

## Weaving web data into a diachronic corpus patchwork

Andrew Kehoe and Matt Gee

Research & Development Unit for English Studies, Birmingham City University

### Abstract

*This paper offers a reassessment of the role of web data in diachronic linguistic analysis. We introduce the diachronic search facilities provided by the WebCorp Linguist's Search Engine, including the use of a new 'heat map' graph for the analysis of changes in collocational patterns over time. We illustrate how web data can be used to supplement data from standard corpora in lexicological studies. Our focus is on the vogue phrase credit crunch and the paper compares examples from standard corpora (BNC, Brown, LOB, Frown, FLOB) with those found in web-accessible newspaper texts. Contrary to previous studies, we do not rely on the web solely for the most up-to-date usage examples. Instead, we show how web-accessible texts dating back to the beginning of the 20<sup>th</sup> Century can be used to fill gaps in and sharpen the picture provided by standard corpora.*

### 1. Introduction

The original WebCorp project (Kehoe & Renouf 2002) was an experiment to see whether we could develop a system to extract linguistic data from web text efficiently and present this to the linguist in as usable a fashion as it is presented in traditional corpora. The system (<http://www.webcorp.org.uk>) receives a word or phrase and other requirements from the user, passes these to a commercial search engine (Google, AltaVista, etc), and extracts the 'hit' pages from the search engine results. Each page is accessed and processed and the extracted concordances are presented to the user in a choice of formats. The WebCorp tool established that web text, though problematic, is nevertheless a resource that can complement corpus evidence with examples of usage that is rare, re-emergent, new or productive.

The WebCorp Linguist's Search Engine (WebCorpLSE) is designed to bypass the commercial search engines upon which WebCorp relied as gatekeepers to the web.<sup>1</sup> WebCorpLSE is crawling and processing the web to build a 10 billion word (or 7 terabyte) text corpus, including a multi-terabyte 'mini-web', designed to act as a microcosm of the web itself (Kehoe & Gee 2007). In addition to the mini-web, WebCorpLSE has built a newspaper sub-corpus, containing daily issues of UK broadsheets from 1984-present and recent issues of other UK and international newspapers. We have also worked with our university colleagues to build collections to assist in their research and teaching, including sub-corpora of blogs, science fiction and major English literary works. All collections are searchable via linguistically-tailored front-ends.

**Final version to appear in A. Renouf & A. Kehoe (eds.) *Corpus Linguistics: Refinements & Reassessments*, Amsterdam: Rodopi (2009), pp. 255-279.**

It is now generally accepted that web data are of value in supplementing evidence from traditional, or 'standard', corpora when examining linguistic change over time. Previous work has tended to turn to the web as a source of evidence of the very latest trends in language use and of new coinages not found in standard corpora. Mair, for example, in a study of change and variation in present day English, states that the best way to 'minimise the risk' of relying on the web as a corpus is to

use it not as a stand-alone source of data, but in conjunction with tried and tested closed corpora. In diachronic work, such corpora are positively indispensable because they add the necessary element of time depth to the web. (Mair 2007: 236)

The approach described by Mair is, in part, necessitated by the bias in commercial search engines, like Google, toward the most recently updated pages and the difficulty in extracting older data from the web through these search engines (cf. Kehoe 2006). In this paper, we shall describe the corpus search tools available in WebCorpLSE and the new possibilities which these open up for diachronic linguistic study. We shall illustrate that carefully selected web data can, in fact, provide the necessary 'time depth' by overlapping with and filling gaps in the data provided by standard corpora. The web data can, thus, sharpen the diachronic picture presented by standard corpora rather than simply widening it at the most recent end of the timeline.

Our analysis will focus on a phrase which, given media preoccupations at the time of writing, may initially seem to be a perfect example of the kind of vogue construction for which linguists have, thus far, turned to the web for evidence: *credit crunch*. The phrase does not appear in the Oxford English Dictionary (OED) and was named as the Oxford University Press 'Word' of the Year for 2008<sup>2</sup>, an honour frequently bestowed on new coinages. One may therefore assume from this that *credit crunch* will not be found in standard corpora but, as we will show in the next section, there are examples of the phrase in corpora. In fact, the corpus-based *Cambridge Advanced Learner's Dictionary* (CALD) includes *credit crunch* in its entry for another phrase, *credit squeeze*:

**credit squeeze** UK noun [C] (US **credit crunch**) INFORMAL  
a period of economic difficulty when it is difficult to borrow money from banks (<http://dictionary.cambridge.org/define.asp?key=18170&dict=CALD>)

This dictionary (latest printed version 2008) is based on the Cambridge International Corpus: just over 240 million words of UK and US writing and speech, with an emphasis on business, legal, academic and financial English. It is likely that the supposed distinction between US *credit crunch* and UK *credit squeeze* was drawn from the last of these, a 'collection of books, journals, newspaper articles relating to economics and finance'<sup>3</sup>. This distinction between UK and US usage is one which will be investigated in this paper. We shall use

standard corpora of British and American English and data from the web to examine usage patterns over time and determine to what extent the US/UK distinction in the *CALD* entry holds true. Initially, our analysis will focus on *credit crunch* and we will return to *credit squeeze* in section 4.

## 2. Evidence from standard corpora

The phrase *credit crunch* appears seventeen times in the British National Corpus (BNC) in texts from 1991-3, shown in Figure 1.

● <b><i>The Economist</i>, 1991</b>	
1	The Federal Reserve is struggling to allay fears of a “ <b>credit crunch</b> ” – when banks are reluctant to lend except to the most creditworthy borrowers. (ABD 81)
2	Fears of a “ <b>credit crunch</b> ” have prompted policy changes at the Federal Reserve in recent months. (ABD 2335)
3	A <b>credit crunch</b> is the name economists give to a sudden reluctance among banks to lend money. (ABD 2339)
4	Typically, a <b>credit crunch</b> happens when banks start to worry about the creditworthiness of their borrowers. (ABD 2341)
5	A <b>credit crunch</b> – mild, as yet – is undoubtedly under way in America. (ABD 2347)
6	There is a risk, though, that the supply of credit will start to fall faster than the demand; in other words, that a <b>credit crunch</b> will start to drive the process of credit contraction. (ABD 2361)
7	The Bank of England, responding to fears of a <b>credit crunch</b> , has asked banks to think twice before turning away would-be corporate borrowers. (ABD 2367)
8	Such frightening costs undermine the credibility of the FDIC, because, if a banking crisis were to start, the government might find itself facing a <b>credit crunch</b> of its own. (ABD 2381)
9	This demand on the international capital markets raises interest rates, aggravating the problems of debt and <b>credit crunch</b> . (ABD 2386)
10	The Federal Reserve’s Alan Greenspan said the Fed would do what it could to ease America’s <b>credit crunch</b> . (ABG 3211)
11	Yet they, too, complain of aches and pains, of being squeezed by a “ <b>credit crunch</b> ” under which borrowing has become harder even while interest rates have been falling. (ABJ 3178)
12	There is no generalised <b>credit crunch</b> in Japan, but particular firms are being hurt. (ABJ 3982)
13	That suggests that a <b>credit crunch</b> is taking place, especially since banks are still under orders from the central bank not to increase lending to property companies beyond the overall rate of loan growth. (ABK 2395)
● <b><i>Daily Telegraph</i></b> , electronic edition of 15/04/1992 (AKJ 453)	
14	That would cause a severe <b>credit crunch</b> .
● <b><i>Unigram X</i>, 1993</b> (CTG 399)	
15	Debt-laden Tustin, California-based business systems supplier MAI Systems Corp appears to have hit a <b>credit crunch</b> according to the German weekly Computerwoche.

<ul style="list-style-type: none"> <li>● <i>Keesings Contemporary Archives</i>, Longman, 1991 (HLC 632)</li> </ul>
16 It also took measures to ease the so-called “ <b>credit crunch</b> “, mainly by relaxing regulatory pressures in order to encourage bank lending.
<ul style="list-style-type: none"> <li>● <i>The Scotsman</i>: Business section, <b>unknown date</b> (K59 3187)</li> </ul>
17 Writing in the February issue of the Lloyds Bank Economic Bulletin, he says: “The restoration of financial balance will mean that, far from there being a <b>credit crunch</b> , banks are likely to continue to find very little net demand for loans from companies.”

Figure 1: All examples of *credit crunch* from the BNC<sup>4</sup>

It is clear from the BNC concordances that the phrase was current in British English in 1991, and also that a credit crunch was underway in the United States at that time and was in danger of occurring in the United Kingdom. However, the fact that the phrase occurs in double quotes, complete with a full gloss, in three articles from *The Economist* (whose readers may be expected to be more familiar with economic terms than readers of general audience newspapers) indicates that the phrase was still new and unfamiliar to the majority of UK readers.

