

Implicit dialogical premises, explanation as argument: A corpus-based reconstruction

KIERAN O'HALLORAN

*Senior Lecturer in Linguistics
Centre for Language and Communication
Faculty of Education and Language Studies
Open University
Milton Keynes
MK7 6AA
UK*

k.a.ohalloran@open.ac.uk

Abstract: This paper focuses on an explanation in a newspaper article: why new European Union citizens will come to the UK from Eastern Europe (e.g., because of available jobs). Using a corpus-based method of analysis, I show how regular target readers have been positioned to generate premises in dialogue with the explanation propositions, and thus into an understanding of the explanation *as an argument*, one which contains a biased conclusion not apparent in the text. Employing this method, and in particular 'corpus comparative statistical keywords', I show how two issues can be freshly looked at: implicit premise recovery; the argument/explanation distinction.

Résumé: J'emploie un logiciel qui analyse quantitativement un article (dont le but est d'expliquer pourquoi des citoyens de la nouvelle Europe vont quitter l'Europe de l'Est pour aller habiter au Royaume-Uni) de journal pour démontrer que son auditoire cible régulier est influencé à produire des prémisses dans un dialogue contenant des explications, et ainsi à interpréter des explications comme si elles étaient des arguments. Cette approche apporte une nouvelle optique sur l'identification des prémisses implicites et sur la distinction entre des arguments et des explications.

Keywords: Argument, corpus-based analysis, corpus comparative statistical keyword, cultural keyword, explanation; implicit dialogical premise recovery; plausible fuzzy reasoning

1. Introduction

1.1 Orientation

This article focuses on an explanation in a newspaper article, from the popular tabloid United Kingdom (UK) newspaper, *The Sun*, which relates to the European Union (EU) enlargement on May 1st 2004. On this date, the EU expanded to include 10 new countries,

eight of which are located in Eastern Europe.¹ The piece of text I focus on explains, via a few propositions, why new immigrants from the accession countries would want to come to the UK. From a linguistic analysis of previous, related newspaper stories, I show the following:

- how regular target readers have, in fact, been positioned over time into an understanding of the explanation *as an argument* with a biased conclusion. This conclusion includes the understanding that large numbers of Eastern European migrants will place strain on UK social services, which is, in part, due to government incompetence. This understanding is not apparent in the textual explanation, but could well be generated in reading by regular target readers.
- how, in understanding the explanation as an argument, regular target readers are likely to generate implicit premises in a dialogical manner—i.e., in dialogue with the explanation.
- how representation via an explanation rather than an argument is a state of affairs which helps the UK newspaper in question to evade charges of bias—since the bias is not in the text as such, but would be in the mind of the reader.

To reveal the argument, I draw on corpus-based text analysis, and in particular the concept of corpus comparative statistical keywords, which I explain below. Using this method, I highlight how two perennial issues in informal logic study can be freshly looked at: implicit premise recovery and the argument/explanation distinction.

1.2 Two perennial issues in informal logic

1.2.1 Recovering implicit premises

Informal arguments frequently are elliptical, since they usually leave out premises that according to formal logic are necessary for valid conclusions (Anderson *et al.*, 1997; Govier, 1987). They are thus ‘enthymemes’ (in Aristotelian language) or ‘incomplete’ (Walton, 2006)². On the issue of implicit premise recovery, there

¹ On May 1st 2004 the EU expanded from 15 to 25 countries. On January 1st 2007, the EU expanded to 27 countries when both Bulgaria and Romania joined. On the history of the EU, see: <http://www.europa.eu>

² Walton and Reed (2005) have a defeasibilistic view of enthymemes (‘A defeasible generalization, in contrast to an absolute universal generalization, is one that is subject to exceptions and that is defeated (defaults) in a case where one of the exceptions occurs (Walton and Reed,

has been interesting work recently. For example, Katsav and Reed (2008) focus on the modelling of implicit premises for computational analysis of arguments. Walton and Reed (2005) show how argumentation schemes (e.g., Walton, 1996) based on defeasible generalisations can be useful in recovery of premises. Furthermore, Araucaria software (Reed and Rowe, 2001), ‘given the user’s markup of an argument in a given text of discourse’, can assist in determining implicit premises (Walton and Reed, 2005: 340).

The cited works are fascinating approaches to implicit premises revelation. They are based, though, on analysis of arguments outside of the reading conditions in which they might be processed. Once an analyst tries to take account of how a text might be processed by a target readership, this creates a problem: to what extent is the recovery of implicit premises a reflex of the perspective/background knowledge of the analyst? The analyst runs the risk of: (i) over-interpretation (O’Halloran, 2003; O’Halloran and Coffin, 2004; Coffin and O’Halloran, 2008), i.e., the implicit premise recovery might say more about the analyst’s perspective than the target readership’s; (ii) arbitrary recovery of implicit premises which leads to circularity; in other words the analyst arbitrarily focuses on a part of an argument where s/he supposes there is an implicit premise, then goes on to interpret the content of the implicit premise which, in turn, reinforces the original analysis that there was an implicit premise in the first place.

1.2.2 Argument versus explanation

It has been noted often that there can be difficulty in distinguishing argument from explanation. One problem is because explanations and arguments share indicator words such as ‘so’ and ‘then’, and because ‘both explanations and arguments consist of groups of statements where some of the statements are taken as starting

2005: 367))’. This would seem to be at odds with the Aristotelian view on enthymemes where implicit premises are recovered deductively. However, Walton and Reed (2005: 339-340) argue:

‘[Our] defeasibilistic view of enthymemes is not as new as it may sound to many readers. It can be shown to be very close to what may have been Aristotle’s original doctrine of the enthymeme, according to the view of some commentators (Burnyeat, 1994). According to this original view an enthymeme is not an argument with a missing premise, but it is a plausible argument based on a defeasible generalization, as opposed to a deductive argument based on a universal generalization of the type represented by the universal quantifier of deductive logic.’

points and lead to others as end points' (Walton, 2006: 76). Despite these difficulties, several theorists have contended the distinction between argument and explanation is an important one to retain (e.g., Govier, 1987; Walton, 2006; Schwarz and Asterhan, forthcoming). An important reason why can be found in Walton's (2006: 76-77) illuminating and useful distinction:

...the purpose of an argument is to give a reason to support a claim made by one party in a dialogue. The claim is something that is doubted by the respondent in the dialogue. It is a proposition that is at issue or is unsettled. An argument is supposed to present a good reason for the respondent to come to accept this proposition as true, thus removing the doubt...

...explanations and arguments are different because each has a different purpose in a dialogue. The goal of an explanation is not to convince or persuade the party that a particular proposition is true [i.e., unlike an argument] but to express the queried proposition in some more familiar terms or relate it to another set of propositions that can be put together so that it is more familiar or comprehensible to him.

I concur with this distinction. However, in practice as this paper shows, it is possible that a reader might blur the distinction between argument and explanation, not because of obtuseness or ignorance, but because they have been primed to read an argument *into* an explanation from regular exposure to a particular set of arguments.

1.3 Corpus methods of analysis

To show how target readers are potentially positioned into reading an argument into an explanation in the May 1st *Sun* text, I draw on the following: a database, or corpus, of texts on related topics from *The Sun* in the six weeks preceding May 1st. A corpus is a large electronic database of texts in a particular language, e.g., The Bank of English is a corpus of 450 million words; The British National Corpus consists of 100 million words of English. In the last 10 to 20 years, technology has developed to allow quick searches of corpora using software known as concordancers. Many advantages can be gained through employing large corpora. For instance, instead of relying on intuitions of how language is used, or on small amounts of data, one can now make empirically robust statements about widespread language use.

I investigate *The Sun* corpus with quantitative analysis software, and analyse how regular target readers would have been exposed to

repeated patterns of argument over the six week period. An important focus in my analysis of this *Sun* corpus is *corpus comparative statistical keywords*, which are explained in Section 3. It is on their basis that I generate premises which are potentially implicit in the processing of target readers of the May 1st text, and thus show the following: not only that the explanation can be read as an argument, but that arbitrariness of implicit premise generation by the human analyst is constrained because the software generates such keywords.

Corpora are more and more being drawn on in argumentation study (e.g., Degano, 2007; Reed, 2006; Stubbs, 1996, 2001; Zagar, 2007; Coffin and O'Halloran, 2008, 2009) and for a variety of different purposes. Michael Stubbs is probably the key pioneer of corpus-based text analysis. He was also one of the first scholars to use corpus-based approaches in relation to argumentation analysis (Stubbs, 1996; 2001) (though this is not a sustained focus). Stubbs' approach has been criticised by the argumentation theorists, Eddo Rigotti and Andrea Rocci (Rigotti and Rocci, 2005). In relation to my corpus-based argumentation analysis, this article takes cues from Stubbs' corpus-based work (see Section 3) and its unfolding is affected by engagement with Rigotti and Rocci's (2005) critique of Stubbs's work (see Section 4).

2. May 1st text 2004 in *The Sun*

2.1 *The explanation*

On May 1st 2004, 10 new countries joined the EU: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia (all Eastern European countries) as well as Cyprus and Malta. On that day, *The Sun* reported this event in the following text.

