# A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts

Paul Baker and Tony McEnery

Dept. of Linguistics and English Language, Lancaster University

A corpus-based analysis of discourses of refugees and asylum seekers was carried out on data taken from a range of British newspapers and texts from the Office of the United Nations High Commissioner for Refugees website, both published in 2003. Concordances of the terms *refugee*(*s*) and *asylum seeker*(*s*) were examined and grouped along patterns which revealed linguistic traces of discourses. Discourses which framed refugees as packages, invaders, pests or water were found in newspaper texts, although there were also cases of negative discourses found in the UNHCR texts, revealing how difficult it is to disregard dominant discourses. Lexical choice was found to be an essential aspect of maintaining discourses of asylum seekers — collocational analyses of terms like *failed* vs. *rejected* revealed the underlying attitudes of the writers towards the subject.

**Keywords:** corpus, collocation, concordance, refugee, discourse

## Introduction

Critical social research is a form of academic enquiry which aims to achieve a better understanding of how societies work. Fairclough (2003: 202) defines a number of starting questions for critical social research, such as "how do existing societies provide people with the possibilities and resources for rich and fulfilling lives, how on the other hand do they deny people these possibilities and resources?" Critical discourse analysis (CDA) is a form of critical social research that can be applied to a range of texts in order to address these and other questions. Wodak and Meyer (2001: 96) refer to CDA as "discourse analysis with an attitude". Van Dijk (2001: 353) notes that CDA does not have

a unitary theoretical framework although there are conceptual and theoretical frameworks (e.g. Marxism) closely linked to CDA. Two basic questions for CDA are "How do (more) powerful groups control public discourse" and "How does such discourse control the mind and action of (less) powerful groups, and what are the social consequences of such control, such as social inequality?" (Van Dijk *ibid* 355).

While there is no single method of analysis in CDA, many researchers working in this area have tended to analyse single texts or small collections of texts. Fairclough (1989, 2003) offers a textual analysis framework for analysing texts, based around features such as metaphor, agency, nominalisation and pronoun use in addition to examining the modes of production and reception of texts. An alternative approach to CDA, using large bodies of texts (corpora) and computational algorithms to determine repetitive patterns of language use, has been carried out by a small number of researchers, e.g. Krishnamurthy (1996), Flowerdew (1997), Piper (2000), Fairclough (2000) and Teubert (2000). It is the intention of this paper to further the contribution of corpus linguistics to CDA, by showing how corpora can play an important role in critical social research, allowing researchers to objectively identify widespread patterns of naturally occurring language and rare but telling examples, both of which may be over-looked by a small-scale analysis. Such language patterns can help to illuminate the existence of discourses that may otherwise be unobserved. In this study we examine discourses surrounding the subjects of *refugees* and *asylum seekers*. The two terms are related, and sometimes used interchangeably, although for others they have different meanings or connotations.[1]

A discourse can be conceptualised as a "system of statements which constructs an object" (Parker 1992: 5). Discourse is further categorised by Burr (1995: 48) as

> "a set of meanings, metaphors, representations, images, stories, statements and so on that in some way together produce a particular version of events… Surrounding any one object, event, person etc., there may be a variety of different discourses, each with a different story to tell about the world, a different way of representing it to the world."

Discourses are not valid descriptions of people's 'beliefs' or 'opinions', and they cannot be taken as representing an inner, essential aspect of identity such as personality or attitude. Instead, they are connected to practices and structures that are lived out in society from day to day.

One way that researchers can be confident in their claims about the existence of discourses is to highlight "patterns of association — how lexical items

tend to co-occur — are built up over large amounts of text and are often un-available to intuition or conscious awareness. They can convey messages im-plicitly and even be at odds with an overt statement." (Hunston 2002: 109). In order to explore how refugees are constructed in news discourse we chose to build and use corpora of news texts. Clearly, the type of corpus one uses will have a strong bearing on the discourses which occur around any given sub-ject. A general corpus such as the 100 million words British National Corpus (henceforth BNC, Aston and Burnard 1998), which is made up of a variety of texts: novels, newspaper articles, spoken conversations etc. may present an equally diverse variety of discourses associated with refugees. However, it may be difficult, from such a general corpus, to make sense of cohesive positions or understand which discourses are viewed as hegemonic/mainstream and which are resistant or in the minority, particularly as frequency alone is unlikely to be an accurate predictor of the strength or importance of a discourse — the pro-ducers of texts and the ways that they are received are also important.

