical items having a certain degree of keyness (cf. Baker 2004). A lexico-semantic, corpus-assisted approach to (critical) discourse analysis could be made more precise by introducing the concept of ‘discourse keywords’ (see Section 1.2).
We agree with Teubert’s (2007, 2010) premise about meaning in discourse: “All that has been said about a discourse object contributes towards its meaning” (Teubert 2007: 68). This means that whatever has been said about e.g. the financial crisis is reflected in every instance it is referred to. It has been acknowledged also from a more lexicological than discourse analytical view that “the meaning of a word is (some kind of summation of) the conceptual content made accessible by the use of that word (as opposed to any other) in particular contexts” (Cruse 2000: 30); otherwise put, “there is no such thing as ‘the meaning of a word’ in isolation from particular contexts: decontextualization of meaning is variable, and in principle, always incomplete” (ibid.: 51). Similarly, Teubert (2007: 70) maintains that “[w]hat […] lexical words […] mean, is what we learn about them in the discourse”. The discourse we are dealing with here is that of the financial crisis in the British public sphere, as it is sustained to a large extent – though certainly not exclusively – by media discourse, in our case exemplified through newspaper articles. We collected texts that contained references to financial crisis as a definable and researchable subset of the overall, indefinite ‘discourse’ mentioned above. This approach builds on the assumed strength of the following links: that the key lexical item financial crisis, which labels the discourse, actually works as an index for finding texts in the related discourse about the financial crisis, and that it is also a node in the semantic network of the related discourse. Thus, when we study semantic scope through the use of financial crisis within the contexts of the related discourse, we can generate information about the latter, too.

1.2 The concept of Discourse Keyword

A dedicated conceptualisation of Discourse Keyword (henceforth DKW) does not play a distinct role in Anglophone academia (but see e.g. Stubbs 2001: 145ff., 2010; Wierzbicka 2006, 2010; Ayto 19991), but it is firmly incorporated into approaches to political discourse analysis in German linguistics,2 using predominantly qualitative hermeneutic approaches. It is understood as comprising:

1. Ayto (1999) is actually a documentation of new words that entered the English language in the 20th century and therefore allows insights into socio-political/economic change from a lexical perspective. However, the dictionary deals neither exclusively nor specifically with DKWs, and it also merely documents the words without analysis.

2. Here, the phenomenon is labelled ‘Schlagwort’ or ‘Schlüsselwort’. ‘Schlüsselwort’ literally translates as keyword, but does not commonly include the notion of ‘keyword’ as a tool for finding or looking up information, e.g. in database searches. This aspect is covered by ‘Schlagwort’; however, ‘Schlagwort’ also includes a meaning similar to ‘buzzword’, indicating its use as a rhetorical device for persuasive purposes.
– lexical items of notable frequency, at least in certain periods of the salience of its relative discourse (e.g. the frequency of financial crisis since 2008);
– lexical items that function as semantic node in texts and/or discourses and which exhibit a complex, discourse-determined semantic structure (such as globalisation);
– lexical items that are part of an ensemble\(^3\) of DKWs; there is not just one keyword that labels the related discourse (e.g. migration), but typically there are a number of such nodes, and the DKWs inhabiting them (e.g. multicultural society) often represent certain points of view (e.g. fortress Europe) or are established as a (counter)reaction to others (e.g. illegalised immigrants vs. illegal immigrants);
– lexical items that pertain to controversial issues – either controversial in labelling/conceptualisation (e.g. illegal(ised) immigrants) or controversial in evaluation (people differ in their evaluations as to whether or not multicultural society is something positive, but the label itself is not disputed).

This conceptualisation of DKWs therefore sits between the notion of ‘cultural key words’ (Stubbs 2010), including its conceptual history (following Brunner et al. 1972ff.), and the notion of statistical keyness of lexical items in corpus analytical contexts (e.g. Baker 2004; Bondi & Scott 2010). DKW research is interested in the discursive background and development of DKWs and in discourse-related functions of DKWs (e.g. babycaust as a stigmatising word used by anti-abortionists), mainly on the basis of political and media discourse. Conceptual history looks at broader cultural concepts such as nation, work, democracy, etc. diachronically and mainly in intellectual discourses. In common with corpus-analytical approaches, DKW research is interested in patterns of language use: frequencies of occurrence, co-occurrences, position in texts and intertextual semantic/pragmatic coherence.

Within German political discourse analysis, the DKW concept has triggered not only a number of proposals for more fine-grained typologies of DKWs and descriptions of the kinds of ‘semantic conflicts’ inhabiting them, but also DKW-based lexicographical projects; in addition, it has informed a number of lexicographical projects on the history of German public discourse (see esp. Strauß et al. 1989; Niehr 1993; Stötzl & Wengeler 1995; Felbick 2003; Schmitz-Berning 2000). In the past, moreover, DKWs were mostly analysed qualitatively and on the basis of what would now be considered small corpora of newspaper texts and/or political speeches, without corpus analysis software, mostly taking into account the following:

\(^3\) We use the term ‘ensemble’ deliberately and prefer it to e.g. ‘set.’ ‘Sets of DKWs’ would suggest similarity in quality of this number of DKWs whereas ‘ensemble’ emphasises that different DKWs perform different functions and play different roles in discourses, including ‘hero’ (freedom) and ‘villain’ (terrorism).
Patterns of discourse semantics

- frequency of occurrence of DKWs;
- occurrence of similar or opposing DKWs, i.e. the role of the DKW in an ensemble of DKWs that occur in the related discourse;
- context, esp. collocations or notable occurrence of DKWs in specific genres or party-specific usage and metaphorical conceptualisations occurring in/with the DKW in question;
- metalinguistic comments: instances where a DKW is commented on critically or affirmatively in terms of its appropriateness.