- a. Migrants are the first of many
By NICK PARKER CHARLES RAE and CHARLES YATES
- b. HISTORY will be made today as new EU citizens begin leaving their homelands for new lives in Britain.
- c. They will come in their hundreds by coach, train and plane, desperate for decently-paid work—or any job at all.
- d. They are citizens of ten new member states of the EU who now have the right to live and work in Britain.
- e. Those states are the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia, Cyprus and Malta.
- f. *Eight are poverty-stricken former Soviet states in Eastern Europe—with a total population of 75 million.*
- g. *Our government says there are half a million jobs waiting here to be filled.*

- h. *And if the new arrivals find legitimate work in Britain, paying tax and national insurance, they will qualify for a range of benefits.*
- i. *These will include free health care, child tax credit, child benefit, working tax credit and housing benefit.*
- j. *Other EU countries like France and Germany have ruled out free access to jobs for up to seven years.*
- k. *Small wonder that eager men and women used to working for a pittance in countries with huge unemployment are snapping up tickets to the UK.*
- l. People out to get a head start in the exodus crammed into coaches yesterday at stations in Warsaw, Poland, the Czech capital Prague and the Slovak capital Bratislava.
- m. Seventy-one LATVIANS smiled as they boarded a double-decker bus in the capital Riga for a 24-hour journey west.
- n. Coach firm Eco-Lines said they had put on an extra service to London this week for the £75 trip.
- o. All services are booked for the next week.
- p. ESTONIANS are travelling from their capital Tallinn, via Helsinki in Finland.
- q. Budget airlines like Sky Europe, which has services from Slovakia to Britain, have reportedly sold out all flights.
- r. POLAND'S Air Polonia alone has bookings for about 4,000 migrant workers this month.
- s. Warsaw-based coach operator Orbis expects migrants to fill 1,500 of its monthly 4,500 seats to Dover. (source: <http://www.thesun.co.uk> 1st May 2004) (my italics)

Consider what I have italicised above, which I reproduce propositionally below:

- Proposition 1. Eight countries are poverty-stricken former Soviet states in Eastern Europe—with a total population of 75 million. (sentence f)*
- Proposition 2. Government says that there are 500,000 jobs inUK. (sentence g)*
- Proposition 3. There are social benefits for migrants if they find legitimate work in UK. (sentences h and i)*
- Proposition 4. Two other major economies have ruled out free access to jobs. (sentence j)*
- Proposition 5. Eastern Europeans do not earn much. (sentence k)*
- Proposition 6. There is much unemployment in Eastern Europe (sentence k)*

Explanation summary:

All this explains why people from Eastern European countries want to come to UK. (sentence k)

I regard this as an explanation—an explanation of why people from the new member states would want to come to the UK, e.g., because there are available jobs and the social benefit system is an attraction. It is not an argument because, in line with Walton's definition, it is not setting out to persuade readers that Eastern Europeans will want to come to the UK. The marker of explanation summary I contend is 'small wonder' (see sentence k). As evidence for interpreting what I have italicised as an explanation, I draw on a large corpus—the 450 million word Bank of English. In this corpus, there are 625 instances of 'small wonder'. The default is a series of propositions which are followed by 'small wonder', providing a summary explanation such as in:

Of course, paper-based books are a technology that works. They are lightweight, portable, cheap, renewable and recyclable—and free from the bugs, viruses and system crashes that plague your PC. **Small wonder**, then, that the "paperback in your jeans pocket", as Penguin promoted its novels to post-war teenagers, has never been bettered, despite attempts by software and hardware companies to do so for today's e-generations.

In the Bank of English, 405 instances (65%) of 'small wonder' are followed by the relativiser, 'that'. Also, overwhelmingly 'small wonder' begins a sentence (with or without a preceding 'It is'). On the evidence of the Bank of English, language users would seem to be *primed* to expect the use of 'small wonder' as an explicit marker of explanation (paraphrasable as 'it is hardly surprising why') at the beginning of a sentence and followed by a that-clause. This is just the same as happens in the May 1st text.³

2.2 *The Sun Corpus*

I compiled a corpus of all texts in the six weeks prior to May 1st which contain the search terms: '(im)migration', '(im)migrant(s)', 'EU' and 'European'. The reason for using six weeks' worth of texts is because it was six weeks before May 1st that *The Sun* began publishing texts regularly on the imminent expansion of the EU. The corpus consists of 76 texts, a total of 26,350 words, and 1,429 sentences, and is organised in chronological order from March 20th

³ By 'primed', I am referring to Hoey's (2005) argument for 'lexical priming': that words are primed for prototypical lexicogrammatical associations from our experience of how they are used. These associations can include the prototypical role words play in text structure such as with 'small wonder'. Hoey demonstrates primings through exploration of large corpora.

to April 30th 2004. These texts, at least on cursory reading, seemed to form a quasi-campaign to persuade readers of the negative consequences of Eastern European immigration to the UK. In Section 5, my quantitative investigation of the corpus reveals regular associations of words, what is known in linguistics as *collocation*. In Section 6, my qualitative exploration reveals semantic patterns or 'strategies' (see below), in relation to Eastern European immigration, which correlate with these collocations.

In Section 7, I will show that repeated exposure to these strategies potentially primes regular target readers to generate a set of dialogical premises in their reading of the May 1st text's explanation—i.e., generated in dialogue with the explanation—and to derive a biased conclusion through plausible reasoning (Walton, 2006).⁴

2.3 Strategies and their linguistic realization

I take the concept of 'strategies' from Wodak *et al.* (1999). This work, in *Critical Discourse Analysis*, is an examination of the construction of national identity in Europe with a particular focus on Austria. Wodak *et al.* (1999) detail strategies for how national identities are constructed through particular types of language use as well as the linguistic realisations of these strategies.⁵ For example, the authors highlight the strategy of 'emphasising the difference between us and them' (Wodak *et al.*, 1999: 36). Here is Wodak *et al.*'s (1999: 34) definition of a strategy:

⁴ Walton and Reed (2005) show that, while many logic textbooks reconstruct missing premises for enthymemes using deductive logic, this treatment does not work for natural language discourse. For the ten case studies they analyse, they contend that implicit premise reconstruction requires 'a less strict standard of reasoning that is defeasible in nature' (Walton and Reed, 2005: 339), i.e., plausible reasoning.

⁵ Aside from strategies and their linguistic realization, Wodak *et al.* (1999) also indicate *topoi* or argumentation schemes which accompany strategies. They conceptualise '*topoi*' (Greek) or '*loci*' (Latin) classically: '...highly conventionalized parts of argumentation which belong to the obligatory elements of argumentation and take the form either of explicit or inferable premises. They are more or less formal (for example *locus a minore*) or content-related (for example *topos of external constraints*) warrants or 'conclusion rules' which connect an argument or arguments with a conclusion, a claim. As such they justify the transition from an argument or arguments to the conclusion. As standardized argumentation schemes, *topoi* can become integral parts of strategic plans and serve to obtain a specific effect which has been the aim of the strategy' Wodak *et al.* (1999: 34-5).

...we conceive strategies to be more or less ‘automated’ activity plans based on models which are located on the different levels of mental, cognitive organization and which are more or less elaborated. As such the concept of ‘strategy’ is closely connected to the concepts of ‘frame’, ‘scheme’ and ‘scripts’...In contrast to ‘script’, ‘scheme’ and ‘frame’, ‘strategy’ is mainly identified in terms of planned social (in our case, discursive) activities, of the political or socio-psychological aims or functions of these activities, and of the (linguistic) means designated to help realize these aims.

Since in *The Sun*, in the six weeks prior to May 1st, there appeared to be a quasi-campaign—to persuade readers of a set of possibilities, e.g., detrimental effects on social services from huge Eastern European immigration—there seemed to be a ‘social-psychological aim’ behind these predictions. For this reason, I use ‘strategies’ to refer to cognitive models that regular target readers of *The Sun* may have constructed in relation to the predicted effects of Eastern European immigration. Another reason I prefer the concept is because this paper looks at the dynamic relationship between language and the circulation of ideas in relation to immigration. Like Wodak *et al.* (1999), I also make links between strategies and their linguistic realizations.

How I qualitatively identify such strategies in *The Sun* corpus will be constrained by quantitative-based generation of ‘corpus comparative statistical keywords’ from *The Sun* corpus. I do this through using the software, Wordsmith Tools 5.0 (Scott, 2008). There are a number of different uses of the concept of keyword. The next section discriminates types of keyword which are relevant to this article,⁶ and prepares the ground for what I mean by ‘corpus comparative statistical keywords’.

3. Keywords

3.1 Cultural Keywords

One meaning of keyword is associated with Raymond Williams, the Marxist thinker. In Williams (1983), ‘key’ in ‘keyword’ indicates that a particular concept is salient across a culture. So, for example, ‘democracy’ and ‘revolution’ are keywords for Williams.

⁶ There are other uses of the term ‘keyword’. For example, in using search engines to explore the world-wide web, users choose keywords to find matching web pages. See also Bigi (2007) for a different notion of ‘keyword’ in argumentation theory.

Williams (1983) is a socio-historical, diachronic dictionary of keywords where their semantic development over centuries is traced and interrelationships explored. For this work, Williams used the complete Oxford English Dictionary (OED), which runs to several volumes. To help with discriminating different types of keyword, I shall refer to this sort as a *cultural keyword* (as indeed does Stubbs, 1996; 2001). This concept of keyword is used in the argumentation work of Rigotti and Rocci (2005)—see section 4.