Therefore, we decided to build corpora which had texts that were created by a restricted set of authors, and where there was a higher degree of certainty that the discourses of refugees and asylum seekers contained within them were more likely to be received as hegemonic or widely accepted positions. Two sources of corpus data were located. The first was British newspaper articles published in 2003. As Fairclough (1989: 54) observes:

> "The hidden power of media discourse and the capacity of… power-holders to exercise this power depend on systematic tendencies in news reporting and other media activities. A single text on its own is quite insignificant: the effects of media power are cumulative, working through the repetition of particu-lar ways of handling causality and agency, particular ways of positioning the reader, and so forth".

Journalists are able to influence their readers by producing their own discours-es or helping to reshape existing ones. Such discourses are often shaped by cit-ing the opinions of those in powerful and privileged positions. Becker (1972: xx) calls this the 'hierarchy of credibility', whereby powerful people will come to have their opinions accepted because they are assumed to have access to more accurate information on particular topics than everyone else. Hall et al. (1978: 58) say "The result of this structured preference given in the media to the opinions of the powerful is that these 'spokesmen' become what we call the primary definers of topics."

The second type of corpus data consisted of articles on refugees and asylum seekers published in 2003 by The Office of the United Nations High

Commissioner for Refugees. This Office was created in 1950 by the United Nations General Assembly and is mandated to lead and co-ordinate international action to protect refugees and resolve refugee problems worldwide. It is estimated that it has helped about 50 million people, and to date has a current staff of around 5,000 people in more than 120 countries. Therefore, both newspapers and the UNHCR can be said to wield different sorts of power in terms of their ability to influence discourses concerning refugees. Newspapers, which are widely read on a daily basis, can help to shape/reflect public opinion, whereas the UNHCR, as a source of relatively independent expertise, is directly involved in the welfare of refugees and has a decades-old proven track record of success. The way that it chooses to construct refugees and related subjects means that it is likely to be influential on an international level.

The news and UNHCR corpora do not contain texts where refugees and asylum seekers as subjects are self-constructing; rather, they have their identities and discourses surrounding these identities constructed for them by more powerful spokespeople. Van Dijk (1996: 91–4) points out that minority groups are frequent topics of political talk and text, but have very little control over their representations in political discourse. Lack of access to journalists means that minority speakers tend to be less quoted than majority speakers (van Dijk 1991), and those who are quoted tend either to be chosen because they represent the views of the majority, or because they are extremists who are quoted in order to facilitate an attack upon the minority (Downing 1980).

In choosing to analyse the two corpora, a further dimension is introduced into the analysis: the extent of difference. What are the most significant differences and similarities in the ways that these subjects are constructed discursively, and how will this relate to the roles of each of the producers of the different texts and their intended audiences? In addition to the corpora of newspaper and UNHCR articles, we also chose to employ a third corpus of more general English for reference — the BNC. While the BNC is not being used in this study to directly analyse discourses of refugees, it is referred to instead because it can reveal normative patterns of language use which can then be compared against the findings in the two more specific corpora. For example, in one of the news texts, an article refers to "swelling the numbers of refugees", we can use the BNC to see how a word like *swelling* is used in general British English, what its common patterns of association or collocates are, and how such a word therefore helps to construct refugees in a certain way.

## Data and Method

The British newspaper texts[2] (referred to hereafter as the News Corpus) were collected from an internet archive called Newsbank,[3] the criteria for selecting news articles being that the article had to contain the words *refugee*(*s*) or the phrases *asylum seeker*(*s*) and had to have been published in 2003. The second corpus (referred to as the UNHCR Corpus) contained every text published by the UNHCR's website[4] on refugees in 2003. Although there were roughly equal occurrences of the terms *refugee(s)* and *asylum seeker(s)* in the News corpus, the UNHCR corpus had a much higher proportion of references to refugees. This is because of the UNHCR's website is focussed around refugees, although there were still a number of references to asylum seekers on the website, suggesting that the two identities share a common ground.

