The affordances of metaphor for diachronic corpora & discourse analysis

Article (Accepted Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/104995/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
The affordances of metaphor for diachronic corpora & discourse analysis: WATER metaphors & migration

Abstract: This paper examines the utility of metaphor as an investigative tool in “long-distance” corpora and discourse studies. I show that metaphor is both important for understanding discourses and useful for diachronic analysis because it allows us to abstract out above the purely lexical level which enables comparison across contexts where the same concept could be lexicalised differently. The case-study is concerned with the oft-discussed metaphor of MIGRANTS ARE WATER in the UK-based Times newspaper over the period 1800-2018 and the conventionalisation and evaluative patterns are presented. The findings confirm that the WATER metaphor has an extensive discourse history with regards to how migration is represented in the UK press but show the evaluations of may differ significantly. The case-study shows how metaphor can give us a way to find discourse evaluations and framings across different time periods. The use of second-order collocates illustrates how corpus tools can help re-contextualise data to ensure interpretation takes into account contemporary framings.

Keywords: metaphor: migration: diachronic analysis: discourse: collocation

1. Long-distance diachronic corpora & discourse analysis

Corpus linguistics has a long-standing tradition of analysing historical corpora to empirically trace language change and development. And yet, as McEnery & Baker (2016: 3) show, “[i]n spite of a large body of work in linguistics in general, and in corpus linguistics in particular, using corpora to explore the past is still in its infancy”. This infancy is not because researchers have neglected the past – far from it – but because of limitations in access to corpora. The problem holding back research into historical language has been the relatively small size of the corpora available. According to McEnery and Baker (2016: 3-6) this creates two related problems. First, “the issue of size becomes acute when one tries to use such corpora to explore words of what one might describe as moderate or low frequency; there
simply not enough data to begin to make generalizations of any worth” (2016: 4). This means that research has been more likely to focus on high-frequency grammatical items and less likely to address aspects of meaning of greater centrality to the discourse endeavour. Second, the corpora are likely to have sampled short periods which means that change over time may not be addressed in a nuanced way and so the scope for diachronic accounts of stasis and change are inevitably limited. Furthermore, where sampling is used, it may omit a period when a term was mentioned frequently. As a result, “corpora which cover long periods of time with little data and few sample points are of limited use – or no use at all – for exploring changes in language that may be relatively transitory and dynamic. Trying to explore such features using such corpora can lead to highly misleading results” (2016: 10). It is this latter point that is particularly problematic because the research can be entirely transparent and accurate but, if based on partial datasets, may be misleading.

As the history of science shows us, science has leapt forwards when there has been a shift in access to data or tools for viewing data. Two developments have changed the relationship of corpus linguistics to historical data in recent years. First, the kind of corpora to which we have access. The special issue of *Corpora* which launched ‘Modern Diachronic Corpus-Assisted Discourse Studies’ (MD-CADS, Partington 2010) was driven by the compilation (by the contributors) of two corpora containing the entire output of three British broadsheets from 1993 and 2005, and this changed the research avenues available to them:

> Having at one’s disposal very large corpora […] opens up entirely new avenues of research in modern diachronic linguistics: we can study changes in meaning, especially of sets of lexical items in relation to both internal linguistic factors and also in response to external social influences. Being in a position to study lexical patterns and how they differ in the two corpora, we are also able to study changes in discourse processes. Partington (2010: 86)

With reference to corpora covering older data, the impact is even greater given the relative scarcity of data from this time. Crucially, the development of diachronic historical corpora such as Hansard (1803-2005), Times Online (1785-2011) and the Old Bailey Corpus (1720-1913) means that we have access to corpora which are *continuous* and *complete*, as opposed to the time and text-sampled corpora which were used in much previous work. We also have corpora like EEBO which provide a far greater range of text types than has previously been possible for these time periods. The second development we might expect to affect the
relationship of historical data and corpus linguistics is the increased recognition of the importance of discourse-level analysis within corpus linguistics (as witnessed by the *Corpora & Discourse* conference series and newly-established *Journal of Corpora & Discourse*).

A brief survey of recent work shows that the development of “long-distance” corpora, which contain the entire output of a source over an extended time period (50 years or more), has widened the scope of extended diachronic investigations for grammatical and lexical variation. For instance, with reference to the Old Bailey corpus, Claridge et al. (2019) show an increase in the use of maximisers (e.g. “perfectly” or “entirely”) and Widlitzki et al. (2016) show a decrease in swearing over time. And yet, developments in tackling discourses from a historical-diachronic approach remain relatively thin on the ground. Exceptions include studies of discourses around prostitution (Baker & McEnery 2017a) and homosexuality (McEnery & Baker 2017b) across the 17th century (both using EEBO), Islam in the nineteenth and twentieth centuries (Baker, Gabrielatos & McEnery 2013, using the British Library’s nineteenth century newspaper collection), bisexuality (Wilkinson 2019) and migration (AUTHOR 2020) in the 1950s and 2010s (both using the Times corpus).

As these discourse studies show, a major affordance of a long-distance view is the ability to *denaturalise* the discourse in question. As discourse analysts, we are inevitably part of our own object of study when we examine contemporary discourses and our readers are likely to be in the same position. Thus it can be difficult to identify and communicate the ways in which discourses represent choices and that, as Fowler (1991: 4) put it: ‘there are always different ways of saying the same thing, and they are not accidental alternatives. Differences in expression carry ideological distinction (and thus differences in representation)’. When we incorporate historical analyses into a diachronic study, or when we shift entirely to a historical period, we are no longer native speakers of our texts and this distance can bring insights (as well as challenges). The historical insights can help us uncover the ways in which discourses have been shaped as when Wilkinson (2019) traces the development of erasure in representation of bisexuality in the Times. We can empirically test claims about modern-day discourse, as in AUTHOR (2020) who reveals the nostalgic nature of government rhetoric about Windrush migrants. We can show how present discourses “rhyme” with those of the past and whether a “representation is new, or simply a modern version of a representation that has deep historical roots” (Baker, Gabrielatos & McEnery 2013: 230).
Thus, these new continuous and complete corpora offer a rich opportunity for understanding discourse developments and life-cycles but, inevitably, they also present new challenges relating to both magnitude and time-scale. In part, the difficulties are driven by the same features which allow us distance to see discourse – in effect, when we undertake historical discourse analysis we are undertaking a cross-cultural study. Furthermore, as linguists we are well aware that that words change their meanings over time and that meanings change their words (i.e. onomasiological and semasiological variation). The challenges raised include, but will not be limited to, the following:

i. How do we trace change and stasis over the entire period?
ii. How do we know that we are comparing like with like over such an extended time period?
iii. How do we achieve contextually-relevant interpretation for multiple time periods?
iv. How do we identify salient points for detailed analysis?
v. How do we handle significantly different quantities of data in different time periods?