3.2 *Corpus-based Cultural Keywords*

Chapters in Stubbs (1996) and (2001) also examine cultural keywords (and we shall see briefly how such examination bears on an argumentation analysis of his in a moment). The difference from Williams is that Stubbs' investigation of cultural keywords is done in the main synchronically and is informed by corpus-based methods. 'Standard' is one of the cultural keywords which Williams (1983) investigates diachronically using the OED. Stubbs (2001) uses a 200-million-word corpus of contemporary English in order to highlight the most common collocates (i.e., words which regularly associate with other words in collocation) of 'standard': 'living', 'high', etc. In other words, the most common collocations are 'living standard', 'high standard', etc. The method is more rigorous than when Williams focuses on contemporary usage since it provides objective quantitative support for the extent to which cultural keywords are being used, and the lexical company they keep. It thus provides a measure of what meanings are being culturally reproduced. We might refer to the cultural keywords Stubbs looks at as *corpus-based cultural keywords*.

Stubbs (1996; 2001) uses corpus-based methods to examine cultural keywords related to argumentation around education policy of the UK Conservative party during the 1980s and 1990s (e.g., the cultural keywords, 'correctness', 'grammar', 'standards'). As an example of this, here is Stubbs (2001: 158), which is largely a restatement of Stubbs (1996: 162). He contends that since different cultural keywords collocate with one another in the language use of different socio-political groups, in turn the kinds of arguments such groups would make would be constrained:

Keywords often inter-collocate, and ideas gain stability when they fit into a frame. Many everyday ideas about language fit very firmly into a frame which contains terms such as:

- standard, standards, accurate, correct, grammar, proper, precise

For linguists, the same terms mean something quite different because they fit into an entirely different lexical field, which contains terms such as:

- dialect, language planning, high prestige language, social variation

These fields are systems of meaning, which use particular vocabulary, take particular things for granted, appeal to different states of knowledge (for example, lay and professional), and therefore allow only particular argumentative moves.

3.3 Corpus-Comparative Statistical Keywords

A large corpus is not only used to provide quantitative support for cultural keywords. It can be used to find a different type of keyword: *corpus-comparative statistical keywords*. These are defined as being statistically more frequent in a text or set of texts than in a large corpus which is taken to be representative of a language. When a large corpus is used for such comparative purposes it is known as a ‘reference corpus’. ‘Keyness’ here is established through statistical measures such as log likelihood value (see Dunning, 1993). A log likelihood value at ≥ 6.63 ($p < 0.01$) confers keyness on a word. Keyness value is proportional to the size of the log likelihood value greater than 6.63. Corpus comparative statistical keywords can be lexical words (e.g., table, beautiful, psychologically) as well as grammatical words (e.g., she, their, to). Imagine a comparison of a corpus of several thousand mobile phone text messages in English with a large reference corpus of English consisting of millions of words from a variety of different genres. It is likely that the grammatical word, ‘da’ (a short version of the definite article), would have relatively high keyness since it is much more likely to feature in text messages than in most other genres. It would then be a corpus-comparative statistical keyword.

3.4 Relationship between corpus comparative statistical keywords and cultural keywords

Corpus-comparative statistical keywords may or may not coincide with (corpus-based) cultural keywords. This is because the former may be: (i) any lexical word and thus part of a very large set or (ii) any grammatical word (a relatively small, finite set). In contrast, (corpus-based) cultural keywords are all lexical words from a much smaller set. Although corpus-based cultural keywords have the advantage that collocation patterns which include them are based

on quantitative evidence, all the same this is quantitative evidence around cultural keywords which have been pre-established by a human interpreter. This may be uncontroversial as in the case of 'democracy' or 'immigration'. Alternatively, choice of cultural keyword may overly reflect the cultural and political sympathies of the human interpreter. For example, Williams (1983) includes 'peasant' as a cultural keyword. In contrast, an advantage of corpus-comparative statistical keywords is that they are established objectively by software. They are not then intuited by human researchers who are open to accusations that their intuitions are arbitrary. However, it must be stressed that though corpus-comparative statistical keywords are generated objectively by software, this objectivity is always relative to a particular reference corpus.

Recently, in relation to argumentation, Stubbs' work has been criticized by Rigotti and Rocci (2005). Since, as I said in Section 1, this article takes cues from Stubbs' work on keywords and its unfolding is affected by engagement with Rigotti and Rocci (2005), in the next section I will set out Stubbs' perspective and my response to Rigotti and Rocci's critique.

4. Rigotti and Rocci (2005)

4.1 Critique of Stubbs

In an anthology devoted to recent work on argumentation, Rigotti and Rocci (2005: 129) take issue with Stubbs' corpus-based position on the relationship between cultural keyword collocation and argumentative text (see section 3.2):

The problem we have with this type of analysis is that focusing one's attention exclusively on quantitative patterns of lexical co-occurrence, regardless of the intention and structure of texts, results in a dangerously simplified image of culture and cultural reproduction. ...in our view a natural language text, slippery and vague as it may be, is not a sort of vegetable soup where words float free.... A text is a coherent sequence of utterances, where coherence is not ensured by repetition of patterns, but by the *congruity* of the meaning of each utterance with the intended effect of the whole.

Rigotti and Rocci's point is one worth taking account of.⁷ To understand cultural / ideological reproduction through texts, analysis of 'quantitative patterns of lexical occurrence' would need to take account of text structure. This is because text structure will have a significant effect on the types of inference readers and listeners generate around cultural keywords. In relation to this issue, the Critical Discourse Analyst, Fairclough (1992: 84), gives the following useful example (which includes the word, 'job', identified as a cultural keyword in Stubbs (1996)):

...what establishes the coherent link between the two sentences 'She's giving up her job next Wednesday. She's pregnant' is the assumption that women cease to work when they have children. In so far as interpreters take up these positions and automatically make these connections, they are being subjected by and to the text, and this is an important part of the ideological 'work' of texts and discourse in 'interpellating' subjects.

'Ideological work' would not take place through 'job' or indeed through 'pregnant', but through the coherence inference linking these words in the two sentences, i.e., coherence is set up through the structure of the text. See O'Halloran (2003) on different types of coherence inference.

4.2 Enthymematic argument structure and cultural keywords

Rigotti and Rocci (2005: 130) go on to usefully explore the link between cultural keywords, and the enthymematic structure of 'natural language' arguments. They give as one example 'He's a traitor. Therefore, he deserves to be put to death' where they identify 'traitor' as a cultural keyword. They unpack this argument syllogistically:

Major premise: *Traitors* deserve to be put to death
(implicit)
Minor premise: He is a *traitor*

⁷ Rigotti and Rocci have an unusual definition of text as 'a coherent sequence of utterances'. Commonly, in discourse analysis, 'cohesion' is regarded as a property of a text (e.g., the co-reference in 'David is a footballer; he plays for AC Milan'). However, 'coherence'—making coherent understanding of text—is not a property of text since it is a mental property (Fairclough, 1992; Cook, 1994; Widdowson, 2004). Coherence of texts takes place, in part, through activations of mentally stored strategies. On this distinction, it is quite possible to have cohesive text which people find difficult to make coherence of (e.g., text in James Joyce's *Finnegans Wake*).

Conclusion: He deserves to be put to death.

They contend that cultural keywords such as 'traitor' play an 'important role in the recovery of the unstated major premise' (Rigotti and Rocci, 2005: 130). Rigotti and Rocci (2005: 131) have a 'working hypothesis for the discovery and testing of cultural keywords:

words that typically have this kind of function in public argumentation within a community are likely candidates to the status of [cultural] keywords in the culture of that community.

Again, I find Rigotti and Rocci's perspective valuable in its engagement with the concept of cultural keyword, and how they extend the concept into the issue of implicit premise recovery. I support the *spirit* of their hypothesis in that I also maintain that there is a link between identifying cultural keywords and recovering implicit premises.

Nevertheless, there are things to take issue with. The textual examples they provide, such as the example of 'traitor', appear to be concocted rather than taken from a relevant empirical dataset. The latter needs to be drawn on; otherwise, the exemplification seems arbitrary. In any case, other support for 'traitor' as a cultural keyword would be needed; without it, their argument is circular. To establish the judgment that Rigotti and Rocci make in the quotation above, one needs a large amount of relevant data, again *pace* Rigotti and Rocci's non-empirical approach. And in relation to the earlier quote (see Section 4.1), I am not sure how one can make a convincing case that text structure is important in cultural reproduction via cultural keywords if text structure is in effect being concocted in exemplification. A 'natural language text' should be just that. Finally, despite the emphasis on text structure that they make, they do not seek to establish possible patterns between cultural keywords and grammatical words which provide text structure, e.g., 'so' and 'therefore'.

In the next section, I take one cue from Rigotti and Rocci's (2005) critique. Text structure, and thus grammatical word meaning, is an important consideration in how I identify strategies in *The Sun* corpus; in turn, this will feed into how I recover implicit premises. As a way of identifying strategies on *immigration* from *Eastern Europe*, I examine the extent to which these cultural keywords collocate, across stretches of text, with objectively generated corpus-comparative statistical keywords (henceforth CCSKs). Specifically, I investigate whether any grammatical CCSKs are potentially providing text structure in concert with lexical CCSKs. I explore whether such CCSKs help to realise

strategies and, in turn, could indicate the kinds of premises regular target readers may generate in dialogue with the propositions of the explanation in the May 1st text. Thus, in contrast to Rigotti and Rocci (2005), I show that implicit premise recovery needs to go beyond identification of *single* cultural keywords. Moreover, I show that an advantage of identifying strategies round CCSKs is that arbitrariness of text structure identification can be reduced (*pace* Rigotti and Rocci, 2005), and in turn, arbitrariness of strategy identification can also be reduced.