**Table 1.** Size of corpora and frequency of target words.

|  | News texts | UN texts |
|---|---|---|
| Size of corpus (words) | 76,205 refugee(s) | 265,224 overall |
|  | 62,299 asylum seeker(s) |  |
| *refugee* | 53 | 1741 |
| *refugees* | 87 | 2954 |
| *asylum seeker* | 10 | 19 |
| *asylum seekers* | 131 | 206 |

In language there may be many ways of referring to the same idea or subject, either by using different grammatical constructions, or words or phrases which contain similar meanings. For example, the word-form *cases* frequently co-occurs (or colligates) with the grammatical category of 'quantifier', in phrases such as *in some cases*, *in many cases*. And the notion of collocation can be extended even further, to include *semantic preference*. This is a term similar to Louw's (1993) and Sinclair's (1991) concept of *semantic prosody*. Semantic preference is according to Stubbs (2001: 65): "the relation, not between individual words, but between a lemma[5] or word-form and a set of semantically related words". For example, in the British National Corpus the word *rising* co-occurs with words to do with work and money: e.g. *incomes*, *prices*, *wages*, *earnings*, *unemployment* etc. Semantic preference also occurs with multi-word units. For example *glass of* co-occurs with a lexical set of words denoting 'drinks': e.g. *sherry*, *lemonade*, *water*, *champagne*, *milk* etc. Semantic preference is therefore related to the concepts of collocation and colligation, but focuses on a lexical

set of semantic categories rather than a single word or a related set of grammatical words.

Semantic preference is related to the concept of *discourse prosody*, although the difference between the two is not always clear-cut. Stubbs (2001: 65) says it is partly a question of how open-ended the list of collocates is. So it may be possible to list all of the words for 'drinks', indicating a semantic preference, but a more open-ended category such as 'unpleasant things' might be seen as a discourse prosody. Stubbs (2001: 88) later notes that even a category of semantic preference will be open-ended, but will contain frequent and typical members. By examining concordances of the terms *refugee* and *refugees*, it should be possible to uncover semantic and discourse prosodies within the corpora in question.

## Refugees

### *The News Corpus*

Concordances (tables showing all of the examples of a search term in the context that it appears in) of the words *refugee* (53 occurrences) and *refugees* (87) were carried out on the 76,205 word News Corpus. In addition, concordance searches of the words *they* and *them* uncovered anaphoric references to refugees which were also included in the analysis. In order to uncover linguistic patterns surrounding the search words, the concordances were sorted alphabetically using the words directly to the left and right of the search terms, and the descriptive clauses and phrases which were used to refer to the target words were then grouped into categories of similarity.

### *Quantification*

An initial analysis reveals that refugees are commonly described in terms of where they are from (e.g. Sierra Leone, Bangladesh, Afghanistan, Iran, Kosovo, Algeria etc.), where they currently are (France, Sangatte, the Belgian border etc.) or where they are going. A smaller set of terms describe refugees in terms of the circumstances which created them, although these words are vague (*economic*, *political*, *war*, *wartime*).

One of the other most common ways of describing refugees in the News Corpus is by providing a pre-modifying quantification (Table 2). This is often given as a rough estimate by reporters — e.g. *tens of thousands* or *up to 100 refugees*. Sometimes numbers are described as growing: *more and more refugees*.