The first two relate particularly to the semantic variation we will encounter while the third is about our ability to develop intercultural competence. These are the questions that I will aim to address in this paper. The fourth, is addressed in Gabrielatos et al.’s (2012) model of peaks and troughs, and McEnery et al.’s (2019) development of Meaning Fluctuation Analysis which provide replicable and data-driven means of identifying points for attention. The fifth is addressed in Wilkinson (2019) who illustrates how different methodological approaches of concordance and collocate analysis can be combined.

2. Metaphor as a tool for corpora & discourse analysis

Metaphor may be broadly defined as bringing together “different areas of experience and knowledge so that a particular topic is cognitively and communicatively presented in terms of another topic” (Musolff 2016: 7). Metaphors act as a way of understanding the world and so the choice of any given metaphor offers up an interpretation of the target. Interpretation is key because metaphors by their very nature are not neutral. When we liken one thing to another we do so on a partial basis; or, as Semino et al. (2018: 29) put it, metaphor use “highlights some aspects (the similarities that can be established between the two) and backgrounds others (things that are different or irrelevant for the comparison). This helps communicate the […] evaluation, and can facilitate some inference while making others
unlikely”. It is this intrinsic evaluative potential which makes metaphor so well adapted to revealing discourses (see also AUTHOR 2013). A significant part of a metaphor’s persuasive power lies in its ability to evoke emotions (e.g. as discussed in Nguyen and McCallum 2016: 167). Furthermore, experimental research has provided evidence for a legitimating impact of metaphor as in Hart (2018) who found that participants presented with FIRE metaphors in news reports were more likely to consider police use of water cannon as legitimate compared to participants presented with a non-metaphorical text.

As noted above, when we compare data in a long-distance diachronic study, the language used in the different points of the corpus will vary considerably and so identification of comparable units becomes especially salient. In addressing this challenge, we can draw on cross-linguistic work which has had to respond to these challenges already. AUTHOR (2020) survey a number of methods for ensuring that we are indeed comparing like with like. The first of these involves comparison of semantic categories which allows the comparison to move above a focus on words which are very likely to have shifted in meaning and frequency of use. For instance, Archer (2014) makes use of the UCREL Semantic Annotation System (USAS), which assigns semantic tags to lexical items, to identify aggression in the Old Bailey corpus. Alternatively, researchers may manually group collocates of node items into semantic sets in order to abstract out above the purely lexical level. For instance, McEnery & Baker (2017a) categorise collocates of harlot, jilt, prostitute, strumpet and whore in order to identify patterns in representation across an entire century. A second kind of approach would make use of discourse frames, so the comparison occurs after the discourse analysis of individual time periods. For instance, one could apply the moral panic frame (Cohen 1972, adapted for linguistic analysis in McEnery 2005) to representation of a given group in different time periods and then compare the frames to identify stasis and change over the period. Similarly, the analysis of conceptual or macro-level metaphor gives the researcher a means for abstracting out above the lexical level so that findings can be compared. For instance, while the words used to indicate a DISEASE metaphor might change over time (e.g. grippe, influenza or flu), the metaphor itself can remain stable (AUTHOR 2021). Crucially, with metaphor analysis we have an existing theoretical frame which explains the relevance of metaphor for discourse construction and so there is a conceptually sound base for the choice of linguistic features.

As might be expected, there are also challenges associated with this kind of metaphor analysis. In the design stage, as in all corpus-based metaphor analysis, the researcher needs to
decide a) how they will use the corpora to identify the metaphors (for instance, searching by references to the source or target of the conceptual metaphor), and b) what will be included as a metaphor – a decision which is always less clear-cut than it might appear. In the analysis stage, they will need to be cautious in the interpretation of metaphor. First, because it is not the case that the same metaphor will always carry the same evaluation (as shown in Section 7 here) and second because the relevance of a metaphor’s position on the creative-conventionalised (or alive-dead) cline for processing and effect is still under discussion. Work on metaphor from a cognitive perspective leads us to expect a cycle of creativity followed by entrenchment and bleaching, as set out by Croft and Cruse’s life-cycle model:

Once a metaphor takes hold in a speech community and gets repeated sufficiently often, its character changes. First, its meaning becomes circumscribed relative to the freshly coined metaphor, becoming more determinate; second, it begins to be laid down as an item in the mental lexicon; third, it begins a process of semantic drift, which can weaken or obscure its metaphorical origins. [...] As time passes [...] the sense of the expression’s metaphorical nature fades and eventually disappears. Croft and Cruse (2004: 204-205)

However, this has been critiqued as, for instance, in Musolff’s (2014) analysis of the body-politic metaphor where he shows how the well-established metaphor may still result in original lexicalisations and meta-discussion can point towards a collective memory of the metaphor’s historical role.

3. WATER metaphors

WATER metaphors have been extensively discussed in contemporary discourses of immigration in a range of geographical contexts including Austria (el Refaie 2001), Australia (Nguyen and McCallum (2016), Bosnia-Herzegovinia (Mujagic 2018), France (van der Valk 2003), Hungary (Toth et al. 2018), Malaysia (Don & Lee 2014), New Zealand (Salashour 2016), Romania (Neagu &Colipcă-Ciobanu 2014.), Spain (Rubio-Carbonero & Zapata-Barrero 2017), USA (Strom & Alcock 2017), UK (Charteris- Black 2006, Gabrielatos & Baker 2008, KhosraviNik 2009). Abid, Manan, and Rahman (2017) investigate metaphor in the framing of Syrian refugees in four host countries and four non-host countries and find that the dominant metaphor in both sets is WATER. The high frequency and widespread occurrence of the MIGRANTS ARE WATER metaphor across different languages and contexts is especially interesting in terms of pointing towards a shared contemporary perspective on immigration although there is not yet a body of research investigating the presence of the metaphor with
reference to emigration (though see Dervinyte 2009) or in past time periods – both points which this paper addresses.