5. CCSKs and text structure

5.1 CCSKs

Using Wordsmith Tools 5.0, the first procedure is finding the CCSKs of *The Sun* corpus. The reference corpus used is BNC-baby, an approximately four million word sample of the British National Corpus, which consists of around one million words each of academic prose, conversation, fiction and newspaper text. This reference corpus is chosen since its genres are mainstream ones and, as a whole, can be taken as a ‘snapshot’ of mainstream use of English. To ascertain CCSKs, firstly wordlists for both corpora need to be generated. (A wordlist provides word frequencies in a corpus). Then, these wordlists are compared by the software to establish log likelihood values, i.e., measures of keyness. *Table 1* shows the 10 highest values for ‘keyness’. As can be seen, the keyword ‘immigration’ has high keyness, as does ‘EU’. What this means is that ‘immigration’ and ‘EU’ occur statistically more often in *The Sun* corpus than in the reference corpus. Since the corpus texts are selected in part using these search terms, it is not surprising that they should feature as CCSKs.

CCSKs provide a rough snapshot of salient topics in a corpus. CCSKs with relatively high values indicate the UK politicians being recurrently referred to over this period: Blair (Prime Minister), Blunkett (Home Secretary) and Hughes (Home Office minister) and also other stories connected with immigration, immigration officialdom, etc., but not necessarily connected with Eastern Europeans.

CCSK	Frequency	% in Corpus	Reference Corpus Frequency	Log likelihood/Keyness
immigration	135	0.49	40	1,164.79
EU	107	0.39	2	1,051.86
Blair	105	0.38	12	974.53
's	155	0.56	630	781.22
Britain	129	0.47	671	594.45
Hughes	58	0.21	58	420.83
Poland	44	0.16	20	361.40
Blunkett	36	0.13	4	334.56
PM	38	0.14	12	325.59
asylum	36	0.13	15	298.92

Table 1. The 10 highest CCSK values for *The Sun* corpus⁸

5.2 Keyword Dispersion Plot

For a word to occur as a CCSK, it does not necessarily need to occur very frequently. For example, 'Trajce' (a town in Poland) has keyness of 30.04; it only occurs 3 times in *The Sun* corpus but does not occur at all in the reference corpus. Because I am interested in seeing whether repeated strategies in *The Sun* corpus can be identified around CCSKs, keyness is not the only criterion I need to take into account. Another is CCSK frequency. However, it must be borne in mind that out of the 76 texts, some CCSKs may occur very frequently in just a handful of texts. Because my focus is on identifying regularly recurrent strategies over the whole of the corpus, CCSKs which feature in a concentrated burst are of less relevance to me. I need then to find CCSKs which are not only frequent but well dispersed across the six weeks' worth of texts. The dispersion plot facility of Wordsmith Tools 5.0 is useful for my purposes. *Table 2* shows the 10 highest dispersion plot values for CCSKs. The dispersion value is a number between 0 and 1 with those CCSKs with a value close to 1 having the greatest dispersion.⁹ Organisation is in terms of descending dispersion values.

⁸ Wordsmith Tools treats 's' (as in 'Britain's' or 'he's') as a whole word.

⁹ Plot dispersion uses the first of the 3 formulae supplied in Oakes (1998: 190-191).

CCSK	Dispersion	Log likelihood/Keyness
#	0.948	102.57
said	0.919	52.30
But	0.915	28.93
who	0.914	45.43
's	0.884	781.22
By	0.880	41.31
yesterday	0.879	44.88
Britain	0.871	594.45
has	0.869	56.00
We	0.868	29.62

Table 2. The 10 highest CCSK dispersion plot values for *The Sun Corpus*¹⁰

Interestingly, the grammatical word ‘But’ (i.e., with a capital ‘B’) has the third highest dispersion value (0.915); it has keyness of 28.93. This is interesting for my focus since, as a grammatical word, it signals a contrastive relation across sentences and thus creates text structure. On the basis of corpus evidence, Scott and Tribble (2006: 72) indicate that ‘it seems that the items which are most likely to be key are nouns, determiners, prepositions, pronouns’ and thus not a sentence coordinator like ‘But’. So, the keyness of ‘But’ is significant. In contrast, ‘but’ (i.e., with a lower-case ‘b’) does not have keyness since it has a negative value (-30.53). There are 85 instances of ‘But’ and 49 instances of ‘but’ in *The Sun* corpus (0.31% and 0.18% of the total number of words respectively). However, in BNC-baby there are 6,594 instances of ‘But’ and 14,619 instances of ‘but’ (0.16% and 0.36% of the total number of words respectively). The ratio of quantities of ‘But’ / ‘but’ in *The Sun* corpus is almost the opposite to that of BNC-baby. So, on this comparison, that ‘But’ has keyness is even more significant. To sum up: in *The Sun Corpus* ‘But’ has significant: keyness, frequency, and dispersion.

¹⁰ Even though ‘#’ and ‘said’ have higher dispersion than ‘But’, I don’t examine these CCSKs. ‘#’ refers to numbers. I don’t examine numbers, in this article, since these don’t provide a grammatical relation and thus text structure. For the same reason, I don’t focus on the lexical word ‘said’. The grammatical word ‘By’ may, at first sight, seem a candidate for providing text structure. However, the reason for its dispersion and keyness is because it features in author by-lines. The other CCSKs in Table 2 are not good candidates for providing text structure. Naturally, exploration of potential text structural words which have keyness and dispersion but which don’t feature in Table 2 could take place, but this is beyond the scope of this article.

The results for 'But' are interesting especially since, in starting a sentence, 'But' has prominence. This is even more the case if it begins a sentence which initiates a paragraph. And indeed, it is common in *The Sun* for paragraphs to be one sentence in length. So, if 'But' is providing text structure for strategies on a regular basis, in turn the strategies will have prominence. Because 'But' rather than 'but' has prominence, I don't collapse all instances of 'But'/'but' in my analysis. Furthermore, if 'But' is providing grammatical structure for strategies, it will be doing so in concert with lexical words or perhaps with other grammatical words. To reduce arbitrariness of strategy identification and thus circularity, I need to discern whether there are collocations of 'But' with lexical or other grammatical CCSKs. Finding such collocations can be done by employing another function of Wordsmith Tools 5.0—keyword links.

5.3 Keyword links

The 'keyword links' function of Wordsmith Tools shows the number of CCSKs which collocate with a particular CCSK in a designated word span. The highest word span possible is 25 words to the left and 25 words to the right of a search term. I choose this span so as to capture the maximum possible co-occurrences with 'But'. There are 9 CCSK links for 'But' for this span (one which crosses sentence and paragraph boundaries):

Britain, he, home, immigration, Mr, UK, vote, will,
yesterday.

Interestingly, one lexical CCSK is 'immigration'—a cultural keyword. Furthermore, and again in relation to the point made by Scott and Tribble (2006: 72), not all of these CCSKs are nouns or pronouns (there are no prepositions or determiners) since they include the modal verb 'will'. In fact, 'will' (171 instances) has relatively high keyness (152.66).

The next stage is to electronically search through the entire *Sun Corpus* and highlight all instances of 'But', as well as the 9 CCSK links to 'But', and then inspect whether regular strategies can be identified around any of these 'But CCSK links' in relation to the cultural keyword, '(Eastern European) immigration'. The Appendix contains fragments of texts from *The Sun* corpus selected on 'But' and CCSKS links. The amount of text chosen for exhibition in the Appendix is what is sufficient to provide semantic coherence of *The Sun*'s strategies via the 'But CCSK links'. Thus, the amount of text exhibited varies given the range of strategies produced. In turn, while the link span is restricted in Wordsmith Tools to 25 words

before and 25 words after a CCSK, the strategy highlighting may go beyond this span.

6. Strategy types via ‘But’

6.1 *The Five Strategy Types*

In exploring the text fragments of the Appendix, I identified five different but related strategy types associated with ‘But’ and the CCSK links. For the strategies identified, the lexical CCSKs which are most repeatedly linked to ‘But’ are ‘UK’, ‘Britain’, ‘immigration’, and ‘home’ (in ‘Home Office’, the UK government department which deals with immigration). There is one grammatical CCSK which is repeatedly linked to ‘But’; this is ‘will’. Below, I indicate the five strategy types in the corpus via representative examples with the CCSKs bolded. ‘But’ signals negative contrast in all of the 36 strategies identified. All relate to projections of immigration generally, and Eastern European immigration specifically.

Strategy 1 (x 18): for (Eastern European) immigration, there is UK government or official agency incompetence

With this strategy, there is often a reporting of government figures for immigration from new EU countries, the actions of immigration officials, etc. Then ‘But’ begins a sentence or paragraph which contrasts negatively with the UK government perspective and which sets up retrospective negative evaluation, either explicit or implicit, that the government, officials etc are incompetent as in Example 1:

(1) 26 April

The **Home** Office estimates that no more than 13,000 workers **will** come each year.

But others put the figure as high as 54,000 due to high unemployment in the East.