With regard to the study of DKWs in particular, a quantitative corpus analysis or "quantitatively informed qualitative" corpus-assisted analysis (Bubenhofer 2013) could significantly enhance the methodological profile of DKW analysis. We regard 'discourse' as a highly complex social, semiotic, cognitive, and linguistic phenomenon that cannot be reduced to the lexical level; put otherwise, discourse complexity cannot just be boiled down to words. Nevertheless, we consider DKWs to be central phenomena in discourses; to some degree they encapsulate discourse complexity because of the semantically highly charged positions they occupy in discourses, as points of discourse semantic accumulation through (frequency of) usage. The aim of studying the semantics and constellations of DKWs in depth is therefore to unravel this complexity in a bottom-up way rather than following a top-down premise that may serve to reduce it.

The advantages of analysing DKWs are:

- their frequency of occurrence in a variety of texts and genres in the public realm ensures relevance and salience for a sound empirical study;
- by their phenomenologically distinct form – as opposed to e.g. the notion of a persuasive strategy – they lend themselves easily to analyses using corpus-assisted approaches; hence quantitative and qualitative examinations can be fully practiced in corpus-guided DKW analysis;
- their ubiquitousness – the existence of semantic nodes make it easier to treat DKWs in contrastive/comparative projects.

Therefore, we would like to promote four ideas:

1. Keywords have been noted and investigated, especially with the arrival of corpus-driven approaches. However, the conceptualisation of DKWs outlined above transcends the premise of a corpus-driven treatment of the phenomenon in terms of frequency, significance and co-occurrence, since it

Following Tognini-Bonelli (2001), we differentiate between individual corpus methods such as corpus-driven and corpus-based, thus focusing on exploiting a corpus in specific ways.
also includes looking at DKW ensembles and at the positions occupied by individual DKWs in their related discourses, and the perspectives and evaluations they entail. This requires some general observation of, and insight into the related discourses – requirements which cannot always be fulfilled using data-driven procedures alone.

2. Corpus-driven approaches should be integrated into the study of DKWs. Lexical approaches to discourse analysis and a pragmatic understanding of lexical semantics are ideal points of convergence between corpus and discourse analysis.

3. With Busse (2008) and Ziem (2008a, b), we argue that DKWs are not just linguistic phenomena with discourse-determined semantics; they correspond to epistemically relevant phenomena that reflect frames of socially shared knowledge generated in and through discourses (see Section 1.3 below).

4. The results of such analyses may be formatted lexicographically (see Section 3.3 below for a suggestion of what this could include). This format could also be a good way of disseminating the kind of insights driven by pragmatics and corpus-based discourse analysis. The good relationship between analytical processability of DKWs and the scope of insight into the related discourse they can provide, renders them ideal for empirically-based contrastive discourse analyses (see Section 3.1 below and cf. e.g. Wierzbicka 1997; Wierzbicka & Harkin 2001; Bassi 2010; Schröter 2013).5

1.3 Cognitive dimensions of Discourse Keywords

We argue that the semantic structures of DKWs at the same time influence, (re-establish) and reflect ‘what we know’ about the issue in question; they are not only indicators of, but also factors in historical development. DKWs trigger the discursive background in which they are positioned and therefore function as a shorthand. Their use activates ‘knowledge’ that has been established in the progress of the related discourse (cf. Busse 2008), and therefore ensembles of DKWs reflect discursive formations.

All segments of knowledge that are triggered through the use of an expression by members of the speech or discourse community make up its meaning. These segments of knowledge are not essentially bound up in the expression as such nor have they been lexicographically defined (…) but they are constantly constituted through usage in discourses – sometimes affirmed and sometimes changed. The

5. See also the contrastive project on the lexicalisation and semantics of emotions across languages, Emolex. http://w3.u-grenoble3.fr/lidilem/projets/EMOLEX//spip.php?article17&lang=fr
methodology of frame semantics gives us an idea how these segments of knowledge can be found in a text corpus.

(Wengeler 2010: 140; translated from the German by Melani Schröter)

In and with discourses, a body of socially shared knowledge about the social, political and cultural reality is established. Analysing DKWs as nodes in the semantic networks of discourses allows the unravelling of structures of discursively established knowledge, as they are reflected and condensed in the semantic structures of DKWs. The frame concept, as it has been utilised in semantic theorising (cf. Fillmore 1976), cognitive linguistics (cf. Ziem 2008a; Hart 2010: 108ff.; both with a discourse analytical view), lexicography, and in relation to discourse-determined semantics and epistemology (Busse 2008; Ziem 2008a and b), allows for thinking about structures of knowledge in terms of semantic structures/valency. Ziem describes frames as “tools for the analysis of knowledge that guides understanding” (2008a: 367) and characterises them as

typicalised and structured segments of collective knowledge which result inductively or abductively from the intersection of similar individual experiences. (…) An activated frame regulates language and activity in that it triggers expectations regarding information that fits into it, or more precisely, regarding potential elements of knowledge in the available slots.