The existing research has generally found a negative evaluation of migrants embedded within the metaphor. Santa Ana (2002: 76) emphasises that three presuppositions are packed into the water metaphor. First, ‘aggregates of human beings are reduced to or remade into an undifferentiated quantity that is not human. Second, as this mass moves from one contained space to another, some sort of kinetic energy is released. […] Third, such movements are inherently powerful, and if not controlled, they are dangerous’. The reference to control also evokes a closely associated metaphor in migration discourses which is the pervasive container metaphor (Charteris-Black). Nguyen and McCallum (2016: 167) suggest that “[s]uch a conceptual construction potentially legitimises the need to ‘stem’ the water – or, more concretely, to toughen legal and social policies concerning the movement of [maritime asylum seekers]”. They suggest that the high frequency of water metaphors also means the fears and solutions which the metaphor arouses become embedded within mental models of migrants and these, we can hypothesise, would affect the schema activated in future textual encounters (it becomes part of the discourse or evaluative prosody, in fact).

One counter voice is KhosraviNik (2009) who notes the presence of common metaphors of large quantities like “floods of” within topoi of “humanisation and individualisation” and comments that “these discursive characteristics function differently from those of other periods/events and on the contrary, here they create a sense of urgency and appeal for support and help” (2009: 21). Salashour (2016) emphasises the presence of favourable metaphors in economic discussion of immigrants in the New Zealand press as in lexicalisations like a “flow” which has “dried up” or “dwindled”. In these instances, we have an overarching frame of migrants as a natural resource. Nguyen and McCallum (2016: 168) also find water metaphors in news articles they classified as either neutral or advocating for better treatment of asylum seekers but they conclude that “[a]lthough different newspapers show some nuanced differences in metaphor choices, those distinctions are minimal in terms of their potential impacts” (2016:168). This suggests it would be overly-simplistic to assume presence of water metaphors signals a hostile attitude to migration and we need to distinguish between evaluation at a more nuanced level: for instance, a news article can argue for a humanitarian approach but still frame migration as a problem. We should also be clear for whom an evaluation is positive or negative – migrants as victims or resources may indicate a positive attitude from the speaker, but neither is particularly positive for the people
thus described. This is why in a corpus and discourse approach, we have to engage with the texts in context as well as at the abstracted viewing points of collocates.

WATER or LIQUID metaphors have also received extensive attention outside migration discourse. One area of focus has been in emotion talk. Building on Kovecses’s (2003) conceptualisation of a master metaphor EMOTION IS FORCE, which encompasses well-known metaphors such as FIRE, OPPONENT, and NATURAL FORCE, Omori’s (2008) corpus study finds that the dominant source metaphor for emotion (in the structure “x of emotion/s) / feeling/s”) is WATER. In addition, they report that the WATER metaphor is frequently used as a source for emotions such as anxiety, relief, desire, pleasure, sadness and fear (but not, for instance, hope or despair). In these instances, the WATER metaphor is used to signal control and release of control (which may be viewed favourably or not by the speaker). A second area of attention surrounds MONEY IS A LIQUID metaphors (e.g. O’Connor 1998) which has been documented across various languages (e.g. English, Serbian and Romanian in Silaški and Kilyeni 2011). Once again, control is found to be a central feature as speakers want to be able to ensure ‘flow’. Control is proposed as the core of evaluative prosody in Duguid (2011) and AUTHOR (2013). It is also identified as a significant feature in migration discourse in Charteris-Black’s analysis of disaster metaphors because “[f]ear of loss of control and resistance to social change contribute to the centre-right word-view” (2006: 569).

5. Methodology

The corpus used in this study is the c.10.5 billion-word Times Online Corpus which covers the period 1785-2011 and was created at Lancaster University, using optical character recognition files from Gale-Cengage. The size of the Times Online sub-corpora are shown in Figure 1.
It was supplemented by a search-term corpus which was compiled from the Times archive on the Lexis Nexis database (19.3 million words) using the seeds: alien/s, asylum seeker/s, boat people, colonist/s, émigré/s, emigrant/s, evacuee/s, exile/s, expat*/s, immigrant/s, migrant/s, refugee/s, settler/s. These terms were chosen as they cover the historical period and refer to different forms of migration (e.g. movement towards or away from the deictic centre of the speaker, forced vs lifestyle and so on) and had been identified and tested for precision and recall in an earlier phase of the larger research project (see AUTHOR forthcoming).

The broad method used in this study combines corpus linguistics and discourse analysis (e.g. as discussed in Baker 2006; Partington et al. 2013; Mautner 2016) and applies this approach to historical corpora (following McEnery & Baker 2017). Corpus linguistics and the analysis of discourse are described by Sinclair (2004: 11) as “the twin pillars of language research” and they are especially well-matched to analysis of metaphor because they allow for a different viewing points onto the data, thus increasing the ability to capture metaphor in both creative and conventionalised forms. In any metaphor study, the principle challenges are identification and retrieval of metaphor. The Pragglejaz Group (2007) set out a systematic and replicable process of Metaphor Identification Procedure in which an expression may be classified as a metaphor when “(a) its “contextual meaning” contrasts with a “basic meaning” that is more physical and concrete (although not necessarily more frequent), and (b) where the contextual meaning can be understood via comparison with the basic meaning” (Semino et al. 2018: 5). Regarding retrieval, as I was interested in everyday public discourses
of migration, I chose to look for metaphor candidates in collocates of migrant names because their presence would signal a degree of conventionalisation indicating that the metaphor was indeed part of the discourse. Furthermore, for a metaphor to be included, I determined that there should be multiple lexicalisations per category which could signal a degree of metaphorical “animacy”.