Strategy 2 (x 8): for (Eastern European) immigration, (fear, worry, challenge from) large numbers predicted to arrive in the UK

In this strategy, ‘But’ is used to provide negative contrast in relation to the UK facing huge (Eastern European) immigration which is a challenge to the government, or a fear/worry for various reasons, as in Example (2):

(2) 27 April: Stay Strong

THE SUN welcomes Tony Blair's assurance that **he will** keep a tight grip on **immigration** from Eastern Europe.

He is right to recognise that people are worried about what **will** happen after May 1.

But the PM faces a huge challenge.

Thousands are heading for **Britain** in search of a better life.

Strategy 3 (x 4): for (Eastern European) immigration, there will be strain on social services

With this strategy, there is a projection of numbers of Eastern European 'migrants' coming to the UK. Then, the strategy via 'But' indicates, through negative contrast, that this will lead to overstretched UK social services as in Example (3):

(3) 26 April

THOUSANDS of plane, train and bus seats had been snapped up last night by poor East Europeans seeking a better life in **Britain**.

They are free to come here when ten new countries join the EU this Saturday, May 1.

But it is feared that in some areas overstretched **UK** services like schools and hospitals **will** be unable to cope with the influx.

Strategy 4 (x 4): for (Eastern European) immigration, there will be some illegal (EU) status

Strategy 4 relates to stories about illegal immigrants in the UK, or predictions that there will be immigrants arriving in the UK from Eastern Europe who will have illegal EU citizenship (posing as Polish below) as in Example (4):

(4) 2 April

Lieutenant Miroslaw Szacillo, 46, of the Polish Border Guards, assured The Sun: "Only the most serious criminals with big money can afford to buy false documents in Poland. We make stringent checks on our borders and we are getting more equipment to detect fake documents."

But Poles who do get in and find work in **Britain will** qualify for a range of benefits

including free healthcare, child tax credit, child benefit, working tax credit, housing benefit and council tax benefit.

‘But’ is used, in negative contrast, in predicting that people falsely claiming to be a member of a new EU country will be able to claim benefits.

Strategy 5 (x 2): for (Eastern European) immigration, there will be some criminality

In Strategy 5, ‘But’ is used to signal, through negative contrast, the prospect of serious criminality characterising some immigration from Eastern Europe, as in Example (5):

(5) 31 March

Many new arrivals **will** be good news for **Britain**.

But some **will** be up to no good. Gun-happy crime syndicates have already set up vicious vice and drug rackets.

Others want to do our nation harm. The Wall Street Journal says Islamic fanatics are using **immigration** as a “Trojan horse to expand jihad, or holy war”.

In sum: strategies relating to Eastern European immigration have been identified around CCSK links with ‘But’, i.e., they have been grounded in text structure. I hope I have shown the value of grounding interpretation of strategies in corpus-generated lexicogrammatical patterns since arbitrariness of identification, and circularity of analysis to interpretation to analysis, has been considerably reduced. It should be said that there seems to be other evidence for these 5 strategies in the dataset but where the strategies are *not* realised through ‘But’. However, to avoid accusations of arbitrariness, identification of further strategies would also need to be grounded in lexicogrammatical patterns constituted by CCSKs. Further investigation along these lines has been beyond the scope of this article.

6.2 The interrelatedness of the strategies and plausible reasoning

The most common strategy is Strategy 1 at 18 instances; the next is Strategy 2 at 8 instances. Individually, Strategies 3, 4, and 5 are fewer in number than Strategies 1 and 2. However, they are different in kind to Strategies 1 and 2 since they are predicted

negative consequences of projected Eastern European immigration; taken together, Strategies 3, 4 and 5 amount to 10 instances.

Interestingly, while the strategies are semantically distinct in the texts, they are often in close proximity to each other, linked to one another around the 'But CCSK links'. Consider Example (6):

(6) 26 April

DON'T blame the people of *Eastern Europe* for heading for **Britain**. (Strategy 2).

The Government has put out the welcome mat. (Strategy 1) And to *hard-up foreigners* this looks the land of milk and honey.

But most will end up in unskilled low-paid jobs in the South East.

How **will they** afford to live? And how can schools and hospitals which are already at breaking point find room for *them*? (Strategy 3) [my italics]

Strategy 1, seemingly so different to Strategies 2-5, links in the data to other strategies as can be seen in the example above. Because of regular interrelating of strategies around 'But CCSK links' over the 6 weeks prior to May 1st, potential has been created for the reader to plausibly reason, in reading of a particular strategy, other strategies. This is because regular interrelating of strategies sets up a series of linked contextual assumptions from which can be plausibly reasoned contextual effects (Sperber and Wilson, 1995). For example, on the basis of exposure to text fragments such as Example 6, when:

- a reader is exposed to Strategy 1, they may well plausibly reason Strategy 2; i.e., highlighting of government incompetence could lead to the inference that large numbers of migrants will arrive in the UK;
- when a reader is exposed to Strategy 1 and/or Strategy 2, they may well plausibly reason Strategy 3; i.e., highlighting of government incompetence could lead to the inference that large numbers of migrants will arrive in the UK, and the (further) reasoning that this will lead to strain on social services.

Consider also Example (7):

(7) 2 April

The huge number of illegal immigrants who have sneaked into **Britain** is a major worry to

voters because of the strains being placed on schools, hospitals and housing.

But for their first five years in power, all the Government did was make matters worse with a tangled mess of rules that didn't work. (Strategy 1)

For the past two years they have talked tough but achieved what? Police don't even arrest lorryloads of illegals (Strategy 4)—they ask them to make their own way to the **immigration** HQ in Croydon. (Strategy 1)

Due to the interrelating of these strategies in this fragment and in others, when:

- a reader is exposed to Strategy 1, they may well plausibly reason Strategy 4, i.e., highlighting of government incompetence could lead to the inference that there will be illegal migrants amongst new Eastern European migrants.

Finally, consider Example (8):

(8) 2 April

Almost 10,000 forged passports and other ID documents were discovered at British ports in 2002, the latest year for which figures are available—an increase of 46 per cent on 2001.

But an unknown number of illegals with false papers make it into **Britain** each year. (Strategies 1 and 4)

Forged passports are commonly used by terrorists to slip in and out of countries undetected.

Two al-Qaeda terrorists, jailed in **Britain** in 2003, were discovered with hundreds of false travel documents. (Strategy 5)

Due to the interrelating of these strategies in this fragment, when:

- a reader is exposed to Strategy 1 and/or Strategy 4, they may well plausibly reason Strategy 5, i.e., highlighting of government incompetence could lead to the inference that there will be illegal immigrants amongst new Eastern European migrants, and further reasoning that serious criminals will enter the UK.

See the Appendix for further evidence of interrelatedness of strategies along 'But CCSK links', and thus potential for the above types of plausible reasoning, since many of the text fragments exhibit more than one strategy. The preponderance of Strategy 1 (18 instances) across the corpus might seem to protect *The Sun* from accusations of bias against migrants (i.e., on the whole, it seems to blame the government, not migrants). However, the interrelatedness of Strategy 1, along the 'But CCSK links', with other strategies means that in practice it would be disingenuous for *The Sun* to claim that it is only criticizing the government.

It is also worth noting that all of the plausible inferences detailed above are causal consequent inferences. A consensus in current psycholinguistic evidence indicates how causal consequent inferences are only weakly generated at best in the absence of relevant information, i.e., absence of relevant contextual assumptions. However, causal consequent inferences can be generated if 'necessary information can become available quickly' (O'Halloran, 2003: 181). The strategies then are unlikely to be generated in the reading of a non-regular, non-target reader. However, as the corpus analysis has shown, they stand a good chance of being generated, and chains of plausible reasoning established, if readers have been regularly exposed to these interrelated strategies in the six weeks leading up to the May 1st text.

6.3 'But' and negative contrast priming

The corpus investigation shows there is evidence that 'But' has potentially been 'primed' for regular *Sun* readers, in the Hoey (2005) sense, to indicate negative evaluative contrast, specifically in relation to predictions around future migration to the UK. Significantly, 'But' is never used in the corpus to indicate a positive evaluation of migrants as contrasting with a previously stated negative one (e.g., 'Some migrants from Eastern Europe may be criminals. But the majority will be good for the economy'). The priming of 'But' relates also to the grammatical word, 'will', and lexical words such as 'immigration'; the priming is actually lexicogrammatical. So, regular readers of *The Sun* during the six-week period could well expect 'But' to preface a negative prediction (most likely around 'will') about Eastern European migration to the UK.

Conversely, because of the nature of the repeated lexicogrammatical CCSK collocations, readers could well have been primed to negatively evaluate information on Eastern European immigration in related subsequent texts, even if the information is expressed neutrally. In other words, they have been potentially primed to make a contrastive negative inference from

repeated exposure to the five interrelated strategies. In subsequent related texts, references to ‘Eastern Europeans’ could thus be regarded as *But strategy activators*. For the May 1st text (see Section 7), I will highlight how this could work. Another point to make is that, given the high level of interrelation of strategies via the ‘But CCSK links’, regular readers are actually being subtly primed (should they accept this positioning) into interrelating the strategies.