The BNC data seem to indicate that the phrase *credit crunch*, like the economic situation it describes, first occurred in the US, thus confirming the *CALD* definition. This opens up the possibility of turning to another set of standard corpora: the Brown family, ‘corpora equivalently sampled from the language, though different in temporal as well as geographical provenance – as a means of identifying rather precisely how the use of the language developed over a period’ (Leech and Smith, this volume). The 1961 Brown and LOB corpora, with 1 million words each of written American English (AmE) and British English (BrE) respectively, contain *crunch* only in a literal sense. FLOB, the 1 million word BrE corpus from 1991, does not contain any instances of *crunch* (though it does include literal *crunching* and *crunchiness*). However, Frown, the 1 million word AmE corpus from 1992, includes three instances of *crunch*, all of which are used in a metaphorical sense to refer to financial situations, including one occurrence of *credit crunch* (Figure 2).

1.	Hallinan introduced the legislation following an Examiner story that revealed that some city bureaucrats were commuting in style at taxpayer expense despite a severe <b>budget crunch</b> that has required reduction of some vital health services. (A25 15-18)
2.	For all of Mr. Kornbluth’s cultural observations, the book is not yet written that closely tracks [US financier Michael] Milken’s persecution with the <b>credit crunch</b> and recession. (C12 89-91)
3.	They lend legitimacy to the racist and misogynist stereotypes so popular with conservative politicians and disgruntled taxpayers who feel an <b>economic crunch</b> and are looking for someone to blame. (G23 160-163)

Figure 2: All examples of *crunch* from Frown corpus<sup>5</sup>

The limitations of the Brown family for lexical rather than grammatical studies, as pointed out by Leech and Smith (this volume), are clear from these results. A

study of *crunch* based solely on Frown would have little choice but to conclude that the word is used only to refer to negative financial situations (the semantic prosody of *severe*, *disgruntled* and *blame* is clear).

It is worth noting, however, that the authors of these AmE texts from Frown, unlike the authors of BrE texts from the same period in the BNC, do not feel it necessary to provide a gloss for *credit crunch* and use *crunch* in a wider sense to refer to a variety of financial situations. Indeed, the example of *credit crunch* in this corpus is mentioned in passing in an article which focuses on a different topic; it is given rather than new information.

### 3. Turning to the web

Nesselhauf (2007) makes the distinction between two types of web-based diachronic linguistic analysis. The first is the approach taken by us with the original WebCorp system: the analysis of short term changes in texts produced specifically for the web (Kehoe 2006). The second is the analysis of changes in ‘larger and/or earlier time-spans based on texts written for other media and later made available on the internet’ (Nesselhauf 2007: 287). WebCorpLSE moves us toward this second approach to web-based diachronic analysis. As outlined in section 1, the system provides access, via the web, to a variety of sub-corpora, many of which were compiled from web-accessible text collections such as Project Gutenberg.

In this paper we focus on the WebCorpLSE newspaper sub-corpus. With regard to this text-type, Nesselhauf’s distinction between the two kinds of diachronic analysis becomes somewhat blurred in that modern newspaper articles are not produced ‘specifically for the web’ but nor are they made available on the web only at a later date. For the past decade, printed newspaper texts have been made available *simultaneously* on the web.

We shall return to this point in section 3.2, which provides details of our newspaper sub-corpus and the kinds of diachronic search possible using WebCorpLSE. Before looking at this, we outline in 3.1 the restricted (though useful) provision for diachronic linguistic analysis in the web-based Google newspaper archive. Throughout our analyses in section 3.1 and 3.2, we will attempt to confirm the accuracy of the *CALD* definition of *credit crunch* by examining:

- i) what web data can tell us about *credit crunch* in AmE, including first occurrence
- ii) what web data can tell us about the introduction of *credit crunch* into BrE

#### 3.1 Google News

Google News (<http://news.google.com>) is a ‘news aggregator’: a website that collates, from multiple sources, news stories which may be of interest to an individual user and presents these on a single page. In addition, the Google News

site contains an archive of major international newspapers and magazines dating back over 200 years. More specifically, Google News provides a master index to several existing newspaper archives (*New York Times*, *Washington Post*, etc) and has begun to digitise print newspapers which were not previously available in electronic form.<sup>6</sup> Google is working with publishers to make ‘millions of pages of news archives’ available, in facsimile and in a form searchable by keyword.

The Google News Archive is not a corpus in the sense used by linguists. Accurate word frequency information is not available and only very limited word contexts are provided, as we shall show in the examples below. However, Google News does allow us to pinpoint when a particular word or phrase entered the lexicon of newspapers in the English-speaking world.<sup>7</sup>

By default, the Google News Archive search interface<sup>8</sup> shows results in ‘relevance’ order, in a similar manner to a standard Google search. A secondary ‘timeline’ option allows the results to be viewed in date order, as shown in Figure 3 for the phrase *credit crunch*.

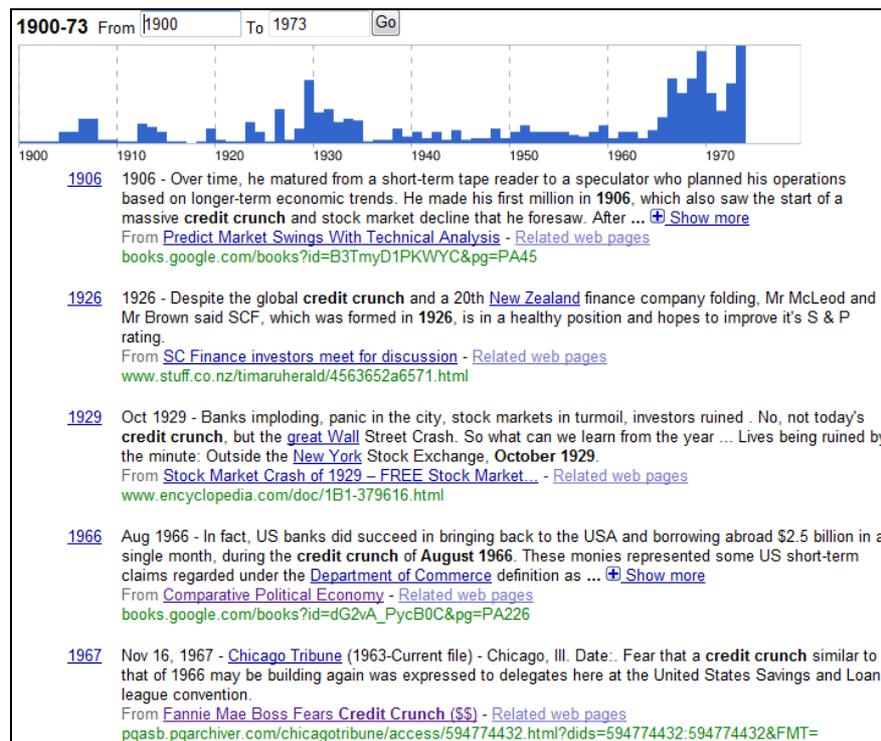


Figure 3: ‘Timeline’ results from Google News Archive for *credit crunch*

Figure 3 would initially seem to indicate that there are examples of *credit crunch* dating back to 1906. However, this output highlights a severe limitation of

Google News for linguistic search. Many of the years associated with articles in the results list are not the year the article was written but the year in which the event being discussed took place. For example, the first result in Figure 3 (listed as 1906) is actually from a book published in 2002, and the second result (1926) is from an article dated May 29<sup>th</sup> 2008 in the New Zealand newspaper *Timaru Herald*.<sup>9</sup> The fundamental difference between the dates required for informational search and for linguistic search (cf. Kehoe 2006) makes Google News an inadequate search interface for the latter. It is undoubtedly useful to know that there was a credit crunch in 1906 but it is also clear that the term itself was not used at that time. The last example in Figure 3 encapsulates this as it is a genuine example from the *Chicago Tribune* of November 16<sup>th</sup> 1967, which states that there was a credit crunch the previous year. The point is that this 1966 credit crunch appears to have been referred to as such only in retrospect.