**Table 2.**  Quantification of refugees in the Nws Corpus.

| | | |
|---:|:---:|:---|
| Pakistan hosts **3m** | refugees | and Iran 2.5m, while a further 1.5m a |
| She was concerned for the **thousands of** Afghan | refugees | and she wanted to spotlight their terr |
| **Tens of thousands of** Afghan | refugees | are streaming home in the vanguard o |
| - 50,000 dead **and more than 300,000** internal | refugees | because of Pakistan-sponsored te |
| Now a fascinating documentary has traced **five** | refugee | children who were rescued from p |
| in Afghanistan, which is **swelling the number of** | refugees | daily, and tightened security by ferry |
| ill be more grieving widows and children, **more** | refugees | fleeing in terror and dying in agony. |
| Our coverage of the deaths of **eight** | refugees | found in a container near Wexford |
| **Four million** | refugees | have fled in a quarter of a century of |
| terror network rages, **two** Afghan | refugees | in northern Iran seek solace in the scr |
| was attacked and beaten up by **a mob of** Afghan | refugees | in Pakistan yesterday. |
| **EIGHT** | refugees | , including two children, were found |
| "And the tragedy of **22 million** | refugees | is another deep wound in our world - |
| juries after being pelted with stones by **up to 100** | refugees | . Last night Fisk was recovering in hi |
| TV and video equipment to help Sighthill's **1,200** | refugees | learn English. |
| shawar, home to **300,000** Afghan | refugees | . On the ground floor two gynaeco |
| months before the bulk of the Shomali's **200,000** | refugees | return: "The major obstacle is the pre |
| **Eight** | Refugee | Stowaways Die In Lorry |
| in March after **19** Kurdish | refugees | were found in the back of his lorry. |
| end to the problem is in sight as **more and more** | refugees | are delivered to northern France by t |

In some cases, these types of quantification suggest that the volume of refugees is troublesome. In the example below, the reference to the large number (*a mob*, *up to 100*) of refugees in the article serves to enhance their danger.

> "BRITISH journalist Robert Fisk was attacked and beaten up by a mob of Afghan refugees in Pakistan yesterday… He suffered head, face and hand injuries after being pelted with stones by up to 100 refugees."

Many of the attempts to quantify refugees suggest an underlying discourse concerning alarm over their growing numbers.

> "Mr Endres said Kabul risked straining under the weight of refugees in transit."

> "The camp… is currently filled far beyond capacity. Because of the upheaval in Afghanistan, which is swelling the number of refugees daily…"

In the first example, refugees are likened to a dangerous mass or heavy load (note the use of the words *straining* and *weight*) while in the second, refugees are constructed as liquid (*filled beyond capacity*, *swelling*). Such a feature may not be peculiar to British newspaper texts; Reisigl and Wodak (2001: 59) note

that the description of immigrants and the effects of immigration in terms of negative metaphors is a common feature of German and Austrian discourses. They list nineteen different types of metaphors, including plants, fire, blood, disease and food. We look at some of the metaphors in the News and UNHCR Corpora in the following sections.

*Movement*

A set of words which describe refugees are concerned with their movement (Table 3), often using verb phrases which suggest a range of evaluative responses which construct refugees as victims or a collective group undergoing suffering: e.g. *fleeing refugees*, *refugees trudge aimlessly, hunched against a biting wind*, *roads heave with refugees*. As with the *filled far beyond capacity* example noted earlier, a number of movement metaphors liken refugees to water in some way: *swelling the numbers of refugees*, *the flood of refugees*, *refugees are streaming home*, *refugees are streaming back to their homes*, *overflowing refugee camps*.

Streaming collocates[6] (or significantly often co-occurs) in the BNC with the words *tears*, *blood*, *sweat*, *water* and *rain* and often occurs in a negative context, e.g. "tears streaming down his face". *Overflowing* collocates strongly in the BNC with *leaking* and *water*. *Swelling* (462 occurrences in the BNC) collocates most strongly with words which suggest medical contexts: e.g. *redness*, *bruising*, *pain*, *chest*. However, *swell* which occurs more often (566 times in the BNC) collocates with words connected to water: *waves*, *Atlantic*, *ocean*, *sea*, *water*. The phrase *flood of* collocates most strongly with *refugees* in the BNC, with *tears* and *immigrants* occurring second and third respectively. Again, *flood* is connected to water and to tragedy.

In a sense then, refugees are constructed as a 'natural disaster' like a flood, which is difficult to control as it has no sense of its own agency. Again, this is not peculiar to British newspaper texts; similar water metaphors were also found by Refaie (2001) in an analysis of Austrian newspaper articles about Kurdish asylum seekers in Italy.