In the first stage (reported in Section 6.1), collocates of the node items period (*alien/s, asylum seeker/s, boat people, colonist/s, émigré/s, emigrant/s, evacuee/s, exile/s, expat*/s, immigrant/s, migrant/s, refugee/s, settler/s) were calculated. The relative frequencies of these items are shown in Figure 2. Collocates of node items were calculated per decade (although this is a unit I would describe as necessarily arbitrary - see Marchi 2018 for a critical review of time units). Collocates were identified using a 5 l/r span, minimum frequency of 5 and loglikelihood as the statistical measure. For consistency and repeatability, the same default settings in CQPweb were used in all decades. Loglikelihood was chosen because the aim of the study was to find the banal and non-deliberate (Steen 2008) metaphors and this measure prioritises frequent and non-exclusive items (Brezina 2018: 74). The collocates from the 2010s sub-corpus were calculated using the same measures as those from the Times Online corpus but they are not directly comparable because they were calculated in a search-term corpus rather than a discourse-complete corpus. In earlier stages of research, this was tested on a sample for which I had the discourse-complete corpus and a search-term corpus could be created using the same procedures as those for the ad-hoc corpus, as summarised in Table 1.

<table>
<thead>
<tr>
<th>Statistical measure of collocation</th>
<th>Ratio of number of collocates retrieved</th>
<th>Overlap between search-term corpus collocates and discourse-complete collocates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loglikelihood</td>
<td>1: 3.37</td>
<td>67%</td>
</tr>
<tr>
<td>MI</td>
<td>1: 3.84</td>
<td>72%</td>
</tr>
<tr>
<td>MI3</td>
<td>1: 2.06</td>
<td>19%</td>
</tr>
<tr>
<td>LogRatio</td>
<td>1: 3.77</td>
<td>47%</td>
</tr>
</tbody>
</table>

Table 1. Comparison of collocate retrieval for refugees in 2009 subcorpus

The main effects are that fewer collocates were generated for the search-term corpus and there is not a full overlap in the collocates. The kinds of collocates which were omitted from the collocates generated in the search-term corpus include the terms used to be build the corpus. In contrast, the terms identified in the search-term corpus collocates included items
relating to the corpus metadata and grammatical items. Thus, it would seem that the 2010 collocates are unlikely to produce false positives for metaphor candidates but there is a risk of false negatives that must be considered in the analysis.

Given the size of the corpora, the collocate lists were very long in many cases (e.g. over 2000 per item per decade) and so the process of identifying lexical items which could indicate WATER metaphors was an iterative and triangulated one which involved searching the lists for items identified by consulting contemporary and historical dictionaries and thesauruses, scanning the collocates, reading extended concordance lines, and referring to previous work on WATER metaphors. Once candidate items had been identified, they were manually checked by reading sample concordance lines to see if they were used metaphorically (e.g. checking whether river was mentioned with reference to a real body of water or as a metaphor for framing movement of people).

In the second stage (Section 6.2), the aim was to retrieve contextual information by identifying the company kept by a sample of WATER metaphor lexicalisations (influx, pour, flow, stream, wave, surge). Therefore, these items were moved to node position, and using the same time span and collocate criteria, their collocates were identified and categorised by semantic field (another means of abstracting out above the lexical to aid comparison when language may be expected to diverge). The aim of this stage was to re-contextualise the lexicalisations of the WATER metaphor by showing what kinds of frames the terms evoked at that point in time.

In the third stage (Section 7), I move from collocates to concordance lines. Here, three decades in which there was a peak of frequency in discussion of migration were selected for a more detailed analysis. The three time periods were 1850-1859, 1900-1909 and 2010-2018. The concordance lines were coded according to the speaker’s evaluation of the migrants described: favourable, unfavourable or unclear. The “unclear” category was used when there was any doubt from an extended reading on the basis that it is better to be confident of the polar codings (favourable or unfavourable) than to force all lines into binary categories. It is possible that an additional category of ‘neutral’ could have brought to the fore instances where evaluation was not central which may signal conventionalisation. Examples of all three categories are given below:

Favourable: The climate was delightful, the soil of the valley of the Torrens fertile, and emigrants of capital poured in, burning to commence realising the golden dreams
they had been enjoying during a four months' voyage (1852) Unfavourable: I do not speak of the injustice done to the colony in answering their application for labour by such a polluted stream of emigrants, for the colonists, I perceive, are speaking for themselves (1852) Unclear: One of the Galway papers, noticing the continuance of the exodus from that province, speaks of a counter tide of returning emigrants, persons who have amassed some wealth or who have fallen into a slate of ill-health (1856)

6. Conventionalisation of the MIGRANTS ARE WATER metaphor

6.1 The history of MIGRANTS ARE WATER

As Table 2 shows, the results indicate that WATER metaphors have a long-standing history of use in relation to discussion of migration and migrants.

<table>
<thead>
<tr>
<th>Decade</th>
<th>Water collocates</th>
<th>No. of collocate types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1800</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>1810</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>1820</td>
<td>river</td>
<td>1</td>
</tr>
<tr>
<td>1830</td>
<td>influx, river</td>
<td>2</td>
</tr>
<tr>
<td>1840</td>
<td>river, influx</td>
<td>2</td>
</tr>
<tr>
<td>1850</td>
<td>influx, stream</td>
<td>2</td>
</tr>
<tr>
<td>1860</td>
<td>influx, river</td>
<td>2</td>
</tr>
<tr>
<td>1870</td>
<td>influx, pouring, stream</td>
<td>3</td>
</tr>
<tr>
<td>1880</td>
<td>flow, influx, river, stream</td>
<td>4</td>
</tr>
<tr>
<td>1890</td>
<td>influx, stream</td>
<td>2</td>
</tr>
<tr>
<td>1900</td>
<td>flow, influx, pouring, stream</td>
<td>4</td>
</tr>
<tr>
<td>1910</td>
<td>absorb, flood, influx, pouring, stream</td>
<td>5</td>
</tr>
<tr>
<td>1920</td>
<td>absorb, absorbed, absorbing, flood, floods, flow, influx, inflow, pouring, river, stream, tide</td>
<td>12</td>
</tr>
<tr>
<td>1930</td>
<td>absorb, absorbing absorbed, flood, flooded, floods, flow, influx, pouring, rising, river, stream, streaming</td>
<td>13</td>
</tr>
<tr>
<td>1940</td>
<td>absorb, flood, flooded, flow, influex, pouring, stream, streamed, streaming, tide</td>
<td>10</td>
</tr>
<tr>
<td>1950</td>
<td>absorb, absorbed, absorbing, flood, flooded, floods, flow, inflow, influx, pour, pouring, river, stream, streaming</td>
<td>15</td>
</tr>
<tr>
<td>1960</td>
<td>absorb, absorbed, absorbing, flood, flooded, floods, flow, inflow, influx, pour, pouring, rising, river, stream, streaming, tide, trickle, wave, waves</td>
<td>20</td>
</tr>
<tr>
<td>1970</td>
<td>absorb, absorbed, absorbing, current, flood, flooded, flooding, floods, flow, inflow, influx, pour, pouring, stream, streamed, streaming, tide, trickle, wave, waves</td>
<td>20</td>
</tr>
<tr>
<td>1980</td>
<td>absorb, absorbing, flood, flooded, flooding, floods, flow, inflow, influx, pour, pouring, stream, streamed, streaming, tide, trickle, wave, waves</td>
<td>19</td>
</tr>
</tbody>
</table>
Although the 1850s is the first decade in which we see multiple water collocates with a single node (emigrant* collocating with influx and stream), Table 2 shows the collocates influx and river occur separately before that point. These should not be mistaken as marking the birth of this metaphor though - if we search manually through the concordance lines, we can find occurrences of the WATER metaphor from earlier periods, as illustrated in (1).