Finding the earliest occurrence of a term with Google News is a rather laborious process. After finding the earliest genuine occurrence on the timeline by experimenting with different date ranges, it is necessary to switch back to the default view to determine if there are any earlier occurrences. As the default view does not show results in date order, all results must be examined. By carrying out this procedure, we found the earliest examples of *credit crunch* to be not the November example from the *Chicago Tribune* but the examples from earlier in 1967 shown in Figure 4.<sup>10</sup>

<p><i>New York Times</i>, June 4 1967 avoid a repetition of last year's <b>credit crunch</b></p> <p><i>Washington Post</i>, June 26 1967 highest interest rate since the 1920s - even a little higher than the rates late last summer during the <b>credit crunch</b></p> <p><i>Washington Post</i>, June 29 1967 Is the Nation heading into another <b>credit "crunch"</b> like last year's, with soaring interest rates, competition for savers' funds, and a new slump in the housing industry?</p> <p><i>New York Times</i>, June 30 1967 danger that we will be moving toward another "<b>credit crunch</b>". To avoid this, we urgently need greater fiscal restraint by the Federal Government</p> <p><i>Hartford Courant</i>, July 2 1967 Five change in federal housing laws, designed to prevent a <b>credit crunch</b> of the 1966 type, were proposed last week to a Senate committee by the National Assn. of Real Estate</p> <p><i>New York Times</i>, July 2 1967 Interest rates, the 'topic and concern of the and financial these days, have been climbing steadily and fears of a new <b>credit crunch</b> similar to last summer...</p>
--

Figure 4: Earliest examples of *credit crunch* in Google News, extracted manually

All the examples in Figure 4 refer to the credit crunch as something which happened the previous year. We cannot say so conclusively but, given that the

Google archive contains editions of these and other newspapers from 1966 yet returns no hits from that year, it seems likely that the term did not appear in the public domain in the United States until 1967. In the next section, we shall outline how WebCorpLSE, running on a combination of offline newspaper archives and newspaper data extracted from the web, can be used to trace the introduction of the term *credit crunch* in to the UK.

### 3.2 Newspaper corpora accessible via WebCorpLSE

We know from the BNC that the phrase *credit crunch* was used in the UK in 1991 but was not widespread and required explanation. Using the diachronic search facility in WebCorpLSE, we are able to trace the use of the phrase across a 25 year continuous span of UK broadsheet newspapers, segmented into months. The corpus contains 950 million tokens and consists of:<sup>11</sup>

- i) a complete archive of *The Guardian* (1984-88)
- ii) a complete archive of *The Independent* (1989-99)
- iii) *The Guardian*, downloaded from the web (2000-08)<sup>12</sup>

This corpus combines the two kinds of web-based diachronic analysis outlined by Nesselhauf (2007). On the one hand, the *Guardian* articles from 2000 onwards are pre-existing web texts. On the other, the early *Guardian* and *Independent* articles are off-line resources, being made available online in a form suitable for linguistic study by WebCorpLSE.<sup>13</sup>

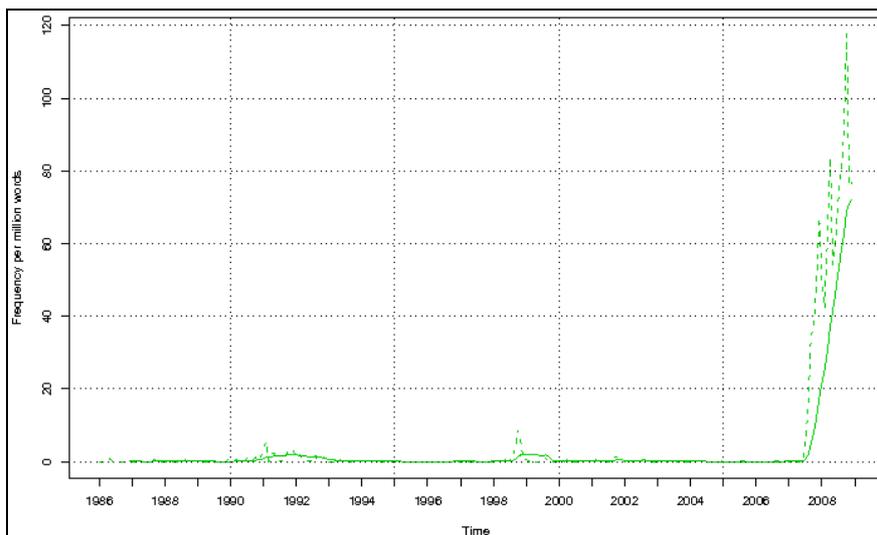


Figure 5: Frequency of *credit#crunch* across time in the WebCorpLSE newspaper archive (per million words)<sup>14</sup>

The graph in Figure 5 shows the frequency of *credit crunch* across time in the WebCorpLSE UK newspaper corpus. All frequencies are normalised to account for the varying size of the monthly segments across the years. The dotted line is the normalised monthly frequency and the solid line is a 12 month moving average. We have been examining such graphs for several years but have never seen a case as extreme as this, where the frequency increases from fewer than 1 occurrence per million words to almost 120 per million words within a single year. One of the earliest occurrences of *credit crunch* in the newspaper corpus is in a sentence from an August 1988 *Guardian* article, which includes a definition of the term and an indication of its origin:

Indeed there is a possibility of a US-style **credit crunch**, where interest rates are pushed up hard for a short period.

However, the phrase is used only 22 times in the 7 years before 1991 and the monthly frequency never rises above 1 per million words. The two noticeable ‘blips’ in Figure 5, prior to the massive upward trend in 2007-8, are accounted for by the concordances in Figures 6 and 7. These concordances were produced by WebCorpLSE, with sentence span selected and the results sorted by date, from earliest to most recent.

7170 instances of 'credit#crunch', case insensitive (sorted) (0 min 2 sec).	
23:	03/01/1991 Sir John also indicated during the programme that a <b>credit crunch</b> might be creeping into Britain, but denied that banks had adopted a policy of refusing to lend to the corporate sector
24:	04/01/1991 If they pull too many plugs on their customers, or if there is a <b>credit crunch</b> in which lending contracts sharply (see below), the recession could become a slump
25:	04/01/1991 WHILE a ' <b>credit crunch</b> ' is developing in the US and Japan because of the shortage of bank capital, the situation is not so dire here in Britain
26:	13/01/1991 The economy is weakening day by day as the <b>credit crunch</b> bites.
27:	14/01/1991 They want fewer doubtful loans and their reduced willingness to lend (compared with earlier cycles) constitutes the so-called ' <b>credit crunch</b> ' which is becoming an important factor in the recession
28:	14/01/1991 The failure last week of the Bank of New England reminds us that the risk of a <b>credit crunch</b> appears greatest in America where, as Professor Alan Budd pointed out some time ago (Barclays Economic Review, November 1990), the average credit rating of money centre banks has been downgraded, over the 1980 s, from triple-A to A3.
29:	17/01/1991 Mr Greenspan hinted at other signs of relief for the US economy, suggesting the <b>credit crunch</b> , which has persisted despite an easing in official US interest rates, may now be on the wane.
30:	20/01/1991 OWING to the world <b>credit crunch</b> and the 'Say no to the overdraft' movement among conscientious bank managers, the Chairman and the MD have still to raise their bail.
31:	29/01/1991 As analysts noted, Alan Greenspan, the Fed chairman, has publicly discussed the possibility of using other means than interest rates to reduce the impact of the <b>credit crunch</b> , in which banks are withdrawing from lending.
32:	29/01/1991 The temptation for the Fed, of course, is to talk confidence up, because lower interest rates have not been enough to get bank lending seriously under way again, to ease the <b>credit crunch</b> .
33:	02/02/1991 The Fed decision, which many expect to be followed soon by additional measures to ease the US <b>credit crunch</b> , led to cuts in the prime rates of leading US banks, from 9.5 per cent to 9 per cent
34:	03/02/1991 Tales of the <b>credit crunch</b> can also be overdone.
35:	11/02/1991 Over the weekend Alan Greenspan, chairman of the Federal Reserve, said the US had some way to go before it ended the ' <b>credit crunch</b> ' in which banks refuse to lend
36:	11/02/1991 He did not refer to lower interest rates, which are only one of the techniques for ending the <b>credit crunch</b> , but his remarks did nothing to counter the view that US rates could soon fall again
37:	11/02/1991 Others say their fellow economists were too blinkered in their methods to spot a <b>credit crunch</b> , in which banks withdrew from their role of financing expansion because of fright at the level of bad debts.
38:	15/02/1991 Although it is difficult to pin down evidence of a <b>credit crunch</b> in macro-economic figures, the Bank of England has decided there probably is one, but of a restricted kind which hurts small companies more than large

Figure 6: WebCorpLSE concordances for *credit#crunch* from *The Independent*, January-February 1991 (case insensitive, sentence span)

Figure 6 shows the occurrences of *credit crunch* in *The Independent* in early 1991 which were responsible for the increase in frequency to a peak of 5.4 per million words in February of that year. Again we see some occurrences in quotes, complete with glosses (lines 24, 27, 31, 35 and 37) and other lexical signals such as ‘so called’ (cf. Renouf & Bauer 2001 on ‘contextual clues’).