(Ziem 2008b: 97f.; translation from the German by Melani Schröter)

Our assumption is that the use of DKWs triggers discourse-determined knowledge, organised in frames.7 Similar to the way in which frames function as categories containing segments of collective knowledge, the semantics of DKWs contain discursive formations which have developed from the uses of the DKWs in their respective discourses; these contexts of usage can be studied through the analysis of collocations and corresponding concordances. In the following, we will report on the findings of a corpus-driven analysis of the DKW financial crisis, conceptualising its semantic structure to reflect a cognitive frame of collectively shared knowledge about the issue in question. This will direct attention to absences, i.e. slots that remain empty in a template matrix frame, capturing what could be (but is not) said about the financial crisis (see Sections 2.2–3.1 below), since what is phenomenologically absent can still be epistemologically relevant. Identifying patterns with a view on absence has so far not been the focus of corpus-analytical studies. Section 3.2 will outline the uses of such a conceptualisation for contrastive keyword analysis.


7. There are some similarities to the concept of ‘local textual functions’ as suggested by Mahlberg (2007).
2. The financial crisis in British broadsheet newspapers in 2009

2.1 Corpus and methodology

We used the database Lexis Nexis\(^8\) to extract articles which refer to the financial crisis from eight British broadsheet newspapers representing different political alignments (The Guardian, The Independent, The Times, The Daily Telegraph and their respective Sunday editions (see Figure 1). The search word financial crisis was used to extract articles covering the issues concerning the global financial/economic/banking crisis as reported by the British press, resulting in a corpus of 4,011,669 word tokens. Thus, a thematically homogenous 'purpose-built' corpus was constructed, within which the DKW financial crisis itself became the object of investigation. The search yielded 5,366 occurrences of financial crisis in articles in the newspaper sources listed above in 2009, a sufficiently large number suitable for different statistical analyses. The material was loaded into the corpus tool AntConc3.3.1 (Anthony 2011),\(^9\) a search and concordance analysis software which allowed us to perform an analysis of collocations using the statistical measure of log-likelihood.

With regard to possible concurrent labelling, we deemed crisis too general, as it continues to be used in all sorts of contexts.\(^10\) Figure 1 illustrates the results of a search conducted within LexisNexis – i.e. not limited to the purpose-built corpus – in order to ascertain that financial crisis was the most salient lexicalisation capturing the related discourse. It shows that financial crisis occurred more often than did economic crisis in every selected newspaper; banking crisis and debt crisis occurred even less frequently.

Another search, performed within the selected newspapers in LexisNexis before finalising the purpose-built corpus, confirmed that 2009 was the year in which financial crisis occurred most frequently as the central DKW (Figure 2).

This high frequency of occurrence in a newspaper corpus supports the claim that we are dealing with a salient lexical item and a relevant issue that is debated and referred to in public discourse. In line with corpus linguistic methodology and in order to find out more about the discourse-determined semantics of the

---

8. LexisNexis of Lexis®Library; see also http://www.lexislegalintelligence.co.uk/intelligence/lexislibrary. It is part of the library resources at the University of Reading, to which we were kindly granted access to aid our collaborative research. LexisNexis is an electronic resource of legal texts and covers a broad range of national as well as regional and local British newspapers.


Figure 1. Frequency of financial crisis and concurrent DKWs in 2009

Figure 2. Frequency of occurrence of financial crisis in selected British newspapers 2007–2010
DKW, we looked at co-occurrences which we understand to be indicative of semantic and thematic formations within the related discourse.

(...) (T)he corpus-based approach enables the researcher to arrive at a more complete understanding of the meanings and functions of certain word choices in texts (...). (...) by looking at the collocational strength of lexical items (...) we are given an objective sense of the themes and associations that are embedded in words due to their continual pairing with other words. By ‘exposing’ the hidden collocations of certain words, we can explain that a certain word or phrase contains a hint of bias.

(Baker & McEnery 2005:223)

However, what do we do with the lists that we obtain of the most (and likewise the least!) significant co-occurrences? How do we group the co-occurring items, categorise them and interpret their (co-)occurrence? One problem that we encountered in employing corpus approaches is the tendency to focus on that which is given, especially in large numbers, but to disregard what is underrepresented, as if, speaking phenomenologically, absence automatically implied epistemological irrelevance.

[A] critical analysis would not only be interested in accounting for what linguistic elements and processes exist in a text or set of texts, but would also need to explain why and under what circumstances and consequences the producers of the text have made specific linguistic choices among several other options that a given language may provide. That is, a critical analysis takes into account absences as well as presences in the data.