(1) Nice abounds with them, and the principal towns of Switzerland overflow with refugees from the neighbouring kingdom (Times 1790)

Table 2 shows that the relationship between names for people who move and WATER metaphors is strong throughout the period and it provides evidence that MIGRANTS ARE WATER continues to be a creative metaphor. While we might expect the conventionalisation process to have narrowed down to a few fossilized instances, the number of lexicalisations increases over time (although this will, in part, be influenced by the increasing size of the corpora, as seen in Figure 1, which means more collocates can be identified). There is a remarkable degree of stasis in the picture of the relationship shown in Table 2. First, there were no WATER collocates from the nineteenth century which do not occur in the present-day data. Second, there are very few collocates from the post 2000 sub-corpora which do not occur in the past corpora. Some exceptions to this are wave (first occurs as a collocate in the 1960s subcorpus), current (1970s) and surge (2000s).

If we represent the same data in a modified format (Table 3), we can also see that some names for migrants appear to attract WATER metaphors more than others.

### Table 2. Water collocates of migrant-related nodes

<table>
<thead>
<tr>
<th></th>
<th>1990</th>
<th>2000</th>
<th>2010-2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>absorb, absorbing, current, flood, flooded, flooding, floods, flow, inflow, influx, pour, pouring, rising, river, stream, streamed, streaming, streaming, swelled, tide, trickle, wave, waves</td>
<td>23</td>
<td>20</td>
<td>23</td>
</tr>
</tbody>
</table>

[Table 2 near here]

---|---|---|---|
migrants | absorb flow | flow flood influx wave | current flood flooded flooding flow flows inflow influx pouring rising river stream streaming surge tide wave waves |
<table>
<thead>
<tr>
<th>settlers</th>
<th>influx river stream</th>
<th>absorb absorbing flow influx influx pouring stream tide</th>
<th>influx</th>
<th>influx river wave</th>
</tr>
</thead>
<tbody>
<tr>
<td>colonists</td>
<td>influx</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>boatpeople</td>
<td>influx pouring</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>refugees</td>
<td>influx pouring</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>immigrants</td>
<td>flow influx stream</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>exiles</td>
<td>flow influx stream</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>expats</td>
<td>flow influx stream</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>evacuees</td>
<td>influx</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>aliens</td>
<td>influx</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>asylum seekers</td>
<td>influx</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
<tr>
<td>emigres</td>
<td>Stream</td>
<td>absorb absorbed flood flooded floods flow influx pour</td>
<td>influx</td>
<td>wave</td>
</tr>
</tbody>
</table>

Table 3. Collocates displayed by node and broad time period

Significantly, this attraction between migrant-naming choices and water metaphors cannot be explained away by greater/lesser frequency of the node items alone. When Table 3 is viewed in conjunction with Figure 2, we see there are names that are highly frequent in certain past periods which do not seem to collocate with water metaphors: such as emigrants and colonists which are the highest peaks in the 1850s, and aliens which is the highest peak in the 1920s. We may, therefore, need to be cautious about making claims for an association with a group rather than specific names for that group or particular categories which are constructed within that group. This is complicated by the fact that relationships between names and concepts are not stable across the time period (e.g. emigrant is not consistently used to refer to migration away from the deictic centre of the speaker).
6.2 Conventionalisation of WATER metaphors

Having established that the MIGRANTS ARE WATER is a long-standing metaphor, the next step is to understand what that means in the cross-cultural context engendered by a long-distance diachronic study. We cannot assume that any given term will have the same denotational and connotational meaning across a period of over 200 years. Nor can we assume that a metaphor positions the reader in the same way across such an extended period. To address this lack of cultural knowledge, we can exploit the corpus in a complementary way by moving the collocates into node position. For instance, if we see that flow is a collocate of emigrant in a given decade, in order to understand how flow framed emigrants, we need to know the contemporary associations of flow i.e. what else did it collocate with, outside migration discourse. By examining these second-order collocates, we can see what associations they had at different points in the past and therefore reach an understanding of the WATER
metaphors of migration in the context in which they were produced. It can also enable us to get an empirical handle on the ways in which these metaphors are conventionalised.