These concordances are contemporary with those from the BNC and it is clear from Figure 5 that, by chance, the BNC compilers captured a phrase, associated with a particular news story, which was at a peak of popularity in BrE.<sup>15</sup> This again highlights the limitations of short time-span synchronic corpora in lexical studies. A study of *credit crunch* based on data from the BNC may overestimate the significance of the phrase in late 20<sup>th</sup> Century BrE. In order to trace the development of a word or phrase fully, it is necessary to use a larger monitor corpus like the newspaper sub-corpus in WebCorpLSE.

138:	02/09/1998	MARTIN TAYLOR, chief executive of Barclays, warned yesterday that the world was heading for a 'credit crunch' because of Russia's financial crisis.
139:	02/09/1998	As Martin Taylor, chief executive of Barclays, argues, we may, after the latest traumatic events in Russia, be looking at a considerable, worldwide credit crunch.
140:	02/09/1998	RUSSIA'S FINANCIAL crisis could lead to a worldwide credit crunch, with serious knock-on effects on the real economy, Martin Taylor, the chief executive of Barclays Bank, claimed yesterday.
141:	04/09/1998	CATTLES, the personal loans provider, could be one of few companies to benefit if Martin Taylor, the chief executive of Barclays, is right about a global credit crunch.
142:	10/09/1998	No wonder so many bankers are talking about a world-wide credit crunch, stretching from the developing world right through to the major industrial nations.
143:	12/09/1998	For the first time since the Asian crisis broke, there seems to be a real risk of a world-wide credit crunch.
144:	12/09/1998	The calls were made as what one banker called 'a serious credit crunch' began to develop in the London and US banking system.
145:	14/09/1998	Struggling banks withdraw credit from good customers as well as bad ones, and there are clear signs of a 'credit crunch'.
146:	15/09/1998	WHEN BARCLAYS' chief executive, Martin Taylor, warned last month that the losses sustained by Western banks in Russia could spark a global credit crunch that could drive us into recession, not everyone was convinced.
147:	15/09/1998	There is definitely a credit crunch going on,' says credit analyst Daniel Sankey at Greenwich Natwest, the bond broking arm of National Westminster Bank.'
148:	15/09/1998	We now have a global credit crunch.
149:	19/09/1998	We already see evidence of this in plunging equity values, and we are beginning to see it in the onset of what some bankers regard as a very serious credit crunch.
150:	24/09/1998	Certainly looking at his evidence to the European Parliament, there is no recognition that (a) the world might be facing a crisis of deflation, and that falling prices would need different policies to rising prices, (b) there was danger of a 'credit crunch', where perfectly credit-worthy borrowers were denied funds because of a more generalised banking malaise, and (c) more than one-third of the world economy was already in recession.
151:	25/09/1998	A massive credit crunch is on its way.'
152:	27/09/1998	In particular there is the gravest danger of a 'credit crunch'.
153:	03/10/1998	But the crisis afflicting the banking sector is a financial one, requiring co-ordinated interest rate reductions and a loosening of monetary policy to avert a catastrophic credit crunch.
154:	04/10/1998	Of even greater immediate danger than a further slide in stock markets is a looming credit crunch.
155:	04/10/1998	Beyond the fading dreams of Christmas bonuses in the City, there are wider fears of a credit crunch.
156:	04/10/1998	Domestically the credit crunch keeps getting worse, and globally Japan's financial and economic problems are rippling through other markets.'
157:	04/10/1998	But is there nothing to be done, except to keep talking hot air and bailing out losers until funds are exhausted and the world economy is throttled by a credit crunch of proportions that have not seen since 1930?

Figure 7: WebCorpLSE concordances for *credit#crunch* from *The Independent*, September-November 1998 (case insensitive, sentence span)

After 1991, *credit crunch* appeared rarely (fewer than 50 occurrences in 6½ years) until late 1998 when it appeared 67 times in 3 months, including the cases shown in Figure 7. This 1998 peak in the frequency of the phrase appears to have

been sparked by comments from the chief executive of Barclays Bank (mentioned by name in lines 138-141 and 146). As in 1991, this peak in *credit crunch* was fleeting and the frequency of occurrence had fallen back below 1 per million words by December 1998. It then remained at that level until July 2007, when the massive increase in frequency began.

Turning to concordances from July 2007 (Figure 8), one is struck initially by the lack of quotation marks around *credit crunch* and lack of any explanation of the term.<sup>16</sup> It may appear that, by this point, the phrase has entered the lexicon of the newspaper to the extent that the journalists no longer feel it necessary to provide an explanation when using it. However, if we then look at a selection of concordances from later in 2007 and into 2008 (Figure 9), with the frequency of *credit crunch* continuing to rise, we find further examples where *credit crunch* is defined by the writer. We also see early evidence of the increasing trend for meta-linguistic discussion of the phrase *credit crunch* and its meaning.<sup>17</sup>

270:	01/07/2007	There are also fears a global <b>credit crunch</b> , sparked by the US sub-prime mortgage crisis, could have knock-on effects.
271:	08/07/2007	In the old days, if you had a <b>credit crunch</b> , the authorities knew where to go: they went to the lenders, and looked at their books,' says Andrew Clare, professor of portfolio management at the Cass Business School. "
272:	20/07/2007	There are now fears of a wider <b>credit crunch</b> . "
273:	20/07/2007	But we're not having a <b>credit crunch</b> yet."
274:	20/07/2007	Some analysts said they feared a broader <b>credit crunch</b> if a collapse in confidence in the US mortgage market rippled out to other parts of the debt markets.
275:	21/07/2007	There are now fears of a wider <b>credit crunch</b> . "
276:	21/07/2007	But we're not having a <b>credit crunch</b> yet."
277:	26/07/2007	A private equity <b>credit crunch</b> is now clearly putting at risk both the timing of the American beverages disposal and the consideration," said the broker. "
278:	26/07/2007	With worse-than-expected new housing sales in the US adding to the <b>credit crunch</b> concerns, Wall Street had slumped around 200 points by the time London closed.
279:	26/07/2007	Fears of a global <b>credit crunch</b> prompted by contagion from the crisis-hit US housing market led to panic selling and big falls on all the world's stock markets today.
280:	27/07/2007	By most definitions, that's a <b>credit crunch</b> .
281:	27/07/2007	That hasn't happened yet with EMI at 249 p, the market is implying there's a one-in-five chance but suddenly it is possible to see how a <b>credit crunch</b> could play out in real situations.
282:	27/07/2007	The White House tonight made a concerted attempt to inject fresh confidence into the world's battered stock markets as share prices suffered a fresh day of falls on fears that a looming <b>credit crunch</b> will end an era of cheap funding for corporate takeovers.
283:	27/07/2007	Sterling, which at the start of the week set fresh 26 year highs against the dollar of \$2.0655, fell to as low as \$2.0277 during trading today as investors dumped risky assets due to growing fears that the slump in the US housing market might spark off a global <b>credit crunch</b> .
284:	27/07/2007	Growing fears of a global <b>credit crunch</b> sparked panic selling on stock markets around the world yesterday, as investors rushed to dump shares.
285:	27/07/2007	Fears of a global <b>credit crunch</b> prompted by contagion from the crisis-hit US housing market led to panic selling and big falls on all the world's stock markets yesterday.
286:	27/07/2007	A private equity <b>credit crunch</b> is now clearly putting at risk both the timing of the American beverages disposal and the consideration," said the broker. "
287:	27/07/2007	Worse than expected new housing sales in the US and losses announced by US housebuilders added to the <b>credit crunch</b> concerns, pushing Wall Street down by about 200 points by the time London closed.
288:	27/07/2007	Worries about a <b>credit crunch</b> in the wake of the US sub-prime problems and difficulties in financing debt-laden deals like KKR's takeover of Alliance Boots continued.
289:	27/07/2007	One of the fears about the <b>credit crunch</b> is that it will prevent many rumoured takeovers from taking place, which is had news considering that much of the market has been undermined by bid hopes. "

Figure 8: WebCorpLSE concordances for *credit#crunch* from *The Guardian*, July 2007 (case insensitive, sentence span)