(Baker, Gabrielatos, Khosravinik, Krzyzanowski, McEnery & Wodak 2008:275)

We explored the analytical potential of a so-called matrix frame for the superordinate/hyperonym ‘event’ (cf. Konerding 1993) under which financial crisis could be subsumed. We used this matrix to identify presence and absence in the data. The matrix frame contains semantic roles/slots and provides logically possible statements/predications about an event, such as its phases, its causes, the people involved in it, factors that support it or help to repeat or end it, consequences of the event for people, the meaning of the event for people, characteristics of the event and similar events.11 Consequently, we treated the possible statements indicated above in a frame semantic context as slots, and the collocations that we retrieved as concrete lexical fillers of the notions captured in the slots, as suggested

11. These possible predications resemble the ‘ways-of-seeing’ suggested by Cruse – interestingly with reference to cognitive grammar (cf. Langacker 2008), which suggests that there is a shared interest here from the perspective of lexical semantics, cognitive linguistics as well as discourse analysis.
by Ziem (2008a). This matrix frame was not developed by Konerding in order to aid corpus-driven collocation analyses in a CDA framework, but to provide a repository-cum-analysis instrument for the description of meaning of nouns from a frame-semantic point of view. As the frame lacked a data-driven input when it was developed, some of the frame’s slots (e.g. slots 13, 14 and 15 in Table 1) seem conflated and difficult to separate logically, while some others (e.g. 4a and 4b) seem not as closely related as suggested by the model. Their generality renders some fillers from the empirical data ambiguous in their assignment to the given slots.

We used the matrix as a logical template which allows us to categorise concordances reflecting mandatory semantic predications and high frequency of occurrences of DKWs, not only on the basis of their assumed salience, but as occurring within a matrix of logically possible statements about an event. Doing this demonstrates which of the possible aspects of an event are over-emphasised and which are under-emphasised in our corpus. It is an attempt not only to systematise the analysis of the retrieved concordance list, but also
to determine the nature of the excluded – of what is not said – where there is literally an infinite range of concerns that are not taken up on any given occasion of talk. After all, what is not said can range anywhere from the sublime to the ludicrous. (McKenzie 2005: 452)

Since the use of the matrix here is explorative, we adopted the matrix frame as it appears in Konerding (1993) and found ourselves able to simultaneously identify its weaknesses. This reflects a general problem with the idea of using a generic frame, namely that of predefined vs. data-driven frames, derived from empirical studies. The slots are general and abstract, making it somewhat difficult to match them unambiguously with empirical data. However, abstract and general are exactly the qualities needed for the purposes mentioned above, since data-driven frames (i.e. defining the slots ad hoc on the basis of classifying co-occurrences from the data) would be prone to circular logic and would not offer a matrix for e.g. checking absence against logically possible presence (as ‘data-driven’ implies the presence of data) or for systematic comparison (as ‘data-driven’ would mean classifying on the basis of (different) results). The unrefined use of

---

12. The matrix frame was developed by Konerding in a process of abstracting towards a set of the smallest possible number of the most general hyperonyms in order to create classes of entities by abstraction, not to accommodate the results of empirical analyses (since at the time, no comprehensive corpus was available).

13. Its inadequacies with respect to certain semantic slots first emerged after a similar analysis of the German *Wirtschaftskrise* (‘economic crisis’) was performed (cf. Storjohann & Schröter 2011).
Table 1. Collocations of *financial crisis* as fillers of slots in the generic matrix frame for ‘event’

<table>
<thead>
<tr>
<th>Semantically salient fillers</th>
<th>Semantically less salient fillers</th>
<th>Semantically non-salient fillers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. conditions of origin or under which the financial crisis takes place</td>
<td>system, credit, cash, stock, shares, bonuses, Lehman, buy, boom, blame</td>
<td>Goldman-Sachs, America, sub-prime mortgages, US sub-prime mortgage crisis, profits, problems, funds, hedge funds, capitalism, billion, million</td>
</tr>
<tr>
<td>2. superordinate context in which the financial crisis plays a role</td>
<td>era, capital</td>
<td>capitalism</td>
</tr>
<tr>
<td>3. function/roles fulfilled by the financial crisis in this superordinate context</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4a. important phases/partial results</td>
<td>wake, start, onset, severity, in the face of, height, at the time of</td>
<td>wave</td>
</tr>
<tr>
<td>4b. characteristics of the financial crisis</td>
<td>global, economic, banking, unprecedented, worst, biggest, greatest, recent, international, worldwide, serious</td>
<td>current, depth, the most serious, severe, unfolding, largest</td>
</tr>
<tr>
<td>5. important players within the financial crisis</td>
<td>Gordon Brown, Barack Obama, America</td>
<td></td>
</tr>
<tr>
<td>6. important contextual items which characterise players within the financial crisis</td>
<td>summit, meeting</td>
<td></td>
</tr>
<tr>
<td>7. conditions under which the financial crisis changes characteristically</td>
<td>responsibility</td>
<td></td>
</tr>
<tr>
<td>8. typical duration of the financial crisis</td>
<td>extent, scale</td>
<td>amid, during, unfold</td>
</tr>
<tr>
<td>9. conditions under which the financial crisis typically begins</td>
<td>system, credit, bonuses, Lehman, begin</td>
<td>hedge funds, greed, responsibility</td>
</tr>
<tr>
<td>10. conditions supportive of the financial crisis</td>
<td>system, profits, banks, power, business</td>
<td>capital, capitalising, contributing</td>
</tr>
<tr>
<td>11. conditions under which the financial crisis repeats itself</td>
<td>repeat</td>
<td></td>
</tr>
<tr>
<td>12. conditions under which the financial crisis stops or ends</td>
<td>bail-out</td>
<td>way out, end</td>
</tr>
</tbody>
</table>
Table 1. (continued)