Influx is consistently one of the most cited realisations of the Water metaphor in discussion of migration and, as shown in Table 2, occurs as a collocate of migrant names in every decade from 1830 onwards. It is also one of the most problematic because it is so highly conventionalised today. Appendix 1 shows the lexical items contained in the 25 most significant collocates of influx across the extended period. In terms of ‘liveness’ of this metaphor, it rarely co-occurs with lexical items from the semantic fields of water or other liquids at any point in time. Thus, we do not have a progressive bleaching of metaphorical meaning. However, it is highly conventionalised for discussing people whose name indicates movement (e.g. strangers) and other categories of people (e.g. paupers). The second main semantic field is finance and this pattern exists throughout the period occurring from 1820 onwards. This provides a way into interpreting use of influx in migration discourse as a mapping of the movement of people onto an existing frame of movement of economic resources. In terms of description, collocates referring to scale consistently focus on large size. The findings for manner are mixed with both constant and sudden recurring in the nineteenth century, while the twentieth and early twenty-first centuries associate it with sudden movement. This more recent pattern is noted in Nguyen and McCallum’s (2016: 166) study of migration discourse: “[i]nflux is particularly threatening because it moves quickly, without warning, and often in great volume”. In terms of interpreting the effects of influx to describe migration movements we have magnitude and suddenness as core features. Finally, the recurrence of causality markers in the collocates (e.g. caused, owing, result) of influx in the earlier period indicate it is newsworthy because of effects. Furthermore, we see indicators of negativity through references to (lack of) control throughout the period (prevent, control, cope).

Pouring appears to maintain “liveness” as it collocates with lexical items relating to water and other liquids throughout the period (see Appendix 2). It also shows a limited but long-standing conventionalisation for description of people (troops appears as a collocate from the 1810s onwards) and a subset of people who move (refugees appears from 1900 onwards). The other main semantic field which co-occurs throughout is that of fire (flames, and more persistently, smoke). Money emerges as a consistent collocate from 1920 onwards (joined by cash in 2000), thus marking an emerging metaphorical use of describing finance which shows
similarity to influx. The movement of pouring is described as continuous in the earlier period and not noted in the later period.

The collocates of flow (Appendix 3) indicate this term is still “live” as it co-occurs with lexical items indicating water and other liquids throughout the time period. The conventionalisation for metaphoric use is most evident in a semantic preference for the lexical items associated with finance (e.g. capital, funds, gold), and this association becomes stronger over time. It is first indicated in 1830 (capital) and gains multiple collocates in 1910. Interestingly, the pattern weakens as we come closer to the present day with cash being the sole financial collocate for the most recent decades. Other items that are described with flow can also point towards the evaluative prosody. In the 1940s and 1950s it is heavily associated with movement of supplies (good, materials, orders), in more recent times it is traffic, information and problems. Like influx, when applied in migration discourses it offers the potential of framing migrant movement as economic goods. The nature of the flow is consistently described in terms of continuity, unlike influx which had that dual representation of suddenness and constancy. There is no consistent pattern of it being used to refer to people and the only migration-related collocate is emigration in 1890. There is little here to indicate a prosody of threat.1

The collocates of stream (Appendix 4) show that, like pouring and flow, it retains a literal use (e.g. water, tide and trout) but this becomes weaker over time. In the two most recent decades, we see items relating to a semantic field of finance, (revenue, income) echoing the patterns noted above but this not well established. Where collocates point towards the nature of the stream, the main feature is constancy (steady, constant) and this feature is established from mid-nineteenth century onwards (prior to that there is more variation e.g. powerful). There are few references to people and these are largely in the proper name Cold-stream Guards (a British Army regiment) though we also note refugees in the 1950s subcorpus.

The picture for wave (Appendix 5) is rather mixed because of the polysemous meanings (e.g. as a hand movement, radio waves etc.). However, three clear patterns can be discerned. First, there are collocates relating to water throughout the period (most strongly, tidal, but also ocean and sea) thus indicating a “liveness” for metaphoric use. The presence of items like swept and crest also suggest that if it is used metaphorically, then the metaphor is extended rather than a fossilised remnant. Second, there are no consistent references to people, or

---

1 It has been observed that flow implies a one-way direction and excludes bi-directionality (Charteris-Black 2006; Nguyen and McCallum 2016), but this is not observable at the collocational level.
migration, Third, from 1920s onwards, where it is use metaphorically there is a semantic preference for crime (e.g. crime, attacks, arrest, violence). The conventionalisation for crime clearly has important implications for the framing of migration when it is used in migration discourses.

The collocates of surge (Appendix 6) show that there is a consistent collocation with water although this is not richly diversified (waves, sea, tidal) and in the earlier period it collocates with nautical items (freight, stewardess, bow) indicating non-figurative use. The strongest pattern is a semantic preference from 1950s onwards for finance and this is richly lexicalised (buying, demands, profits, and so on), mirroring findings for the other water items. Where surge is modified (from 1930s onwards), the emphasis is on news values (Bednarek & Caple 2017) such as strength (great, mighty), recency (recent, latest) and unexpectedness (sudden, late).

<table>
<thead>
<tr>
<th>Used across period</th>
<th>Used in later periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>influx</td>
<td>pouring</td>
</tr>
<tr>
<td>PEOPLE WHO MOVE</td>
<td>LIQUID</td>
</tr>
<tr>
<td>PEOPLE</td>
<td>MONEY</td>
</tr>
<tr>
<td>MONEY</td>
<td>refugees [later period]</td>
</tr>
<tr>
<td></td>
<td>troops [earlier period]</td>
</tr>
</tbody>
</table>

Table 4. Summary of semantic fields in collocates

These semantic preferences of the water collocates are summarised in Table 4. In short, pouring, flow, stream, wave and surge all retain literal usage according to the collocates indicating liquids. The exception is influx which has a weak collocational relationship with literal items throughout. In metaphorical terms, influx, pouring, flow, stream and surge all
exhibit a preference for the semantic field of finance which suggests that when people who move are framed metaphorically using these WATER metaphors they are being placed into an existing framing relating to economic worth. In contrast, wave, which does not co-occur with finance-related terms, shows a conventionalised semantic preference for crime lexis. Therefore, here we can anticipate a negatively-charged framing of migrants being established when then they are represented using this lexicalisation and it is worth noting that wave becomes a collocate of migrant names (shown in Table 3) around the same time period that it develops the crime semantic preference. Finally, only influx and pouring seem to have a conventionalised relationship with lexical items referring to people. In the case of influx, and to a limited extent, pouring, this is more specifically people who move. Where these two differ significantly is in the drama associated with the movement (sudden vs gradual).

When we talk about conventionalisation of metaphors we need to be attentive to what kind of relationship is conventionalised. In the case of the WATER collocates, we can identify two clines. First, we can ask whether the items are conventionalised for framing migration. Second, we can ask whether these terms are conventionalised for metaphorical use more generally. These relationships are visually depicted in Figure 3. Knowledge of these relationships can tell us how to evaluate and interpret the different lexical realisations of a single metaphor.