Phrases such as "trudge aimlessly" help to construct refugees as having no real understanding of their situation or what motivates them. Consider the phrase "desultory groups of refugees" (Table 3, line 8). *Desultory* is a fairly rare word in general British English, occurring only 103 times in the BNC, and only collocating with two lexical words, *fashion* and *conversation*. Again, the word implies lack of motivation or pattern: e.g. "Because the session is informal it is liable to fragment into a desultory conversation with no clear direction."

**Table 3.**  Movement of refugees in the News Corpus.

| | | |
|---|---|---|
| and the Holy Family **fleeing** as | refugees | to Egypt. Jesus relives exodus and ex |
| l troubles of 1950, Shaw was engaged in **carrying** | refugees | from East Pakistan to Calcutta and t |
| They can also help to **attract** | refugees | back to their homes and thousands of |
| rgest oilfield, **and provide a haven for** | refugees | and defecting soldiers. |
| xt night the Taliban bombed Talin Chee, **forcing** | refugees | into another village and finally the ha |
| heated and broke down close to a village **housing** | refugees | who had **fled** across the border fro |
| which is **swelling** the number of | refugees | daily, and tightened security by ferry |
| and tinsel, desultory groups of | refugees | **trudge aimlessly**, hunched against a |
| alert to deal with the **flood of** | refugees | . Border residents also left their hom |
| h huge brightly painted trucks bearing **returning** | refugees | sitting on tottering piles of their pos |
| of Turkish gangsters who have been **transporting** | refugees | from Ankara and Istanbul via Zeebru |
| mbed a bridge and a tractor convoy **packed with** | refugees | . This time it is thought the Americ |
| nother day. The roads **heave with** | refugees | , the taxis are full and the drivers still |
| , the Holy Family was forced to **take** Him as a | refugee | to Egypt. |
| ed mullah had **fled** across the Pakistan border as a | refugee | in the company of his sen |
| N Pushtuns who were **stranded** in nearby | refugee | camps when the Taliban fled may be |
| starvation or disease in the **overflowing** | refugee | camps, then clearly the number of d |
| e problem is in sight as more and more | refugees | **are delivered to** northern France by |
| Tens of thousands of Afghan | refugees | **are streaming back** to their homes a |
| Tens of thousands of Afghan | refugees | **are streaming home** in the vanguard |
| hospital hit, or children crippled by landmines, or | refugees | **driven** from their homes by military |
| will be more grieving widows and children, more | refugees | **fleeing** in terror and dying in agony. |
| Four million | refugees | have **fled** in a quarter of a century of |
| n at Guantanamo Bay, where Cuban and Haitian | refugees | have previously **been held**, and is |
| | Refugees | … **have returned** to rubble |
| es said Kabul risked straining under the weight of | refugees | **in transit**. |
| ing a mercy mission to **bring** the young European | refugees | into homes all over the country. |
| | Refugees | **Return**. Tens of thousands of Afgh |
| months before the bulk of the Shomali's 200,000 | refugees | **return**: "The major obstacle is the |
| , and those shots of | refugees | **smuggling themselves** on to Eurost |
| s mother's necklace and **trafficked in** our wartime | refugees | . Yet women fell for him. On reflecti |

The movement of refugees is constructed as an elemental force which is difficult to predict and has no sense of control. If refugees are likened to the movement of water, then they are dehumanised and become something that requires control in order to prevent disaster to others (e.g. non-refugees).

As well as describing the movement of refugees as being almost random, a number of phrases focus on movement in terms of large quantities: e.g. *roads heave with refugees*, *packed with refugees*. The word *packed* collocates with words suggesting quantity or places containing large numbers of people in

the BNC — *tightly*, *densely*, *closely*, *tight*, *crowd*, *courtroom*, *cinemas*, while the verb lemma ʜᴇᴀᴠᴇ[7] collocates with words which suggest weight — *bulk*, *broad*, *deep*, *great* as well as the expression of emotional burdens, e.g. "Uncle Wafter heaved a sigh and slumped back in the chair, his hand covering his eyes" (example taken from the BNC).