<table>
<thead>
<tr>
<th>Semantically salient fillers</th>
<th>Semantically less salient fillers</th>
<th>Semantically non-salient fillers</th>
</tr>
</thead>
<tbody>
<tr>
<td>13. different states of affairs or further events which the financial crisis may lead to</td>
<td>unscathed, unable</td>
<td>eurozone crisis, chaos</td>
</tr>
<tr>
<td>See also slot 14 below</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. different consequences that the occurrence of the financial crisis may have for people</td>
<td>storm, problems, pressure, losses, disaster, aftermath, consequences, emerge, revenue, decline, cuts, failure</td>
<td>bill, risk, changes, pensions shortfall</td>
</tr>
<tr>
<td>result, recession, impact, debt, collapse, prices, effects, support, failure, damage, costs, downturn, result, fall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. meaning of the occurrence of the financial crisis for people</td>
<td>property loss, job loss, raise, opportunity, increase, income, hurt, hard, fragile, handling, understand, cope, learn, lessons, victim, risks, pressure, casualty, struggle</td>
<td>enduring, battered, come through, dealing with, survive, threat, reputation, instability, inflation</td>
</tr>
<tr>
<td>problem, debt, fear, lack, unemployment, response</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. similar events, differences to such events and general categories into which the financial crisis falls</td>
<td>recession</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. characterisation of theories in which the financial crisis plays a role</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. characterisation of information which is initiated by the occurrence of the financial crisis (such as naming further events as consequences)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. other means of reference to the financial crisis</td>
<td>economic crisis, global crisis</td>
<td>global economic crisis, cash crisis, the world economic crisis, banking crisis, debt crisis, depression</td>
</tr>
<tr>
<td>economic crisis, global crisis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Konerding’s frame as it is, with added categories from previous empirical research, is a compromise which illustrates this difficulty and ambiguity. Larger empirical investigations of predicate types might be able to refine existing categories as established by Konerding. The emphasis on the most generic conceptualisation of ‘event’ also overlooks the possible influence of the more specific meaning of ‘crisis’ as “a time of intense difficulty or danger” or “a time when difficult decisions have to be made” (Oxford English Dictionary 2010).
2.2 Findings

The range of possible statements about events (the financial crisis) were conceptualised as slots (arranged in rows in Table 1); the lexical co-occurrences with financial crisis were treated as contextual fillers that could be assigned to these slots. The analysis of collocations (with a word span of five) performed by AntConc resulted in 1953 collocates, which we primarily arranged according to frequency, while excluding function words. As well as wanting to find out which slots were contextually filled by which lexical filler, we also wanted to judge the degree of regularity and salience of a pattern. Therefore, we split the table into categories in terms of frequency of co-occurrence with regard to a second parameter indicating semantic salience. Table 1 is split into columns reflecting three categories:

- semantically salient fillers (for our purposes, more than 10 occurrences, e.g. worst (84), result (43), bank (29), impact (26), people (18));
- semantically less salient fillers (for our purposes, fewer than 10 occurrences, e.g. action (8), government (8), risk (6));
- semantically non-salient fillers, (for our purposes, fewer than five occurrences, e.g. pension (3), inflation (2), support (2)).

The matrix frame is not intended as a way of making absolute claims about the semantic structures but rather as an indicator of semantic distribution. The slots support observations regarding absence and presence, as well as degrees of presence. Some slots are filled more than are others, which points to the existence of thematic foci. The frame is also useful when accounting for statistically less significant co-occurrences contribute to ‘constituting the meaning’ of a DKW by mapping its discourse semantic hinterland, irrespective of an in some way arbitrary cut-off point defined by statistical significance. We do not fully endorse the implication that what is beyond this cut-off point would be phenomenologically irrelevant, especially since other factors – e.g. proximity within the space of the five words to the left and to the right – also need to be considered. When looking at the distribution over slots, the frame illustrates patterns of usage. Even where individual fillers are not statistically significant, the presence of other fillers in the same slot points to the salience of the specific semantic aspect represented by that very slot. We have provided translations of the slots in the original frame proposed by Konerding.

14. Most of these collocates (ca. 1300) co-occurred with the search word only once.

15. We are aware of the subjective aspect of a division according to frequency. We are not interested in absolute numbers; rather, our aim was to obtain an impression of the degree of dominance of lexical realisations for specific semantic categories.
(1993) for the hyperonym ‘event’. Since the original matrix frame is generic, it refers to ‘event’ where we specified ‘financial crisis’ for our current purposes.

Some of the co-occurrences are ambiguous in their relation to the given slots. For example, responsibility (a lack of it) occurs in the context of the beginning of the crisis (slot 9), but also as something that needs to be resumed in order to deal with the crisis (slot 7). In such cases, the fillers were grouped into slots relevant to the different ways in which they were used. Due to the generality of the slots, it was difficult to determine whether the co-occurrences amid, during and unfold referred to the duration (slot 8, which is where we list them) or to the phases/partial results of the crisis (4a). In our previous, similar corpus study of German

---

**Table 2. Collocations of financial crisis as fillers of slots added by the authors to the generic matrix frame ‘event’**