1:	05/08/2007	A " <b>credit crunch</b> ", properly understood, occurs when would-be borrowers with plans for productive capital expenditure are denied access to loans," he said. '
2:	13/08/2007	If the feared " <b>credit crunch</b> " happens where lenders get so nervous about extending credit that they stop lending then that would be a problem.
3:	04/11/2007	If they don't, they have to lend less and that is called a <b>credit crunch</b> .
4:	29/12/2007	When anxious queues of pensioners began forming outside Northern Rock branches in August, few of us even knew what sub-prime lending meant, but as the crisis in America's housing market crossed the Atlantic, money markets froze, banking bosses resigned, rumours of recession spread and "sub-prime" and " <b>credit crunch</b> " became the unhappy buzzwords of the year.
5:	30/12/2007	It was certainly a horrible year for housebuilders, with interest rate rises at the start of the year followed by the turmoil of the <b>credit crunch</b> , which raised the cost and reduced the availability of mortgages, leading to clear signs of a falling housing market and big falls in the shares of all building-related companies.
6:	16/03/2008	August 2007: The term ' <b>credit crunch</b> ' enters the national vocabulary as America's problems cross the Atlantic.
7:	24/03/2008	But since the subprime mortgage markets of the US started to go bad early last year and Northern Rock began to implode in September, the term " <b>credit crunch</b> " has been cropping up with depressing frequency.
8:	24/03/2008	It seems every piece of bad news can somehow be traced back to the " <b>credit crunch</b> I've even tried to blame it for my recent defeats on the squash court.
9:	29/03/2008	Once again you present the case of people with financial problems as if they are "victims" of the banks and the " <b>credit crunch</b> ", without any real consideration of how their plight developed.
10:	06/04/2008	Nine months into the <b>credit crunch</b> , which began with risky loans in America's sub-prime mortgage market, Britain's banks are still reluctant to lend to each other, and ordinary borrowers are paying the price.
11:	07/04/2008	This is the most concrete evidence to date that the esoteric " <b>credit crunch</b> " has moved out of the so-called "interbank money markets" and into the consciousness and pockets of the British people.
12:	09/04/2008	The " <b>credit crunch</b> " problem is international, originating in the US, and centring on interbank lending which has largely dried up as a consequence of the inability of creditors to locate bad debts toxic" mortgages securitised into complex instruments and now permeating the system.

Figure 9: Filtered WebCorpLSE concordances for *credit#crunch* from *The Guardian*, 2007-8 (case insensitive, sentence span)

A possible explanation for this lies in Figure 10, which shows the proportion of occurrences of *credit crunch* which appeared in each sub-section of *The Guardian*.

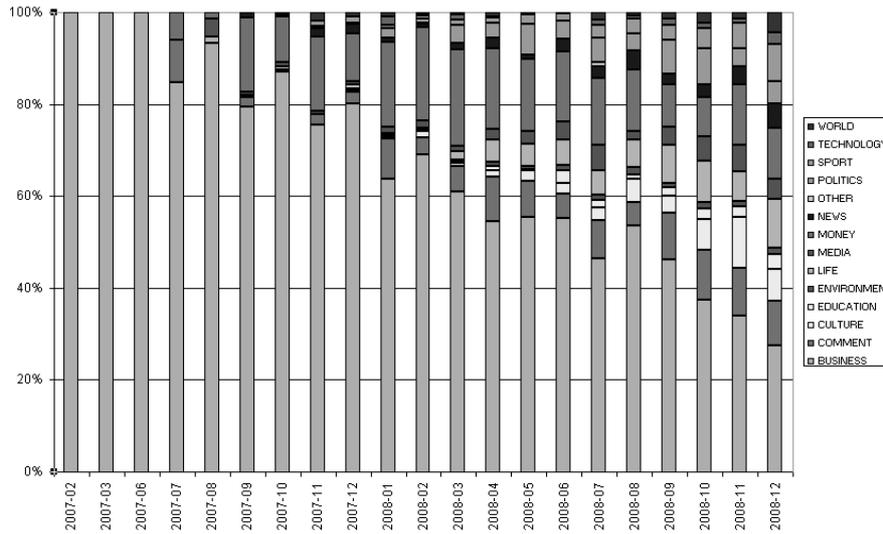


Figure 10: Proportion of occurrences of *credit#crunch* across sections of *The Guardian*, 2007-8<sup>18</sup>

In the early months of 2007, the phrase appeared only in the ‘Business’ section. By July 2007 it was also appearing in the ‘Money’ and ‘Comment’ sections, and by August it had spread to ‘Media’ and ‘Life’. Eventually, in December 2008, *credit crunch* was appearing in **all** sections of the newspaper, including ‘Sport’, ‘Education’ and ‘Culture’, thus confirming the notion in Figure 9, concordance 11 that ‘the esoteric “credit crunch” has moved out of the so-called “interbank money markets” and into the consciousness and pockets of the British people’.

### 3.3 Collocational analyses

Although the filtering options in WebCorpLSE can be used to make manual data analysis a more manageable task, the number of results can be prohibitively large when dealing with frequent lexis. In her 1987 study of ‘lexical resolution’, using a corpus of 13 million words, Renouf concluded that ‘eventually a point may be reached in corpus development where all word forms in which there is a lexicological interest are sufficiently exemplified’ (Renouf 1987: 130). It could be argued that we have now gone beyond this point, to a situation where corpora are so large that, for all but the rarest word forms, we are presented with more concordance data than can be analysed manually. As a result, statistical analyses have become increasingly important.

One way to examine the growth of *credit crunch* over time is to produce span 1 collocational statistics for one or both of the words which constitute the phrase. We have chosen to take *credit* as this is the more frequent of the two words in our corpus and we felt that an analysis of its collocates may provide more information about *squeeze* and other related words.

Figure 11 shows the span 1 collocates of *credit* for all months up to and including December 1988 (with a stopword filter switched on), whilst Figure 12 shows the same information but with the time period extended 20 years to the end of the corpus (December 2008). A z-score calculation is used to compare the expected frequency of collocation (based on the frequencies of each word) with the actual, observed frequency. Such collocational statistics are now standard in corpus linguistics and they are undoubtedly useful, as in this case where they reveal that *crunch*, which did not appear as a statistically significant collocate of *credit* in 1988, had become its most significant collocate by 2008. (In fact, viewed from the opposite perspective, *credit* accounts for 90% of the significant collocates of *crunch* in L1 – immediate left – position in the corpus as a whole.)

WebCorpLSE provides an enhanced collocation tool which allows the tracking of changes in collocational patterns across time. We refer to this as a collocational ‘heat map’, where heat is used as a metaphor for collocational strength. To generate a heat map, WebCorpLSE ranks all collocates of the target word in the whole corpus by z-score, and selects the top 200 significant collocates for further analysis. These are then broken down into groups by month and year to create a diachronic table of collocation frequency. The monthly z-scores are used to plot the strength of collocation on a graph by translating them into shades of red.

Collocate	L1	TOT	R1	Z-score
card	1	807	806	619.54
cards	6	507	501	419.92
consumer	225	225		183.95
Suisse		104	104	101.68
Family	102	102		92.72
family	193	194	1	87.76
boom		90	90	80.21
Consumer	80	80		76.48
rating		82	82	75.50
export	83	83		75.42
Export	72	73	1	71.41
Guarantee		64	64	62.78
Act		67	67	51.52
bank	71	74	3	46.36
Lyonnais		44	44	42.67
reference		48	48	41.36
scoring	1	47	46	40.94
controls		40	40	34.97
facilities		41	41	34.52
tax	62	65	3	33.65
lines		42	42	31.65
limit	2	35	33	28.89
insurance	1	37	36	26.85
balances		26	26	24.45
unions		31	31	24.30

Figure 11: Significant span 1 colls. of *credit*, up to end of 1988

Collocate	L1	TOT	R1	Z-score
crunch	3	7031	7028	4149.97
card	32	17249	17217	4010.65
Suisse	1	3311	3310	2739.34
cards	54	8825	8771	2478.06
Lyonnais	2	1638	1636	1461.61
rating	8	2053	2045	964.25
Consumer	1183	1186	3	774.05
tax	4329	4333	4	597.99
consumer	1430	1435	5	389.45
Agricole	1	369	368	355.72
Tax	486	490	4	334.70
Counselling	2	349	347	325.30
reference		751	751	260.96
deserves	518	518		256.05
pension	1048	1050	2	248.81
ratings	2	551	549	239.53
squeeze	4	424	420	233.01
export	508	509	1	217.58
balances	2	266	264	196.68
Card	1	236	235	196.14
interest-free	209	209		185.52
unions	9	684	675	183.78
Family	429	431	2	181.24
facility	2	343	341	178.28
markets	15	856	841	174.02

Figure 12: Significant span 1 colls. of *credit*, up to end of 2008

Figure 13 is a heat map for the span 1 collocates of *credit* from 1985 to 2008.<sup>19</sup> This output highlights the fine-grained approach to collocation provided by WebCorpLSE heat maps. We see *Lyonnais*, a strong span 1 collocate of *credit* for over 10 years, disappear from the map in 2003, at the point when the French bank Credit Lyonnais became known as LCL. We also see *Family* disappear and *Tax* appear in 1998-9, when the 'Working Families' Tax Credit replaced 'Family Credit' in the UK welfare benefit system. These are not linguistically interesting examples in themselves but they indicate that the methodology is sound and allow us to draw more meaningful conclusions when, for example, *reference*, *ratings*, *histories* and *limit* become strong collocates of *credit* (relating to 'debt worthiness').