Closely related to the notion of packed refugees are another set of movement descriptors connected to refugees, associated with the transportation of objects and goods. Refugees are *delivered*, *transported*, *carried*, *trafficked* and *smuggled*.

> "For the locals, no end to the problem is in sight as more and more refugees are delivered to northern France by traffickers."

The highest collocate of ᴅᴇʟɪᴠᴇʀ in the BNC is *goods*. *Goods* also occurs as a strong collocate of *transport*, along with *cattle*, *supplies* and *materials*. *Carry* occurs much more frequently and has a wider range of collocates, many of which are object-based nouns: *placards*, *firearms*, *suitcases*, *torches*. Finally, *smuggle* collocates most strongly with the following lexical words: *cocaine*, *heroin*, *drugs*, *drug* and *arms* while ᴛʀᴀꜰꜰɪᴄᴋ collocates with *narcotics*, *drugs* and *arms*.

Therefore, as well as being described as an elemental force that cannot be reasoned with (water), refugees are also constructed in terms of metaphors and connotational verbs which construct them as transported goods, particularly illegal substances — again, as a token of their dehumanisation.

*Tragedy*

Another discourse of refugees is to construct them as 'tragic' (see Table 4). This involves using phrases such as the *plight of*, *despair of* and *tragedy of*. *Plight* collocates in the BNC with groups such as *homeless*, *refugees*, *blacks*, *women*, *unemployed* and *children* — all identities that could be constructed as powerless in different ways. However, three of the four strongest collocates of *plight* in the BNC are *highlighting*, *highlight* and *highlighted*, in sentences such as "It is 10 days since our shock issue highlighting the plight of the starving in Somalia". *Plight* is therefore often connected to attempts to heighten awareness about a group which is oppressed or unfortunate in some way.

Other words which suggest tragedy in connection with the refugee data are *scrounge*, *beg*, *tedious*, *tottering*, *solace* and *stricken*. Refugees are reported as starving, dying while locked in containers, seeking solace in religion, queuing for food and being attacked by soldiers.

**Table 4.** Tragic circumstances of refugees in the News Corpus.

| | |
|---|---|
| After seeing the **plight of** these refugee | children, I could not do any |
| king undercover to report on the **plight of** Afghan refugees . But her f | |
| "And the **tragedy of** 22 million refugees is another deep wound in our world - | |
| s with ; cultural vandalism and the **despair of** its refugees , reports Peter Foster in Bamiyan | |
| **wounded** early today as Israeli tanks moved on a refugee | camp in the Gaza Strip. |
| Israel **fires** on refugee | camp |
| Israelis **attacked** the Rafah refugee | camp — thought to be a stronghold of |
| For in the **poverty-stricken** Palestinian refugee | camps of Lebanon, giving your life t |
| ghans, the ones who haven't made it to the refugee | camps and **have no relatives** in |
| Cambodian refugees refugees continue to **starve.** | |
| widows and children, more refugees fleeing in terror and **dying** in agony. | |
| Our coverage of the **deaths** of eight refugees found in a container near Wexford | |
| None of the refugees **has enough calcium.** | |
| two Afghan refugees in northern Iran **seek solace** in the sc | |
| EIGHT refugees … were found **dead** | |
| huge brightly painted trucks bearing returning refugees on **tottering piles** of their possession | |
| THE net is closing on the gang that sent eight refugees to their **deaths** in a lorry container, it | |
| Unusually, the refugees were **not advised** to carry mobile tele | |
| Eight Refugee Stowaways **Die** In Lorry | |
| It's here you find refugee | camps that have turned into **shabby** c |
| " They | **have so little medical assistance** tha |
| e body, not spirit, that must take precedence as they | **scrounge** for scrap metal and **beg** for |
| Instead they | live a **tedious** routine, queuing up to |
| Not only were they | **hungry**, they were **freezing**. |
| inquests into the deaths of the eight recorded that they | **died** at sea between Belgium and Ir |

## Official attempts to help

Related to the presentation of refugees as tragic victims is another set of collocates which are more concerned with external efforts to help them (Table 5). These involve phrases such as *refugee action*, *refugee service*, *refugee agency*, describe official bodies involved in running organisations, and discuss attempts to enable refugees to 'integrate into society', particularly by learning the language of their host country or by going to school. The grammatical pattern *X for refugees* is a relatively common example of this 'helping' discourse trace. In addition, terms such as *shelter*, *help*, *concern*, *mercy* and *rescue* contribute to the construction of this discourse.