<table>
<thead>
<tr>
<th>Semantically salient fillers</th>
<th>Semantically less salient fillers</th>
<th>Semantically non-salient fillers</th>
</tr>
</thead>
<tbody>
<tr>
<td>20. who/what is affected by the financial crisis</td>
<td>banks, world, people, US, (chief) executives, economy, business, UK, public, Europe, market, banking, minister, money, London, Britain, country, bankers, seniors, nation, life, industry, investors, housing market, companies, (financial) services authority/industry, stock market, director</td>
<td>world, trade, taxpayers, sector, system, overseas, management, government, family, currency, Treasury, firms, shareholders, savings, society, economists</td>
</tr>
<tr>
<td>21. financial crisis as agent</td>
<td>cost, cause, effect</td>
<td>hit (hard), strike</td>
</tr>
<tr>
<td>22. conditions under which the financial crisis will develop further</td>
<td>system, profit, credit</td>
<td></td>
</tr>
<tr>
<td>23. events/states of affairs/measures that people take with regard to the financial crisis</td>
<td>plan, need, pay, warned, investment, forced, recovery, policy, growth, regulation, bail-out</td>
<td>action, working, toll, response, rescue, reporting, demand, challenge, boom, behaving, tackle, solution, answer, recovery, prevent, confidence, plans, hope, gold, efforts, attempts, trust, survive, success, reforms</td>
</tr>
<tr>
<td>24. hyperonyms – what the financial crisis is seen to be/is compared with</td>
<td></td>
<td>challenge, problem</td>
</tr>
</tbody>
</table>
Among the striking results of our analysis is the emphasis on the overwhelmingly negative consequences of the financial crisis: recession, impact, debt, collapse, failure, damage, downturn, fall, problems. Slots 14, 15 and 20 show a large number of fillers, i.e. collocates of financial crisis. This is in contrast with the relatively small number of fillers pertaining to people and agency (5, 6, 7, 23), causes (1, 9) and influence or development (7, 8, 9, 10, 11, 12). A perception of the crisis emerges as a phenomenon that will affect many in a negative way, but which remains opaque in terms of why it occurred and what could be done about it. This impression is supported by the fillers in 4b, which characterise the financial crisis as wide in scope (global, worldwide, international) and scale (biggest, greatest, worse, unprecedented), and by occurrences of financial crisis as an agent that does something to someone (slot 21: cost, cause, effect).

The over-emphasis on those (negatively) affected by the crisis, and on its (negative) consequences, along with a neglect of cause, influence, development and agency can be seen as characteristic of newspaper discourse, which not only constructs narratives about what is going on in the world, but also points out what that may mean for its readers – i.e. it brings such events closer to them in more than one sense. Focusing on consequences may also be a feature of the human way of perceiving situations and events, which – as already pointed out by Dewey (1929, esp. 100ff.) – anticipates course and consequence. Our finding is of interest, since the media-sustained public sphere is also the realm in which discourse communities’ understanding of themselves is formulated and perpetuated; it would make equal sense to reflect on reasons and responsibilities and on scope for action. The structural semantic bias towards negative consequences could be seen as reflecting, but also as inducing, angst (angst-inducing discourse can prepare people for political ‘necessities’ as such). The fact that specific collocates denoting negative consequences occur more frequently than those denoting cause and reason might also be due to prior contextual priming (Hoey 2007): speakers might be more strongly primed to associate events such as the financial crisis, and hence its lexical representation, with personal negative consequences, while other thematic roles might be contextually less strongly primed and hence contain fewer or no fillers. The corpus texts can therefore provide us with linguistic evidence for shared conceptual associations.
of a discourse community and demonstrate how these are lexically constructed through collocates. The regularities and patterning of lexis and grammar reveal the degree of conventionalisation of such primed associations.\textsuperscript{16}

3.2 Implications for the contrastive analysis of DKWs

An earlier, similar study (Storjohann & Schröter 2011) used Konerding’s matrix frame, applying it to the German term \textit{Wirtschaftskrise} (‘economic crisis’), albeit on the basis of a larger German newspaper corpus including regional papers and using a different corpus analysis tool.\textsuperscript{17} In German, it was a slightly different lexicalisation (\textit{economic crisis} vs. \textit{financial crisis}) that dominated the discourse in 2009 as a lexical node. The following account of similarities and differences illustrates the potential of systematic contrastive DKW analysis on the basis of such a matrix frame.

A striking similarity between German and British crisis discourse lies in the emphasis on the overwhelmingly negative consequences. As is the case for the British corpus, the German corpus, too (in slots 14, 15 and 20), shows a large number of ‘negative consequence’ fillers; e.g. \textit{Auswirkungen} (\textit{effects/repercussions}), \textit{Folgen} (\textit{consequences}), \textit{Arbeitslosigkeit} (\textit{unemployment}), \textit{sinkende Nachfrage} (\textit{decreasing demand}), \textit{wegbrechende Steuereinnahmen} (\textit{eroding tax revenues}), \textit{Entlassungen} (\textit{redundancies}).\textsuperscript{18} In both discourses, the crisis is characterised by attributes such as \textit{global}, \textit{international}, \textit{worldwide} (German slot 4b), and with superlatives such as the \textit{most serious}, \textit{greatest}, \textit{largest}, \textit{biggest}. In both corpora, references are made to phases and developments (slots 4a and 8). There are also similarities in measures to fight the crisis (slot 23).

Reasons and responsibility for the crisis and the sociopolitical and economic circumstances surrounding it (slots 1, 9, 10) are largely absent in the German

\textsuperscript{16} However, within a corpus-derived framework, Hoey’s priming theory is more suited to account for the presence of fillers/collocates rather than for their absence, unless it could be argued that the absence of collocates is contextually primed in the same way.