Figure 3. Illustration of conventionalisation patterns in water collocates
What this section has shown is how the corpus can be a tool for re-contextualising the metaphors with their contemporary framings; the other collocates of *wave* etc. are part of the meaning of ‘*wave of [migrant group]*’ and so on.

### 7. Evaluation of Migrants Are Water

#### 7.1 Overview of evaluation in peaks/troughs

To gain insight into the evaluation performed by the *Water* metaphor, the third stage involved returning to the concordance lines in which migrant names are the node. Concordance lines from periods of peak discussion of (1850-1859, 1900-1909 and 2010-2018, as described in Section 5) were coded for evaluation by the speaker. The findings are summarised in Figure 4. As can be seen, there is a clear trend in shift of evaluation with more favourable evaluations co-occurring with Migrants Are Water metaphors in the earlier periods compared to a much greater proportion of unfavourable evaluations in the most recent time. The reasons for this are not straightforward and, like the newsworthiness of migrants in the corpus, do not correlate simplistically with migration flows.

![Figure 4](image_url)

**Figure 4. Summary of evaluation in concordance lines of Migrants Are Water**

We can dig deeper to test whether evaluations are tied to either names for migrants (e.g. *immigrant vs emigrant*) or the choice of lexicalisation of the metaphor (e.g. *wave vs stream*).
Naming choices were classified as more favourable when there were at least 50% more favourable occurrences than unfavourable and when there were at least five occurrences for each evaluation. The same was applied to classification as unfavourable.

Regarding choice of name for migrants, as Table 5 illustrates, it is not the case that certain naming choices correlate with favourable evaluations and others with unfavourable evaluations over time. The strongest pattern is one of favourable evaluation in earlier periods and unfavourable in later periods. It is perhaps worth noting that even *refugees*, one of the more sympathetic terms for inward migrants in the current day (AUTHOR 2014), occurs in more unfavourable contexts.

<table>
<thead>
<tr>
<th>Date</th>
<th>More favourable</th>
<th>More unfavourable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1850-1859</td>
<td>emigrants immigrants</td>
<td>aliens refugees</td>
</tr>
<tr>
<td>1900-1909</td>
<td>colonists emigrants migrants settlers</td>
<td>flood influx swamp</td>
</tr>
<tr>
<td>2010-2018</td>
<td>immigrants refugees asylum seekers migrants</td>
<td>flood flow influx pour tide wave</td>
</tr>
</tbody>
</table>

Table 5: Summary of evaluation correlating with naming choice

Table 6 provides a similar finding insofar as to the principal factor on the balance of evaluation is the time period. The terms *influx* and *flow* appear with more favourable evaluations in the 1850s and 1910s respectively and more unfavourable evaluations in 1900-1909 and 2010-2018. In contrast, the terms *absorb* and *stream* only appear in more favourable evaluations in the earlier periods while *flood, swamp, pour, tide* and *wave* only appear in the later period where they have a more unfavourable setting.

<table>
<thead>
<tr>
<th>Date</th>
<th>More favourable</th>
<th>More unfavourable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1850-1859</td>
<td>absorb influx</td>
<td></td>
</tr>
<tr>
<td></td>
<td>stream</td>
<td>flood influx</td>
</tr>
<tr>
<td>1900-1909</td>
<td>absorb flow</td>
<td></td>
</tr>
<tr>
<td></td>
<td>stream</td>
<td></td>
</tr>
<tr>
<td>2010-2018</td>
<td>flood flow</td>
<td></td>
</tr>
<tr>
<td></td>
<td>influx pour</td>
<td></td>
</tr>
<tr>
<td></td>
<td>tide wave</td>
<td></td>
</tr>
</tbody>
</table>
The key point to make here is that if we base our assumptions about the evaluative potential of a given metaphor on our current-day contextual knowledge, we risk gravely mis-reading data from the past. The WATER metaphor is a long-standing frame for describing people who move, but it is not consistent in the evaluative polarity.

7.2 Frames underpinning evaluation

Using the concordance lines from the three different time periods, we can then start to identify common frames which correlate with the favourable/unfavourable evaluations. One way in which migrants are evaluated favourably in WATER metaphors is with an economic resource frame. This fits the documented use of WATER as a source metaphor for financial issues (in both previous research and the collocational analysis in Section 7.1) and shows how the different approaches to metaphor may triangulate. In the first economic resource frame, migrants provide labour, as illustrated in (2) and (3) and this is attested in all three sample periods.

(2) A considerable addition having been made to the population of Geelong by the influx of Immigrants, the labour market had been well supplied both with female domestics and almost every other description of servant, as well as mechanics (1855)

(3) Plans for special visas for London are being prepared by Sadiq Khan as he tries to preserve a steady flow of migrants into the capital after Brexit. The mayor of London is considering plans that he said would “ensure we can carry on recruiting and attracting talent (2016)

The second economic resource frame sees the migrants as bringing capital, illustrated in (4) and (5) and this was attested in all three periods.

(4) The future of this Colony, so richly endowed by nature, depends upon a steady influx of enterprising colonists with brains and capital to develop its immense mineral, forest, and agricultural resources (1909)
(5) For a small country it's punching far above its weight and the ambitions of the government are clearly paying off; the hotels are full, the marinas bustling and there's an influx of wealthy expats that is unmatched anywhere else in the Balkans (2018)

In the two more recent periods, there is a thin pattern of favourable evaluation occurring in a frame where migrants are presented as a cultural and racial resource as attested in (6) and (7).

(6) Nothing but a steady influx of English emigrants will counter-balance the numerical superiority of non-English inhabitants of South Africa-a population the annual increase of which is abnormally large (1903)

(7) Cheese has been a staple of the American diet for centuries. Cheesemaking was brought to the new colonies in the 17th century by Puritan dairy farmers and popularised across the expanding nation by successive waves of European immigrants who brought with them the cheesemaking traditions of Germany, Switzerland, Italy and Scandinavia (2018)

The unfavourable evaluation frames partially mirror those above. In the first unfavourable frame, the metaphor presents migrants as an economic threat and this frame was found in all three time periods sampled. Example (8) sees outward migration as a threat to the ‘home’ country and (9) presents inward migration as the threat.