7. The May 1st text and Implicit Dialogical Premise Reconstruction

7.1 How the text positions regular target readers into negative evaluative dialogue

For ease of reference, here again are the five interrelated strategies which were identified around CCSK links with ‘But’:

Strategy 1 (x 18): for (EE) immigration, there is UK government or official agency incompetence

Strategy 2 (x 8): for (EE) immigration, (fear/worry/challenge from) large numbers predicted to arrive in the UK

Strategy 3 (x 4): for (EE) immigration, there will be strain on social services

Strategy 4 (x 4): for (EE) immigration, there will be some illegal (EU) status

Strategy 5 (x 2): for (EE) immigration, there will be some criminality.

Annotated below in italics is an interpretation of how regular readers of *The Sun* who have accepted the positioning of the interrelated five strategies, over the six weeks, could plausibly infer the following in their reading of the May 1st text:

(i) strategies as ‘negative But contrastive inferences’ in dialogue with strategy activators (i.e., text which refers to Eastern Europeans / semantically relates to the strategies) in the May 1st report.

(ii) plausible inferences from strategies indicated in Section 6.2.

Strategies activated will vary from all to some for such readers. I am not suggesting, then, that all the ‘But strategy negative contrastive inferences’ annotated below would be activated for the above readers, nor all the plausible inferences. Moreover, it is

possible that unannotated parts of the May 1st text will be 'alive' as strategy activators for some target readers. The explanation part of the text is bolded.

- a. *Migrants are the first of many*
By NICK PARKER CHARLES RAE and CHARLES YATES
- b. HISTORY will be made today as new EU citizens begin leaving their homelands for new lives in Britain.
- c. They will come in their hundreds [*ACTIVATE STRATEGY 2 AS CONTRASTIVE INFERENCE: BUT for EE immigration (fear, worry, challenge from) large numbers predicted to arrive in the UK: PLAUSIBLY REASON STRATEGY 3: BUT for EE immigration, there will be strain on social services*] by coach, train and plane, desperate for decently-paid work—or any job at all [*ACTIVATE STRATEGY 5 AS CONTRASTIVE INFERENCE BUT for EE immigration, there will be some criminality*].
- d. They are citizens of ten new member states of the EU who now have the right to live and work in Britain.
- e. Those states are the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia, Cyprus and Malta.
- f. **Eight are poverty-stricken former Soviet states in Eastern Europe—with a total population of 75 million** [*ACTIVATE STRATEGY 2 AS CONTRASTIVE INFERENCE : PLAUSIBLY REASON STRATEGY 3*].
- g. **Our government says there are half a million jobs waiting here to be filled** [*ACTIVATE STRATEGY 1 AS CONTRASTIVE INFERENCE: BUT for (EE) immigration, there is UK government or official agency incompetence: PLAUSIBLY REASON STRATEGIES 2, 3, 4, 5*].
- h. **And if the new arrivals find legitimate work in Britain** [*ACTIVATE STRATEGY 5 AS CONTRASTIVE INFERENCE*], **paying tax and national insurance, they will qualify for a range of benefits** [*ACTIVATE STRATEGY 3 AS CONTRASTIVE INFERENCE*].
- i. **These will include free health care, child tax credit, child benefit, working tax credit and housing benefit** [*ACTIVATE STRATEGY 3 AS CONTRASTIVE INFERENCE*].
- j. **Other EU countries like France and Germany have ruled out free access to jobs for up to seven years** [*ACTIVATE STRATEGY 1 AS CONTRASTIVE INFERENCE: PLAUSIBLY REASON STRATEGIES 2, 3, 4, 5*].
- k. **Small wonder that eager men and women used to working for a pittance in countries with huge unemployment are snapping up tickets to the UK** [*ACTIVATE STRATEGY 2 AS CONTRASTIVE INFERENCE: PLAUSIBLY REASON STRATEGY 3*].

- l. People out to get a head start in the exodus [*ACTIVATE STRATEGY 2 AS CONTRASTIVE INFERENCE: PLAUSIBLY REASON STRATEGY 3*] crammed into coaches yesterday at stations in Warsaw, Poland, the Czech capital Prague and the Slovak capital Bratislava.
- m. Seventy-one LATVIANS smiled as they boarded a double-decker bus in the capital Riga for a 24-hour journey west.
- n. Coach firm Eco-Lines said they had put on an extra service to London this week for the £75 trip.
- o. All services are booked for the next week [*ACTIVATE STRATEGY 2 AS CONTRASTIVE INFERENCE : PLAUSIBLY REASON STRATEGY 3*].
- p. ESTONIANS are travelling from their capital Tallinn, via Helsinki in Finland.
- q. Budget airlines like Sky Europe, which has services from Slovakia to Britain, have reportedly sold out all flights [*ACTIVATE STRATEGY 2 AS CONTRASTIVE INFERENCE: PLAUSIBLY REASON STRATEGY*].
- r. POLAND'S Air Polonia alone has bookings for about 4,000 migrant workers this month [*ACTIVATE STRATEGY 2 AS CONTRASTIVE INFERENCE: PLAUSIBLY REASON STRATEGY 3*].
- s. Warsaw-based coach operator Orbis expects migrants to fill 1,500 of its monthly 4,500 seats to Dover [*ACTIVATE STRATEGY 2 AS CONTRASTIVE INFERENCE : PLAUSIBLY REASON STRATEGY 3*].

(source: <http://www.thesun.co.uk> 1st May 2004)

In sum: while the May 1st text does not explicitly evaluate potential migrants negatively, as a result of prior positioning readers could well be reproducing evaluative meanings such as 'new migrants will be a drain on social services, schools and hospitals' and 'many migrants will be criminals since the government/official agencies will have problems in preventing their entry to the UK'. Audience reception studies would, of course, be needed to establish these claims.

Let me return to the explanation in the May 1st text which I identified in Section 2.1. In Section 7.2, I interpret how potential 'But strategy' dialogue with the explanation by regular target readers can allow the explanation to be read by them as an argument. This is because strategies activated can act as premises from which a biased conclusion can potentially be derived by these readers. By indicating how premises could be generated dialogically, by regular target readers, around the propositions of the explanation, in Section 7.2, I reconstruct the explanation in the May 1st text as an argument.

7.2 *Reconstructing the explanation as argument*

Below, I reproduce from Section 2.1 my propositional analysis of the explanation in the May 1st *Sun* news text (in italics). Where I interpret semantic content in the propositions of the explanation as potentially functioning as strategy activators, I include the inference which regular target readers could well generate (on the basis of Section 7.1). Since these inferences can become premises in a piece of argumentative reasoning which leads to a conclusion, I refer to these inferences as implicit premises.

Proposition 1. Eight [new member states] are poverty-stricken former Soviet states in Eastern Europe—with a total population of 75 million (sentence f).

Premise 1 (implicit): BUT there is (and the target reader potentially has fear of) a large number of migrants (escaping poverty) arriving in the UK. (Strategy 2)

Proposition 2. Government says that there are 500,000 jobs in the UK (sentence g).

Premise 2 (implicit): BUT for EE immigration, there is government incompetence. (Strategy 1)

Proposition 3. There are social benefits for migrants if they find legitimate work in UK (sentences h and i).

Premise 3 (implicit): BUT for EE immigration, potential illegitimate (criminal) behaviour. (Strategy 5)

Premise 4 (implicit): BUT for EE immigration, there is strain on social services. (Strategy 3)

Proposition 4. Two other major economies have ruled out free access to jobs (sentence j).

Premise 2 (implicit): BUT for EE immigration, there is government incompetence. (Strategy 1)

Proposition 5. Eastern Europeans do not earn much (sentence k).

Proposition 6. There is much unemployment in Eastern Europe (sentence k).

Premise 1 (implicit): BUT there is (and the target reader potentially has fear of) a large number of migrants (escaping poverty) arriving in the UK. (Strategy 2)

Explanation summary on basis of propositions:

All this explains why people from Eastern European countries want to come to UK (sentence k).

Argument conclusion on basis of implicit premises:

Therefore (the target reader is afraid that) large numbers of migrants will place strain on social services and this is due, in part, to government incompetence since numbers of immigrants will be greater than predicted.

By grounding implicit premise recovery on the strategies I identified, I have in effect based this recovery on linking a cultural keyword, '(Eastern European) immigration' to lexical CCSKs and to the grammatical CCSK 'But', i.e., to text structure. Moreover, the state of affairs I have uncovered using the corpus-based method is one which helps *The Sun* to evade charges of bias. This is because the bias is not in the text as such (i.e., in the explanation), but would be in the mind of the reader (i.e., in the argument). Another aspect of this corpus-based examination that is worth commenting on is as follows. In the original proposition analysis, there would seem to be an implicit proposition that Eastern Europeans can earn much more money in the UK than they can in their home country; see sentence k ('working for a pittance'). Interestingly, the seeming implicitness of this proposition increases the likelihood that regular target readers will understand that Eastern Europeans are attracted by social benefits—which in turn will lead to a strain on social services—rather than understand that Eastern Europeans are attracted by the higher wages in the UK. In other words, that this proposition seems implicit in the explanation facilitates reading of the explanation as the argument I have reconstructed. Echoing the quote from Fairclough (1992) in Section 4.1, in the right priming conditions, implicit propositions can do 'ideological work'.