\textsuperscript{17} The corpus was compiled on the basis of the text archive of the Institute for German Language in Mannheim. For the analysis of the corpus, the query tool COSMAS II with its integrated collocation tool “\textit{Statistische Kollokationsanalyse und -clustering}” (Belica 1995) was used. This uses the same statistical measure of log-likelihood as AntConc.

\textsuperscript{18} Posch (2010) comes to a similar conclusion after investigating patterns of argumentation, nomination, predication and metaphorisation in 40 newspaper articles dealing with the economic crisis in three Austrian newspapers between October 2009 and October 2010. She observes that the crisis is presented as an external threat, and that there is notable depersonalisation regarding causes and agency, so that the crisis is presented as a natural force or catastrophe.
data. In the English discourse, a comparatively large number of fillers refer to the circumstances under which the crisis began, developed and is possibly being sustained. Fillers in slot 2 suggest that in Germany, the crisis was positioned as a phenomenon that is part of *globalisation* whereas in the British discourse, the financial crisis is part of an era and a phenomenon of *capitalism*. This also pertains to differences (who and what is affected), particularly in slot 20. In both discourses, the impact on banks and industrial branches is mentioned. In German, collocations such as *Arbeitslosenzahlen* (unemployment figures), *Verbraucher* (consumers), *Kommunen* (local councils), *Arbeitsplätze* (jobs), *Bundesbürger* (citizens), *Arbeitnehmer* (employees) and *Länder* (federal states) are significant. Similar fillers are found in the British discourse, but they are far less dominant. In the British press, there is a slight tendency to emphasise the effects of the crisis as primarily affecting *business*, *bankers*, *business people*, *executives*, *companies*, *the stock market*, *trade*, *currency*, *shares*, *brokers* and *analysts*. Reference to the banking sector is more limited in the German data, but fillers with reference to other central branches of industry such as the car industry (*Autohersteller* (car manufacturers), *Autogesellschaften* (automotive companies)) occupy positions similar to those found for the English banking terms.

The fillers in slot 16 referring to similar events or problems seem to also pertain to this difference. In German, the item *Wirtschaftskrise* co-occurs with terms such as *Globalisierung* (globalisation), *Klimawandel* (climate change), *Klimaschutz* (climate protection), *Klimakrise* (climate crisis), *Umweltprobleme* (environmental problems), *Umweltschutz* (environmental protection) and *Terrorismus* (terrorism), all of which can be described as global problems/challenges. They typically occur in coordinated patterns such as *Wirtschaftskrise und Globalisierung* or *Wirtschaftskrise und Umweltschutz*. In Germany, the economic crisis is conceptualised as a global phenomenon, together with other global problems. In the British corpus, the *financial crisis* is not really treated as a problem on a par with other global issues. Looking at coordinated patterns in the British newspapers reveals that the most dominant pattern is *financial crisis* and *recession/economic downturn/slump/slowdown/collapse*. In this light, it seems anything but accidental that the equivalent German discourse finds its lexical centre in *Wirtschaftskrise* (economic crisis) instead of *Finanzkrise* (financial crisis), and that it is the other way round for the UK.

This semantic evidence is in line with the financial services playing a bigger role in the British economy, where there is a whole culture and economy built on property and an associated mortgage market. In Germany, real estate plays a more limited role and mortgage banking is less of a risk-associated business. German banks were affected to a lesser degree and the implications for Germany seem to
be a more generally disadvantageous effect of the economic decline on Germany’s export-orientated economy. In this context, the occurrence of globalisation in the German discourse is interesting – like in the case of globalisation (cf. Storjohann & Schröter 2011; Storjohann 2007), the economic crisis appears as an external agent and not a phenomenon that is rooted in the way the German economy works. It affects Germany as a development that will unleash negative consequences for Germany, even though it was caused elsewhere. The British discourse focuses more on the conditions in the financial sector that brought the financial crisis about (slots 1, 10) and on the political measures taken in reaction to these conditions as well as on the players reacting to them (slots 5, 12). These contrasts indicate that discourses develop differently according to different broader contexts.

The contrastive study of a single but discursively salient lexical item can help carve out differences that make sense in a more general discourse context; in this case, because of its relative remoteness, the crisis is seen as a global phenomenon in the German discourse, in contrast to the British discourse, where the financial sector plays an economically less crucial role than it does in the German context. However, it is not only differences which are of potential interest; the over-emphasis on (negative) consequences in both corpora may indicate either an opportunistic angst-inducing discourse and/or a generic feature of newspaper discourse (focusing on who/what is affected by ‘the news’).

3.3 Implications for DKW lexicography

A lexicographical project on DKWs would have to account for their inseparable relation to the discourse that shapes them. Thus, an entry would have to contain information about the DKW’s discourse (historical) context; in the case of financial crisis, about the discourse emerging after the so-called ‘collapse’ of Lehman Brothers in the US and Northern Rock in the UK in 2007/2008. It should address the role the DKW played in the discourse as the central lexicalisation of the emerging discourse, which receded in frequency and salience when the discourse shifted to the economic downturn/recession more generally and when other issues emerged such as the European debt crisis. Its usage does not indicate the highly controversial conceptualisations of the issues it signifies; metalinguistic reference such as ‘so-called’ or concurring lexicalisations that offer different perspectives (such as, say, ‘collapse of neoliberalism’ or ‘crisis of homeownership’) do not seem to influence its use. The fact that the financial crisis had come into existence following the collapses mentioned above, the identification of the dimensions it affected (the financial sector, the economy, the labour market, consumer strength
etc.), and the fact that it had consequences, all went undisputed. The semantic profile of financial crisis could be described on the basis of the above collocation analysis – overwhelmingly contextualised in relation to who/what is affected in a negative way, conceptualising it as extraordinary in scope (global) and intensity (worst), etc.