Figure 13 also captures the cyclical nature of credit crunches, with *crunch* appearing as a significant collocate of *credit* for specific short periods (1991-2, 1998-9) before 'fading' out of use again. We also see *squeeze* appearing as a span 1 collocate of *credit* at similar, but not identical, points in time (appearing more gradually from 1988-91, and weakly in 1993-4 and 1998-9). We shall examine *squeeze* in section 4.

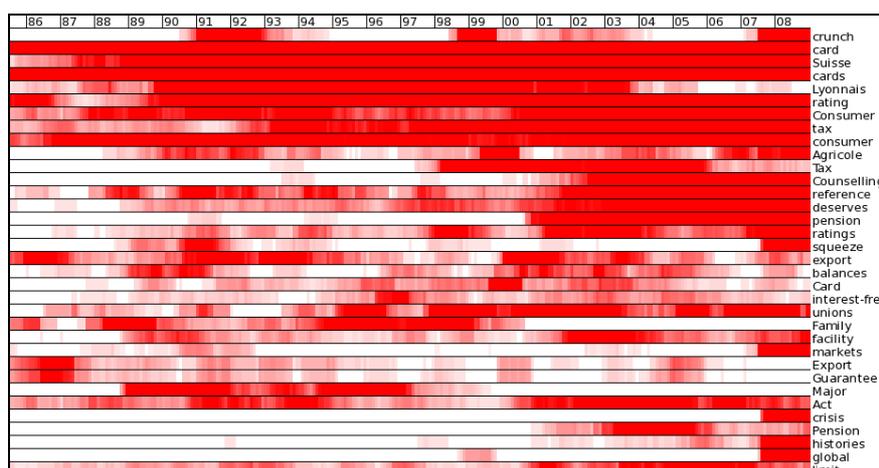


Figure 13: Top of 'heat map' for span 1 collocates of *credit* (case insensitive) in newspaper corpus 1985-2008 (left and right collocates)

Both *crunch* and *squeeze* re-emerge as strong collocates of *credit* in 2007-8 and it remains to be seen how long this particular event will last. Given that the phrase *credit crunch* is being used more frequently than ever before and that collocates indicating severity (*global*, *crisis*) also appear as strongly significant in 2007-8, it would seem that it will be much longer before this particular credit crunch fades from the heat map.

We should also note in our discussion of collocation that WebCorpLSE allows the generation of collocates for any search term and is not restricted to single words searches. Figure 14 shows the span 4 collocates of the phrase *credit crunch* over time. Until 2007, the phrase had few statistically significant collocates, though *banks* first appeared in 1991 and *global*, *fears* and *markets* had appeared by 1999 (the time of the second 'blip' in Figure 5). By 2008, there is a long list of words describing the credit crunch, its causes and effects, some of which are classed as significant as a result of their own newness and rarity (*sub-prime*, *write-downs*). It will be interesting to monitor changes in the collocational profile of *credit crunch* in future years.

#### 4. A brief discussion of *credit squeeze*

Space does not permit a full discussion of *credit squeeze* but we have conducted a diachronic analysis of the phrase and will summarise the main findings here. Unlike *credit crunch*, *credit squeeze* does appear in the OED, under the headword *credit* (Figure 15).



The whole phrase does not appear in the Brown family of corpora but there is one occurrence of *squeeze* in this sense in the BrE LOB corpus:

The big “**squeeze**” means that it is going to be more difficult to arrange a loan or overdraft. (A06 206-207; *Daily Sketch*, 4 August 1961)

The phrase is not quite as frequent as *credit crunch* in the BNC, appearing 13 times in texts from 1976-93 (Figure 16).

1	In 1974 his property and investment group also faced problems brought on by a <b>credit squeeze</b> and downturn in the building market. (AAS 11: <i>Guardian Business</i> section, 31/12/89)
2	The capital standards, negotiated through the Bank for International Settlements (BIS), are a natural scapegoat for the <b>credit squeeze</b> that is deepening the recessions in Britain and America and may provoke one in Japan. (ABE 159: <i>Economist</i> , 1991)
3	The higher interest rates and <b>credit squeeze</b> control used by the Conservatives did, however, slow down growth in the economy overall. (CRD 480: <i>Engineers, managers and politicians</i> , 1993)
4	The Conservatives had clearly let the economy overheat for electoral advantage in 1955, but as soon as the election was over, clamped down with a <b>credit squeeze</b> . (CRD 559: <i>Engineers, managers and politicians</i> , 1993)
5	Foreign business also has a more practical complaint: because of China’s <b>credit squeeze</b> , bills are no longer paid on time. (EDU 578: <i>Marxism Today</i> )
6	In Britain the apparently smooth growth during the long boom was marked by dramatic events that, at the time, seemed to be crises: for example, the 1957 <b>credit squeeze</b> and record interest rate jump (FA0 588: <i>Restructuring Britain: the economy in question</i> , 1988)
7	In a less obvious but equally influential manner, if a <b>credit squeeze</b> is applied as a macroeconomic policy, the resulting high interest rates will reduce the number of people able to take out mortgages. (FB2 719: <i>Rural Britain: a social geography</i> , 1985)
8	It won’t be affected by the <b>credit squeeze</b> ...? (G0F 1376: <i>Sweet dreams</i> , 1976)
9	This is true in that consumer demand has collapsed as a result of the <b>credit squeeze</b> (G38 485: <i>Marketing Week</i> , 17/01/92)
10	Britain therefore experienced a <b>credit squeeze</b> in the early 1990s during a period of recession in much the same way -- and for much the same reasons -- that she experienced a credit boom during the period of growth and “overheating “ in the mid-1980s. (H91 296: <i>A treaty too far</i> , 1992)
11	The government responded to the payments crisis with a <b>credit squeeze</b> . (K8U 225: <i>Capitalism since 1945</i> , 1991)
12	This situation would occur in circumstances as in the late 1960s, when due to a <b>credit squeeze</b> , interest rates rose. (K8W 1292: <i>UK financial institutions and markets</i> , 1991)
13	Second and simultaneously, in order not to release a consumer <b>credit squeeze</b> that would second imports, they should introduce controls on the supply of credit (KRT 3495: Fox FM News: radio programme)

Figure 16: All examples of *credit squeeze* from the BNC

It is noticeable that *credit squeeze* appears far more in the BNC in books, discussing past events (sources underlined in Figure 16), than in news stories. These results are significantly different from those for *credit crunch* and may indicate that *crunch* was in the process of replacing *squeeze* in this context in BrE texts discussing current events. We cannot, of course, draw this conclusion purely from an analysis of the BNC or other standard corpora, for reasons outlined above. However, a diachronic analysis of our UK newspaper corpus using WebCorpLSE (Figure 17) does provide further evidence for this.

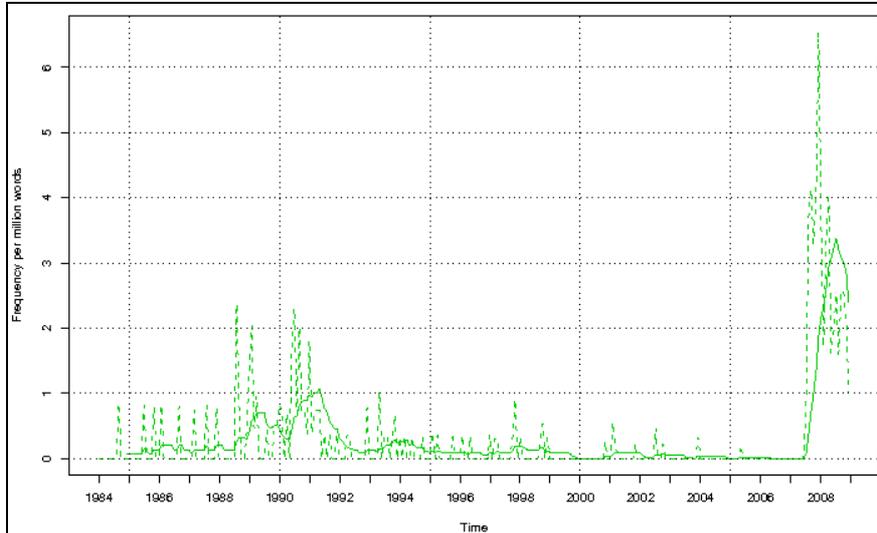


Figure 17: Frequency of *credit#squeeze* across time in the WebCorpLSE newspaper archive (per million words)

The phrase *credit squeeze* appears in the newspaper corpus earlier than *credit crunch* (1984 versus 1987) but there are only 422 occurrences of the former, compared with 7069 of the latter, and *squeeze* does not reach the same peaks in frequency reached by *crunch*.