The derivation of the conclusion is through a reconstruction of plausible reasoning rather than deductive reasoning. It is difficult to say, of course, whether the argument conclusion would be realised so sharply and explicitly in target readers' minds as I have modelled, or whether all the premises detailed would be realised. Given that Strategies 3, 4 and 5 leak into one another, and these strategies and Strategies 1 and 2 are so interrelated, which, in turn, can give rise to plausible inferences (see annotation of text in Section 7.1), the conclusion I have modelled is plausible but could well be realised fuzzily. It is probably better, then, to refer to this reasoning as *plausible fuzzy reasoning*. The degree of fuzziness depends, in part, on the degree of processing effort invested by the

reader, i.e., the extent to which they are reading casually (see O'Halloran, 2003). I would contend that much reasoning by casual readers, on the basis of regular positioning by patterns of evaluative meanings (e.g., in the media), is of this order. So while the goal of the writers of the May 1st text may have been to provide an explanation of why people from new member states would want to come to the UK (e.g., because there are available jobs; the social benefit system is an attraction), regular readers of *The Sun* are likely to read the explanation fuzzily—as a hybrid of an explanation and an argument with a conclusion. Some may give more weight to plausibly reasoning to a conclusion over their processing of an explanation and vice-versa; all will depend on the degree to which they have accepted the positioning of the strategies.

8. Conclusion

In this article, I have shown the following:

- the value of a corpus of texts for highlighting how regular target readers can, on the basis of prior positioning, be primed to read an argument into an explanation. I have indicated how this state of affairs can favour a popular tabloid newspaper, and potentially other types of media, because accusations of bias can be rebutted: a biased conclusion is in the minds of regular target readers and not on the page.
- the value of generating corpus-comparative statistical keywords (CCSKs) and linking them to cultural keywords for the business of implicit premise recovery. One reason that CCSKs are useful is because analysis of them helps to reveal text structure around cultural keywords, which is important for the recovery of implicit premises. Another reason is because their generation reduces arbitrariness in implicit premise recovery. In turn, circularity of analysis to interpretation to analysis is also reduced.
- how corpus-based methods can help to illuminate a form of plausible reasoning—plausible fuzzy reasoning—where premises can attach themselves dialogically to the propositions of an explanation in the reading of regular target readers.

Acknowledgements

Many thanks to Jean Goodwin, Chris Reed, and Catherine Williams for their very helpful feedback on an earlier version, as well as to the anonymous reviewers and the editors for very useful comments that helped to improve this article.

References

- Anderson, R.C, Chinn, C., Chang, J., Waggoner, M., and Yi, H. (1997). On the logical integrity of children's arguments, *Cognition and Instruction*, 15: 135-167.
- Bigi, S. (2007). Keywords in argumentative texts and their persuasive power. In F.H. van Eemeren, J.A.Blair, C.A. Willard, & B. Garssen (Eds.), *Proceedings of the sixth conference of the international society for the study of argumentation*, pp.129-135. Amsterdam: SicSat, International Center for the Study of Argumentation.
- Burnyeat, M. (1994). Enthymeme: Aristotle on the logic of persuasion. In D.J. Furley & A. Nehemas, (Eds.), *Aristotle's Rhetoric: Philosophical Essays*, pp. 3-55. Princeton, NJ: Princeton University Press.
- Coffin, C. and O'Halloran, K.A (2008). New directions, New methods. In C. Coffin & K.A. O'Halloran, K.A (guest Eds.), *Researching argumentation in Educational Contexts: New Directions, New Methods, International Journal of Research and Method in Education*, 31(3): 219-227.
- Coffin, C. and O'Halloran, K.A. (forthcoming) *Argument Reconceived? Educational Review*, 61(3).
- Cook, G. (1994). *Discourse and literature: The interplay of form and mind*. Oxford: Oxford University Press.
- Degano, C. (2007). Presupposition and dissociation in discourse: a corpus study. *Argumentation*, 21: 361-378.
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence in *Computational Linguistics*, 19 (1): 61-74.
- Fairclough, N. (1992) *Discourse and social change*, Cambridge: Polity Press.
- Govier, T. (1987). *Problems in argument analysis and evaluation*. Providence, RI: Foris.
- Hoey, M. (2005). *Lexical priming* . London: Routledge.
- Katzav, J. and Reed, C. (2008). Modelling argument recognition and reconstruction. *Journal of Pragmatics* 40: 155-172.
- Oakes, M. (1998). *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.

- O'Halloran (2003). *Critical discourse analysis and language cognition*. Edinburgh: Edinburgh University Press.
- O'Halloran, K.A and Coffin, C (2004) Checking overinterpretation and underinterpretation: Help from Corpora in critical linguistics. In C. Coffin, A. Hewings & K.A. O'Halloran (Eds.), *Applying English Grammar: Functional and Corpus Approaches*, pp. 275-297. London: Hodder Arnold.
- Reed, C. (2006). Preliminary results from an argument corpus. In Eloína Miyares Bermúdez & Leonel Ruiz Miyares (Eds), *Linguistics in the twenty-first century*, pp. 185-196. Cambridge: Scholars Press.
- Reed, C. and Rowe, G. (2001). Araucaria: software for puzzles in argument diagramming and xml. Technical report, Department of Applied Computing, University of Dundee.
- Rigotti, E. and Rocci, A. (2005). From argument analysis to cultural keywords (and back again). In F.H. van Eemeren & P. Houtlosser (Eds.), *Argumentation in Practice*, pp.125-142. Amsterdam: John Benjamins.
- Schwarz, B. and Asterhan, C. (forthcoming). Argumentation and reasoning. In K. Littleton, J. Kleine Staarman & C. Wood (Eds.), *Elsevier Handbook of Educational Psychology*. London: Elsevier.
- Scott, M. (2008). *Wordsmith tools version 5.0*. Oxford: Oxford University Press.
- Scott, M. and Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.
- Sperber, D. and Wilson, D. (1995) *Relevance*, 2nd ed. Oxford: Blackwell.
- Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford: Blackwell.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.
- Walton, D. (1996). *Argumentation schemes for presumptive reasoning*. Mahwah, NJ: Erlbaum.
- Walton, D. (2006). *Fundamentals of critical argumentation*. Cambridge: Cambridge University Press.
- Walton, D. and Reed, C. (2005). Argumentation schemes and enthymemes. *Synthese* 145: 339-370.
- Widdowson, H.G. (2004). *Text, context, pretext: Critical issues in discourse analysis*. Oxford: Blackwell.
- Williams, R. (1983). *Keywords: A vocabulary of culture and society*, 2nd ed. London: Flamingo.
- Wodak, R., de Cillia, R., Reisigl, M. and Liebhart, K. (1999). *The discursive construction of national identity*. Edinburgh: Edinburgh University Press.

Zagar, I.Z. (2007). Arguing from large corpora: Some epistemological and methodological dilemmas; experimental study aimed at quantifying political argumentation. In F.H. van Eemeren, J.A. Blair, C.A. Willard, B. Garssen (Eds.), *Proceedings of the sixth conference of the international society for the study of argumentation*, pp.1553-1558. Amsterdam: SicSat, International Center for the Study of Argumentation.

Appendix: Strategies (and their interrelatedness) around 'But CCSK links'

Strategy 1 (x 18)

30 March

"The applicants rarely know what is in their business plan, cannot speak English, and have absolutely no knowledge or experience in the type of skills needed for respective businesses," wrote **Mr** Ramsden.

But the applications were "unfortunately" being approved by IND against the "strongest recommendations" of the embassy.

31 March

Last Monday, a senior British diplomat was also suspended for lifting the lid on MORE blunders. James Cameron, who was in charge of visa applications in Bucharest, revealed bogus applications from Romania and Bulgaria—not among the EU eight—were being nodded through by the suitcase-load.

He claimed that many had previously been refused tourist or student visas for fear they were benefit scroungers.

But yesterday it emerged alarm bells were already being sounded in high places 18 months ago.

31 March

He said: "Interviewed applicants rarely knew what was in their business plans. It was often clear that applicants had no idea about the trade they were supposed to be setting up in the **UK**."

But he claimed the dodgy applications were rubber-stamped by **immigration** chiefs in Leeds—even when they were warned of "flaws".

31 March

An official in Romania wrote to the **UK immigration** department "strongly" urging them to refuse him, saying she did not find his statement "credible".

But on July 24, 2003, the **Home** Office granted him a visa.

31 March Lost warning

MR C, a Romanian, applied for a visa in August 2003.

The assistant at the British Embassy in Bucharest said: "We phoned his company and they'd never heard of him".

But he got his visa after the **Home** Office “lost” the letter suggesting his claim was bogus .

1 April

He was sent to a detention centre in Oxfordshire where **he** applied for asylum claiming **he** had received death threats from criminals in his homeland.

But immigration officers who quizzed Kunowski failed to uncover his true identity.

2 April

The huge number of illegal immigrants who have sneaked into **Britain** is a major worry to voters because of the strains being placed on schools, hospitals and housing.

But for their first five years in power, all the Government did was make matters worse with a tangled mess of rules that didn't work.

For the past two years they have talked tough but achieved what? Police don't even arrest lorryloads of illegals – they ask them to make their own way to the **immigration** HQ in Croydon.

Naturally, they do a runner. That's why our cities and towns have an underclass of illegals prepared to work as slave labour.

2 April

Almost 10,000 forged passports and other ID documents were discovered at British ports in 2002, the latest year for which figures are available — an increase of 46 per cent on 2001.

But an unknown number of illegals with false papers make it into **Britain** each year. Forged passports are commonly used by terrorists to slip in and out of countries undetected. Two al-Qaeda terrorists, jailed in **Britain** in 2003, were discovered with hundreds of false travel documents.