In a contrastive lexicographical project, a section similar to 3.1 above (on presence and absence) could be added. Ideally, entries would contain citations from the corpora illustrating typical usage or typical collocations in their local contexts. As the corpus-assisted description of discourse reveals a complexity that is difficult to cover sufficiently in a printed dictionary, it would be worthwhile for such a project to be published in electronic format and to organise entries as tabs that cover the different kinds of information, i.e. discourse (historical) context; the use and position in an ensemble of the DKW within its related discourse; the detailed semantic profile of the DKW; and, in a contrastive project, information about similarities and differences of lexicalisation and/or use and notions within discourses in different languages and discourse communities.

4. Conclusion

Using the example of financial crisis, the present study has employed corpus-assisted approaches to discourse analysis. We have argued in favour of conceptualising salient lexical items in discourses as DKWs to empirically underpin the pragmatic dimension of lexical semantics, while acknowledging the cognitive dimension of DKWs and paying attention to what could logically be said in related discourses, but is not said – with epistemic consequences. Our intention was also to promote the idea of a contrastive discourse analysis, maintaining that frame-based analyses of DKWs could be a way to handle such an approach practically and systematically. Last but not least, our aim was to develop the notion of discourse lexicography, as a means of perhaps engaging a wider public with the insights that can be gained from keywords in discourse contexts. However, in order to pursue the envisaged research on a larger scale, a number of problems would need to be addressed. First of all, while financial crisis is an obvious example of a DKW, other DKWs play a role as well. Conceptually, a line would need to be

19. For comparison, these aspects seem less clear cut in the case of the DKW globalisation (cf. Teubert & Čermáková 2004; Storjohann 2007).
Patterns of discourse semantics

...drawn to indicate which other words (bail-out, collapse, downturn, bankers, greed, etc.) can be considered ‘key’ in this discourse, and a decision would need to be made with regard to (relative) frequency as the main criterion for assuming discursive salience or keyness. Methodologically, on the one hand, it does not seem satisfactory to try to set up a predefined list of DKWs that can be searched in a targeted way in corpora. It would require careful consideration to ascertain how else DKWs can be retrieved from data or even help to construct purpose-built corpora, based on the assumption that certain lexical items will occur regularly. A similar problem, relating to the area of tension between predefined or data-driven matrix frames, was discussed above in Section 2.1. Clearly, there is a need for linguistic models to be developed which could help to map corpus findings more insightfully. At the end of the day, contrastive research requires careful consideration of comparability of sources and corpora.

References


Stubbs, Michael. 2010. “Three Concepts of Keywords.” In Keyness in Texts, ed. by Marina Bondi and Mike Scott, 21–42. Amsterdam: John Benjamins. DOI: 10.1075/scl.41.03stu


Ziem, Alexander. 2008b. “Frame-Semantik und Diskursanalyse: Skizze einer kognitionswissens-
chaftlich inspirierten Methode zur Analyse gesellschaftlichen Wissens.” In Methoden der
Diskurslinguistik – sprachwissenschaftliche Zugänge zur transtextuellen Ebene, ed. by Ingo
DOI: 10.1515/9783110209372.2.89
AntConc.” In Die da oben: Sprache, Wahlen, Partizipation, ed. by Frank Liedtke, 69–90.
Bremen: Hempen.

About the authors

Melani Schröter is Associate Professor in German Linguistics at the University of Reading. Her
research focuses on public and political discourse, including addressee orientation and silence
and concealment in political discourse and public metadiscourse about controversial commu-
nicative behaviour. More recent research interests focus on contrastive analyses of political dis-
course, esp. discourse key words in European migration discourses as well as defiant discourses
in subcultures, esp. punk. Her most recent book publication deals with absence and silence in
discourse (Silence and Concealment in Political Discourse, John Benjamins 2013).

Petra Storjohann works in the lexical division of the Institute for German Language in Mann-
heim, where she is part of the lexicographical project elexiko, which documents the core lexis
of public discourse on the basis of a large corpus. Her research focuses on the empirical, corpus
analytical study of lexical-semantic relations in texts and discourse. Her recent publications
include an edited volume Lexical-semantic relations. Theoretical and practical perspectives (John
Benjamins 2010), articles about the phenomenon of colligation, about paradigmatic construc-
tions, corpus-methodological aspects as well as reports from the lexicographic practice.

Addresses for correspondence

Dr Melani Schröter, Lecturer in German
University of Reading
Modern Languages and European Studies
Whiteknights PO Box 218
Reading RG6 6AA, United Kingdom
m.schroeter@reading.ac.uk

Dr Petra Storjohann
Institut für Deutsche Sprache Mannheim
R5, 6-13
68161 Mannheim, Germany
storjohann@ids-mannheim.de