We also used Google News to extract the earliest occurrence of *credit squeeze* in newspapers, in the same way described above for *credit crunch*. This revealed the earliest occurrences to be in two *New York Times* articles from 26<sup>th</sup> March 1929 (complete with Google OCR errors):

alt of which have been recently b3 the stock market. threw out the intimation that a **credit squeeze** of major proportions was inevitable if the use of ...

The tightest **credit squeeze** in almost nine years tools place On the S:Ork Stock Exchan=a yesterday, when the call loan rate advanced to 74 per cent

These early occurrences of the phrase in AmE are contrary to the claim in the *Cambridge Advanced Learner's Dictionary* (quoted in Section 1) that *credit squeeze* is a UK term, equivalent to the US *credit crunch*. It is, of course, conceivable that *credit squeeze* was once the preferred term in AmE and that, at some point after the coining of the phrase *credit crunch* in the US in 1967 and before the earliest articles in our newspaper corpus (1984), *credit squeeze* was still used more widely than *credit crunch* in BrE. What is certain from our analysis is that, given the recent **global** credit crunch and massive increase in usage of the phrase in UK newspapers, this distinction between UK and US usage no longer holds true.

It is beyond the scope of the current paper to examine the semantics of the two phrases in depth, a task which would require economic as well as linguistic insight, but the phrases *credit squeeze* and *credit crunch* do not appear to be as synonymous as the *CALD* definition implies. It is clear from the OED citations and LOB and BNC concordances that, from the 1950s-1990s, a credit squeeze was a measure applied by a government as a deliberate economic policy. A credit crunch, in its most recent incarnation at least, is something over which governments seemingly have little control.

## 5. From the *credit crunch* to the *crunch*

As we have seen, the vast increase in use of the phrase *credit crunch* in mid-2007 was mirrored by an increase in the less used *credit squeeze*, with both phrases being used to describe the same event.

During the same period, we have also noted an increase in the elliptical form *the crunch* and have examined this by using the date filter option in WebCorpLSE to view all occurrences of the phrase in *The Guardian* from 2007-8. These were then analysed manually and divided into five categories:

- i) *crunch* as a premodifier (e.g. *the crunch vote*, *the crunch game*)
- ii) *the crunch* referring to the credit crunch
- iii) COME+to *the crunch* (including *the crunch came*, etc)
- iv) literal crunch (*the crunch of gravel*)
- v) other

A graph of the results (Figure 18) reveals that, whilst the other meanings have remained constant, *the crunch* as an abbreviated form of *the credit crunch* has increased in frequency following first occurrence in July 2007.

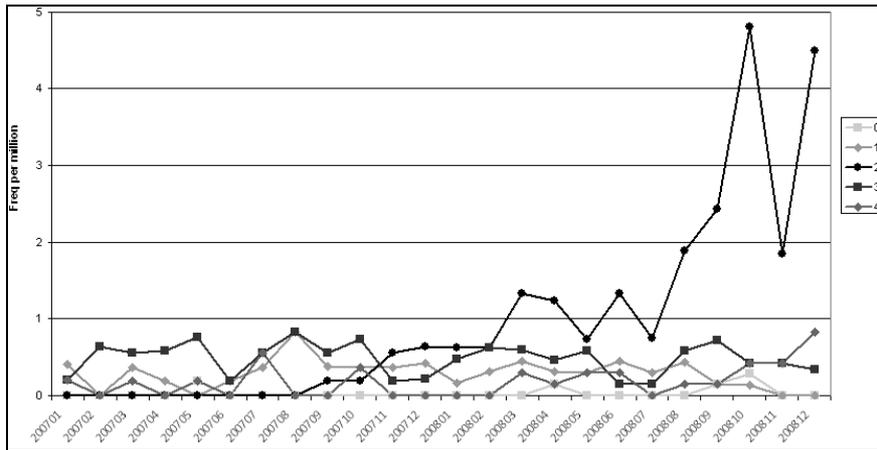


Figure 18: Frequency of *the crunch*, 2007-8 (per million words), differentiated by sense: 0=other, 1=premodifier, 2=credit crunch, 3=COME to crunch, 4=literal

The manual analysis also revealed some creative uses of *the crunch*, where two meanings have been conflated by journalists for effect, including:

1. Analysts at Evolution Securities said the worst was still to come, with **the “crunch”** for Greene King and other licensed retailers arriving this winter and next spring. (04/07/08)
2. When it comes to the crunch, price matters. (07/12/08)
3. The worry is that when it comes to the crunch multinationals will close overseas plants rather than domestic ones and overseas utilities will not pass on cost decreases arising from oil. (08/12/08)

Examples 2 and 3 here are from articles about the credit crunch and the use of the idiom ‘when it comes to the crunch’ appears to be a conscious decision by the writer, certainly so in 2, a sub-headline. The writer of example 1 uses the COME+to *the crunch* construction (and signals the play on words with the double quotes around *crunch*) but then selects *arriving* instead of *coming*. We would suggest that this was a deliberate choice by the journalist (or possibly a sub-editor) to ensure that the ‘credit crunch’ meaning was not ‘lost’ in the idiom.

There appear to be two factors driving the growth in the ‘shorthand’ form *the crunch*. Firstly, journalists tend to tire of ‘buzz’ phrases quickly and begin to look for ‘snappier’ alternatives. Secondly, the vast increase in usage of the phrase (*the*) *credit crunch* over a relatively short period of time has left it (and the associated concept) in the public consciousness to such an extent that the shorthand form *the crunch* is interpretable instantly, without a gloss.

## 6. Conclusion

In this paper, we have illustrated how the web can be used to supplement usage examples from standard corpora in diachronic linguistic analysis. When considering a recent linguistic phenomenon such as the rise of *credit crunch*, the web offers a solution to the restrictions posed by the ‘dearth of corpora of English spanning the whole of the twentieth century, or more particularly spanning the early part of it’ (Leech 2005: 85).

We have shown that, through careful data selection and the use of advanced diachronic analysis tools in WebCorpLSE, it is possible to widen the focus and trace the development of a word or phrase across the twentieth century, in British and American English. Our analysis of *credit crunch* and associated phrases has highlighted the value of Google News as a repository of twentieth century texts, but has also revealed the limitations, for linguistic search, of the search software provided by Google. The ideal solution would be to access the Google News archive via WebCorpLSE or other similar interface, thus allowing full-scale diachronic linguistic search of twentieth century newspaper text. Of course, newspaper corpora are not an ideal data source for the analysis of all kinds of linguistic phenomena but, as Hundt and Mair (1999) point out, newspapers are usually at the forefront of linguistic change and are, thus, a valuable resource in the kind of linguistic analysis carried out in this paper.

Our analysis has focussed on usage patterns rather than semantics but the work has allowed us to make some observations about the meaning and status of the phrase *credit crunch* and of *crunch* individually, as relates to *squeeze*. In fact, our analysis of the ‘shorthand’ form *the crunch* in *The Guardian* uncovered a meta-linguistic discussion of *crunch* to which we now refer in conclusion:

What exactly is a crunch?

Crunch in this context has two meanings, the first being “critical moment”, as in “coming to the crunch”. This is the older meaning of the two, almost certainly dating to Winston Churchill’s use of it in a 1939 *Daily Telegraph* interview. [...] The second, more modern meaning is the sense of “squeeze”, arising from paucity – this is how we get “energy crunch”. [...] Generally, the two meanings bisect, so the word conveys an urgent scarcity. [...] But the two meanings have not yet coalesced entirely.

(Zoe Williams, *The Guardian*, 7 January 2008)<sup>20</sup>

What this journalist refers to as the ‘more modern meaning’ is the wider AmE use of *crunch* which was already apparent in the 1992 Frown examples discussed in Section 2. This meaning has apparently made its way into BrE as a result of the massive surge in frequency of *credit crunch*. Prior to 2008, the ‘paucity’ example, *energy crunch*, had only appeared in our newspaper corpus 7 times, but there were then 13 occurrences in 2008 alone (3 of which appeared days before Zoe Williams’ comments and are apparently what sparked them). As we noted in

section 3.3, 90% of the immediate left collocates of *crunch* in our newspaper corpus are accounted for by case variants of *credit*, so there is little evidence for the wider use of *crunch* to refer to other kinds of ‘squeeze’ at present. Apart from *energy crunch*, we do note a handful of occurrences of other crunches in 2007-8 data (*supply, pensions, housing, oil*). In our analysis of *the crunch*, we also note two examples where *crunch* appears to fill a slot more commonly filled by the semantically related *pinch*:

1. Harriet Harman, has repeatedly and patronisingly said that “ordinary” families are feeling the crunch from rising fuel and food prices (06/05/08)
2. Budget hotels are raking it in as business people feel the crunch (05/10/08)

The second example here could be interpreted as ‘feel the effects of the credit crunch’, but the first is seemingly equivalent to ‘feeling the pinch’ (the use of *from* rather than *through* precluding the interpretation ‘feeling the credit crunch’). This use of *crunch* is reminiscent of its use in a wider financial sense in the AmE concordances from Frown (Figure 2), where it does indeed convey both scarcity and urgency.