(8) From every station on every line of railway-from every packet on whatsoever canal-the multitude of emigrants filling the cars and cabins pour into the seaports, to fill the emigrant ships whose destination is the new world. Where will all this end? The labourers are deserting the fields […] (1853)

(9) The discoveries of science, diminishing the demand for manual labour on the one hand, and that monstrous flood of alien immigrants on the other, will quickly undermine the British worker's economic position and recreate a surplus in the labour market (1903)

There is no clear unfavourable parallel to the second favourable theme (migrants are capital) in the MIGRANTS ARE WATER metaphors – this is more likely to be realised in MIGRANT ARE A WEIGHT metaphors. The third favourable frame is echoed in the second unfavourable frame; migrants are presented as a cultural threat to race, ethnicity, religion or class. This echoes the findings of wave as a source for describing crime which constitutes a cultural attack. This was found in all three time periods, as shown in examples (10) to 12).
There is a noticeable difference between the class of emigrants who go out to Australia and those who are pouring in a never-ceasing stream to America. While the former have a somewhat respectable and substantial appearance, the latter are composed mainly of Irish and German poor, and have squalor and filth for their abiding companions (1854).

If the influx of coloured aliens into a land but sparsely inhabited with whites continues unchecked, Queensland will, in a generation or two, be cursed with a population of half-castes, a misfortune which none greater can overtake a young nation (1901).

Europe's identity is fundamentally threatened by the mass influx of migrants, most of them Muslims, who will live together in parallel communities that won’t integrate (attribution, 2015).

In the third unfavourable frame, there is no explication of the threat posed by migrants. They are unfavourably evaluated simply in terms of excess number. It is assumed that the shared reader/writer knowledge can allow the presuppositions embedded here to be sufficient to communicate meaning. This was found in all three time periods and is illustrated in (13) and (14).

Victoria is still receiving from the Old World more immigrants than, in the present condition of our land system, we can readily absorb (1857).

Government needs help to stem the flood of migrants (2010).

This final stage of qualitative analysis confirms the importance of control in evaluation. The same metaphor of MIGRANTS ARE WATER may be used with the same discourse frame (e.g. in relation to the economy) but with completely differing evaluations depending on the speaker’s view of who controls that resource.

8. Conclusions

In conclusion, the research presented here has shown that MIGRANTS ARE WATER is a long-standing metaphor. What makes it especially interesting is that it is simultaneously conventionalised and creative as a metaphor. The lexical realisations of the metaphor are
relatively stable throughout the period and some elements are highly conventionalized for both non-literal use and for describing migrants. In the case of *influx* this occurs to an extent throughout the period that it is difficult to justify a metaphorical interpretation. The conventionalised nature of the metaphor, and the implications for interpretation, are briefly addressed in Nguyen and McCallum (2016) who conclude their paper stating:

> we acknowledge that most of the metaphors we have discussed are ‘dying metaphors’, simply clichés that have lost their real imaginative power and are used by journalists to spare the effort of inventing phrases for themselves. They resort to them out of habit of using sensational language and catchy phrases. Nevertheless, what one has to bear in mind is that such language still affects the perception of the migration phenomenon discussed in the press and certainly does not help to build a positive picture of it. (Nguyen and McCallum 2016: 220)

And yet, overall, the metaphor of *MIGRANTS ARE WATER* continues to be elaborated, and draws on lexical items which are still used in literal contexts - both factors indicating ‘liveness’. For some users, the metaphors are well-trodden and may well be used unthinkingly (non-deliberate). However, for others, they express a clear world view and are used with the aim of communicating that point. It would be unhelpful to conflate these (or assume there is no middle ground), but the fact the metaphor exists at both ends of the conventionalised cline means it will speak to different addressees differently. Future research in this area could examine whether the conventionalised and novel lexicalisations of *MIGRANTS ARE WATER* elicit differing or similar emotional responses.

In terms of evaluation, the dominance of negative evaluations is relatively recent and the proximity of the *MONEY IS A LIQUID* and *MIGRANTS ARE WATER* metaphors shape the framing of people who move. However, as anticipated by Duguid (2011), the evaluative prosody of the metaphor is underpinned by the speaker’s perception of control. The most significant takeaway from this is that stasis in source-target mapping cannot be assumed to reflect stasis in either framing or evaluation. Future work could also explore more closely at the relationship between *EMOTION IS FORCE* metaphors and the ubiquity of the *MIGRANTS ARE WATER* metaphor. If we look for an explanation of the longevity of this metaphor, we could see it in an association with or realisation of the news value of ‘superlativeness’ (Bednarek & Caple 2017). The examination of the water collocates showed, *influx, wave* and *surge* all collocate increasingly with modifiers indicating magnitude and suddenness. In the move to
more recent time periods, we also see that ‘negativity’ could be a news value playing a role in maintaining the metaphor. The mixed evaluation in the past may also offer a way into understanding how a metaphor which today has a strongly negative prosody may be both deniable regarding intent and unwittingly used.

With reference to the role of metaphor in long-distance diachronic study, I hope to have shown how metaphor can become a tool to both analyse and contextualise. The corpus linguistics tools can provide a) information about saliency of metaphors at different points in time, b) contextual information about what associations a given lexicalisation of a metaphor has in the past and c) identification of sampling periods for more detailed metaphor analysis. At no point, of course, is the researcher taken out of the equation in this process. It is not about automating any process but about providing replicable ways into the data and reflexivity and transparency about those processes is central. The new difficulty that this highly-productive process raises is how to report and present findings spanning multiple time periods. As Hunston (2017) reminded us, ‘corpus investigation software rearranges the data facilitating the observation of latent patterns’ and we need to find creative ways of doing the same kinds of ‘rearranging’ in presenting findings – and it may be that we need more interactive forms of ‘papers’ to enable the richness of ‘long-distance’ diachronic studies to be fully communicated.

**Acknowledgements**

I would like to acknowledge the help and generosity of the Centre for Corpus Approaches to Social Sciences at Lancaster University for access to the Times Online Corpus and Dario del Fante for help in compiling the contemporary Times corpus.

**References**