5 April

Tony Blair is where **he** belongs – slap at the centre of the row over **immigration**.

Home Secretary David Blunkett has borne the brunt of public fury.

But his hands were tied by a Prime Minister who initially saw nothing wrong with a million new people moving into **Britain**.

6 April

Right now the PM is surrounded on three sides — by **illegal immigration**, Iraq and the EU.

His instinct is to choose his own turf and come out fighting.

So far he has always come out on top — from his first tussle with the unions over Clause Four to the Commons vote to go to war.

He has always believed he would win a vote on the EU once the battle was joined.

If his closest allies are right and he is planning a referendum this year, it is a courageous and welcome move.

Most people are deeply cynical about the European Union, the single currency and the constitution.

His best chance of persuading us to embrace even greater control from Brussels is by trying to scare us into it.

He will raise the spectre of lonely isolation outside a flourishing and powerful Europe.

He will warn that Britain will be left behind in a new Dark Age as the powerhouse EU economy surges on without us.

Mr Blair may have got away with that bluff a year or so ago.

People do not like change, even if they are unhappy with the way things are.

But life has been unkind recently to a Prime Minister who once seemed to defy the laws of political gravity.

He no longer walks on air.

Trust has gone, along with the billions taken in tax for failing public services.

The rest evaporated after Iraq and the calculated deceit over **immigration**.

7 April

AN axed Lottery fund gave just £28,000 to war veterans in its last few weeks—but dished out £1.4MILLION to refugee charities.

The Community Fund was wound up on Monday.

But the much-criticised group provoked more anger by snubbing World War II heroes in its last six weeks.

It handed a measly £28,114 to charities which help brave veterans of Dunkirk and Normandy.

But refugee associations received £1,383,664—almost 50 TIMES as much.

The largest single payout went to the Somali Healthy Mind Project...

19 April

Tony Blair is the most charismatic politician for decades.

But his supporters have grown weary of broken or only half-fulfilled promises on health, crime, **immigration** and transport.

26 April

The **Home** Office estimates that no more than 13,000 workers **will** come each year.

But others put the figure as high as 54,000 due to high unemployment in the East.

And there are fears services like schools and hospitals, already strained, **will** not be able to meet the new demands.

26 April

DON'T blame the people of Eastern Europe for heading for **Britain**.

The Government has put out the welcome mat. And to hard-up foreigners this looks the land of milk and honey.

But most **will** end up in unskilled low-paid jobs in the South East.

How **will** they afford to live? And how can schools and hospitals which are already at breaking point find room for them?.

29 April

THOUSANDS of migrants **will** be waved into **Britain** by officials completely unprepared for an invasion by new EU citizens, The Sun can reveal.

Many from countries like Poland, Slovakia, the Czech Republic and Hungary are expected to arrive seeking a better life in the wake of the EU expansion on May 1.

But immigration officials have decided AGAINST taking extra measures to cope with the expected rush.

29 April 39,750 **will** come to **UK**

NEARLY 40,000 people from the ten new EU countries **will** flock to **Britain** each year, experts said last night.

The **Home** Office reckons the figure **will** only be 13,000 immigrants annually.

But The Sun has compared estimates from all the major expert bodies to give a truly representative figure.

29 April

PANIC is a very strong word when used about a Government.

But it is impossible to disagree with Michael Howard's use of it **yesterday**.

It is blindingly obvious that on one of the most important issues affecting **Britain**, Tony Blair's lot have made a complete mess of it.

Strategy 2 (x 8)

2 April

The huge number of illegal immigrants who have sneaked into **Britain** is a major worry to voters because of the strains being placed on schools, hospitals and housing.

But for their first five years in power, all the Government did was make matters worse with a tangled mess of rules that didn't work.

For the past two years they have talked tough but achieved what?

Police don't even arrest lorryloads of illegals—they ask them to make their own way to the **immigration** HQ in Croydon.

Naturally, they do a runner. That's why our cities and towns have an underclass of illegals prepared to work as slave labour.

26 April

THOUSANDS of plane, train and bus seats had been snapped up last night by poor East Europeans seeking a better life in **Britain**.

They are free to come here when ten new countries join the EU this Saturday, May 1.

But it is feared that in some areas overstretched **UK** services like schools and hospitals **will** be unable to cope with the influx.

26 April

The **Home** Office estimates that no more than 13,000 workers **will** come each year.

But others put the figure as high as 54,000 due to high unemployment in the East.

And there are fears services like schools and hospitals, already strained, **will** not be able to meet the new demands.

27 April Stay Strong

THE SUN welcomes Tony Blair's assurance that **he will** keep a tight grip on **immigration** from Eastern Europe.

He is right to recognise that people are worried about what **will** happen after May 1. **But** the PM faces a huge challenge. Thousands are heading for **Britain** in search of a better life.

29 April

THE richer half of former Czechoslovakia has 10.4million people, including 750,000 gipsies.

The republic's turbulent history has now stabilised—and it has relatively low unemployment of 7.5 per cent.

But its young **WILL** be looking to taste a better standard of living.

And the **UK will** be top of their list for migration.

29 April

THOUSANDS of migrants **will** be waved into **Britain** by officials completely unprepared for an invasion by new EU citizens, The Sun can reveal.

Many from countries like Poland, Slovakia, the Czech Republic and Hungary are expected to arrive seeking a better life in the wake of the EU expansion on May 1.

But immigration officials have decided **AGAINST** taking extra measures to cope with the expected rush.

29 April 3 9,750 **will** come to **UK**

NEARLY 40,000 people from the ten new EU countries **will** flock to **Britain** each year, experts said last night.

The **Home** Office reckons the figure **will** only be 13,000 immigrants annually.

But The Sun has compared estimates from all the major expert bodies to give a truly representative figure.

29 April

UNLIKE many of the new entrants, Hungarians already enjoy a good standard of living.

The population of 10million also has relatively low unemployment.

Of those who leave, many **will** pick Austria and Germany – where there are strong historical ties.

But experts believe **Britain will** be a big attraction.

Strategy 3 (x 4)

2 April

The huge number of illegal immigrants who have sneaked into **Britain** is a major worry to voters because of the strains being placed on schools, hospitals and housing.

But for their first five years in power, all the Government did was make matters worse with a tangled mess of rules that didn't work.

For the past two years they have talked tough but achieved what?

Police don't even arrest lorryloads of illegals – they ask them to make their own way to the **immigration** HQ in Croydon.

Naturally, they do a runner. That's why our cities and towns have an underclass of illegals prepared to work as slave labour.

26 April

DON'T blame the people of Eastern Europe for heading for **Britain**.

The Government has put out the welcome mat. And to hard-up foreigners this looks the land of milk and honey.

But most **will** end up in unskilled low-paid jobs in the South East.

How **will** they afford to live? And how can schools and hospitals which are already at breaking point find room for them?

26 April

THOUSANDS of plane, train and bus seats had been snapped up last night by poor East Europeans seeking a better life in **Britain**.

They are free to come here when ten new countries join the EU this Saturday, May 1.

But it is feared that in some areas overstretched **UK** services like schools and hospitals **will** be unable to cope with the influx.

26 April

The **Home** Office estimates that no more than 13,000 workers **will** come each year.

But others put the figure as high as 54,000 due to high unemployment in the East.

And there are fears services like schools and hospitals, already strained, **will** not be able to meet the new demands.

Strategy 4 (x 4)

2 April

Depraved Kunowski, 48, raped 70 women and children in Poland then came to **Britain** and strangled 12-year-old Katerina Koneva.

He even got life-saving heart surgery on the NHS and claimed benefits.

But getting false ID documents can be shockingly easy, as we discovered.

A tip-off from a contact in the Polish underworld led us to Rozyckiego Mark

2 April

Almost 10,000 forged passports and other ID documents were discovered at British ports in 2002, the latest year for which figures are available—an increase of 46 per cent on 2001.

But an unknown number of illegals with false papers make it into **Britain** each year. Forged passports are commonly used by terrorists to slip in and out of countries undetected.

Two al-Qaeda terrorists, jailed in **Britain** in 2003, were discovered with hundreds of false travel documents.

2 April

In four weeks, Poland joins the EU and its citizens **will** have the right to live and work here.

But how many of those who arrive here bearing Polish passports **will** be comnen entering illegally?

2 April

Lieutenant Miroslaw Szacillo, 46, of the Polish Border Guards, assured The Sun: “Only the most serious criminals with big money can afford to buy false documents in Poland. We make stringent checks on our borders and we are getting more equipment to detect fake documents.”

But Poles who do get in and find work in **Britain will** qualify for a range of benefits including free healthcare, child tax credit, child benefit, working tax credit, housing benefit and council tax benefit.

Strategy 5 (x 2)

31 March

Many new arrivals **will** be good news for **Britain**.

But some **will** be up to no good. Gun-happy crime syndicates have already set up vicious vice and drug rackets.

Others want to do our nation harm. The Wall Street Journal says Islamic fanatics are using **immigration** as a “Trojan horse to expand jihad, or holy war”.

2 April also

Almost 10,000 forged passports and other ID documents were discovered at British ports in 2002, the latest year for which figures are available- an increase of 46 per cent on 2001.

But an unknown number of illegals with false papers make it into **Britain** each year. Forged passports are commonly used by terrorists to slip in and out of countries undetected.

Two al-Qaeda terrorists, jailed in **Britain** in 2003, were discovered with hundreds of false travel documents.