This paper has traced the assimilation of the phrase *credit crunch* in to BrE. During the 1990s, the phrase was used periodically but infrequently in UK newspaper texts, reflecting the cyclical nature of the economic phenomenon it describes. As a result, each time the phrase re-emerged, journalists found it necessary to provide a full gloss. Since mid-2007, however, *credit crunch* has increased in usage to such an extent that the elliptical form *the crunch* is now interpretable immediately by the UK public. In fact, as a result of the spread of *credit crunch*, the word *crunch* is itself beginning to take on new meanings, including some not linked directly to the financial domain. It, thus, seems unlikely that the phrase *credit crunch* will require a gloss if it is to re-emerge once again in future years.

#### Acknowledgement

Development of WebCorpLSE was in part funded by the UK Engineering and Physical Sciences Research Council (EPSRC), grant reference EP/E001300/1.

#### Notes

- 1 A recent upgrade to the original WebCorp system (and renaming to ‘WebCorp Live’) has increased processing speed, but the reliance on commercial search engines remains and the range of searches possible is thus still limited. We are maintaining the original WebCorp system for the benefit of those users who wish to conduct ‘live’ searches of the ‘whole’ web, as accessible through commercial search engines.
- 2 <http://www.askoxford.com/worldofwords/wordfrom/wordsoftheyear2008>
- 3 [http://www.cambridge.org/elt/corpus/international\\_corpus.htm](http://www.cambridge.org/elt/corpus/international_corpus.htm)

- 4 For each occurrence, the BNC file and line number are given in parentheses. Concordance lines are grouped according to the publication and article from which they are taken (the latter extracted manually from the source files). The BNC was designed as a synchronic corpus and is not ideally structured for diachronic study. For example, the file ABD contains 9 occurrences of *credit crunch* but it is not immediately clear that the last 8 of these all occur in the same article. Nor is it clear on exactly which day each newspaper article was published and, in some cases there is no date information at all other than a wide range (e.g. the article from *The Scotsman* in figure 1: 1985-1994). Results are presented in figure 1 in BNC file order, which is not necessarily date order.
- 5 The Frown manual (<http://khnt.hit.uib.no/icame/manuals/frown>) reveals the sources of these examples to be:  
**A25:** Press: Reportage: *San Francisco Examiner*: ‘S.F. Supervisors Crack Down on Use of City Cars’ (06/10/92).  
**C12:** Press: Review: *Wall Street Journal*: ‘The Persecution of Milken’ (25/08/92).  
**G23:** Belles Lettres, Biographies, Essays: Ruth Conniff ‘The Culture of Cruelty’, *The Progressive* (09/92).
- 6 See <http://googleblog.blogspot.com/2008/09/bringing-history-online-one-newspaper.html>.
- 7 There are several limitations, some of which we go on to outline below. The main limitation is that, at present, the compilers of the Google News Archive are focussing their attention, for the earlier periods of history, on US newspapers. This is not so much of a problem in our case, since we are searching for a term which we believe to have originated in the US.
- 8 <http://news.google.com/archivesearch>, The searches discussed in this paper were carried out in January 2009.
- 9 The Google News results pages carry the disclaimer “Dates associated with search results are estimated and are determined automatically by a computer program”. Kehoe (2006) detailed the ways in which a computer program could estimate the authorship date of web texts for use in linguistic analysis, with a high accuracy rate. Newspaper articles contain far more reliable dating information than web pages, so it is unlikely that Google’s program is wildly inaccurate when estimating these dates. It is simply estimating dates for a different purpose.
- 10 Note that, in most cases, the full text of matching articles is not available. In some cases, a sentence context is available by following the link to the corresponding newspaper archive. In other cases the limited context on the Google News results page is all that is available. Figure 5 shows the

widest context available. There is an apparent OCR error in the last context shown.

- 11 Though the corpus comes from two different broadsheet newspapers, these are broadly comparable in terms of content, focus and style.
- 12 Including its Sunday sister newspaper *The Observer*.
- 13 Though *The Guardian* has an archive on its website, this is complete only from 1999 onwards. Only a selection of the 1984-88 articles in our corpus is available on the *Guardian* site and *The Independent* does not have a freely accessible archive at all. WebCorpLSE makes limited contexts available from these sources, to registered users only.
- 14 The # operator in WebCorpLSE matches the three variants ‘credit crunch’, ‘credit-crunch’ and ‘creditchunch’, a useful option when searching for compounds. As it transpires, the last of this does not occur in our corpus. We use the *credit#crunch* query syntax throughout this paper. This particular search is also case insensitive.
- 15 The same is perhaps also true, to a lesser extent, for the Frown corpus, its 1992 AmE texts capturing a credit crunch in the US economy at that time.
- 16 Line 280 (‘By most definitions, that’s a credit crunch’) is a possible exception. However, we would not class the sentence immediately before this (‘Right now, big buy-outs are impossible: the debt markets are closed until the jam clears’) as a clear definition of the term. The concept of ‘credit crunch’ is not presented in this article as something which may be unfamiliar to the reader.
- 17 This concordance selection was made possible by the ‘filter’ option in WebCorpLSE, which allows manual removal of individual concordance lines, filtering by date, etc.
- 18 Some of the categories in this chart are composed of several sub-sections on *The Guardian* website: **COMMENT:** Comment, Letters; **CULTURE:** Artanddesign, Arts, Books, Culture, Film, Music, Stage; **LIFE:** Lifeandhealth, Lifeandstyle, Cars, Society, Travel, Weekend; **NEWS:** News, UK News; **TECHNOLOGY:** Science, Technology; **WORLD:** EU, Global, International, USA, World.
- 19 We have included both left and right span 1 collocates for illustrative purposes. WebCorpLSE allows the analysis of right and/or left collocates at spans 1-9 and sentence span. It is possible to conflate the frequencies of case variants, separate part-of-speech variants (e.g. separate entries for *crunch\_NN* and *crunch\_VV*) or view POS collocates only.
- 20 <http://www.guardian.co.uk/business/2008/jan/07/creditchunch.zoewilliams>

**References**

- Hundt, M. and C. Mair (1999), 'Agile and uptight genres: The corpus-based approach to language change in progress' *International Journal of Corpus Linguistics* 4, 221-242.
- Kehoe, A. & M. Gee (2007), 'New corpora from the web: making web text more "text-like"' in: P. Pahta, I. Taavitsainen, T. Nevalainen & J. Tyrkkö (eds.) *Towards Multimedia in Corpus Studies*, University of Helsinki: [http://www.helsinki.fi/varieng/journal/volumes/02/kehoe\\_gee](http://www.helsinki.fi/varieng/journal/volumes/02/kehoe_gee)
- Kehoe, A. (2006), 'Diachronic Linguistic Analysis on the Web using WebCorp' in: A. Renouf & A. Kehoe (eds.) *The Changing Face of Corpus Linguistics*, Amsterdam: Rodopi, 297-307.
- Kehoe, A. & A. Renouf (2002), 'WebCorp: Applying the Web to Linguistics and Linguistics to the Web', in: *Proceedings of WWW 2002*, Honolulu, Hawaii. Electronic publication: <http://www2002.org/CDROM/poster/67>
- Leech, G. and N. Smith (this volume), 'Change and constancy in linguistic change: How grammatical usage in written English evolved in the period 1931-1991'.
- Leech (2005), 'Extending the possibilities of corpus-based research on English in the twentieth century: A prequel to LOB and FLOB', in: *ICAME Journal No. 29*.
- Mair, C. (2007), 'Change and variation in present-day English: integrating the analysis of closed corpora and web-based monitoring', in: M. Hundt, N. Nesselhauf & C. Biewer (eds.) *Corpus Linguistics and the Web*. Amsterdam/New York: Rodopi, 233-247.
- Nesselhauf, N. (2007), 'Diachronic analysis with the internet? *Will* and *shall* in ARCHER and in a corpus of e-texts from the web', in: M. Hundt, N. Nesselhauf & C. Biewer (eds.) *Corpus Linguistics and the Web*. Amsterdam/New York: Rodopi, 287-305.
- Renouf, A. & L. Bauer (2001), 'Contextual Clues to Word-Meaning', *International Journal of Corpus Linguistics*, Vol. 5 (2), Amsterdam/Philadelphia: John Benjamins, 231-258.
- Renouf, A. (1987), 'Lexical Resolution', in: W. Meijs (ed.) *Corpus Linguistics and Beyond: Proceedings of the Seventh International Conference on English Language Research on Computerized Corpora*. Amsterdam: Rodopi, 121-131.