**Table 5.** Official attempts to help in the News Corpus.

| | |
|---|---|
| **helped** the UNHCR return refugees | in East Timor; |
| Margaret Curran welcomes new funding to **help** refugees | integrate into society. |
| Last week we looked at Refugee | **Action**. Here, Belinda Beresford in |
| id expansion was straining resources, said the UN refugee | **agency**. |
| Now a fascinating documentary has traced five refugee | children who were **rescued** from |
| , who is chief executive of the Northern England Refugee | **Service**, said: "If she went into Afg |
| Bob Dylan and Eric Clapton performed **in aid of** refugees | of Bangladesh in 1971. |
| "She was **concerned for** the thousands of Afghan refugees | and she wanted |
| ing a **mercy mission** to bring the young European refugees | into homes all over the coun |
| On the first floor is the **school for** refugees | and orphans. |
| tywest will be turned into a **reception centre for** refugees | and the homeless. |
| hildren in India and **humanitarian assistance for** refugees | in Afghanistan. |
| hief of mission of the **UN high commissioner for** Refugees | . Agencies are racing to provide bla |
| agreements with the **UN High Commissioner for** Refugees | on the role of the peacekeeping forc |
| He was filming for **Comic Relief** in a refugee | camp. |
| the French close the **Red Cross** refugee | camp at Sangatte. However, they op |
| into the love which feeds the hungry, **shelters** the refugee | , and confronts violence and power |

## Crime and nuisance

However, a less common discourse constructs refugees as being connected to crime and as a nuisance (Table 6). In this case, their presence 'pushes down' house prices or causes a steep rise in petty crime. Such fears invoke a more general discourse of capitalism, whereby refugees are seen as a threat to the

**Table 6.** Crime and nuisance in the News Corpus.

| | |
|---|---|
| Newsman Fisk **stoned by** Afghan refugees | |
| uries after being **pelted with stones by** up to 100 refugees | |
| **ufacture the explosive** in laboratories hidden in refugee | camps. Mr Reid is said to have |
| an unknown number of Somali and Djibouti refugees | have been **arrested** in the past week. |
| There have been **complaints** the refugees | have caused a steep rise in **petty crime and pushed down house prices** |
| Four of the refugees | who led the Chunnel charge were **jailed for four months** |
| Palestinians in the Gaza and West Bank refugee | camps **resent their lavish lives**. |
| Israelis attacked the Rafah refugee | camp — thought to be a **stronghold of militants** |
| ushtuns who were stranded in nearby refugee | camps when the Taliban fled may be **hiding arms and plotting revenge**. |
| A RED CROSS boss kept secret a refugee | **plot to storm the Channel Tunnel** |
| e in petty crime and pushed down house prices. " They | **are a pest**," says a woman in th |

capitalist way of life by reducing the value of property. As one woman is reported as saying: "they are a pest".

Refugee camps are also reported as hiding grounds for extremists or militants, and refugees are also involved in plans to enter countries illegally — e.g. "to storm the channel tunnel". The use of *storm* as a verb suggests attempts to conquer — common collocates of the verb lemma STORM in the BNC include *troops*, *castle*, *victory* and *army*. In this sense, refugees are invaders.

A final, rare use of *refugee* in the News Corpus is more metaphorical, where it is used to describe people who look like someone or something else:

> "Coming on like some refugee from the Ricky Lake Show, burly Fred spent much of the programme successfully convincing Sandra that he thought the first marriage had been annulled, and moaning to the camera that he 'shouldn't be put in jail for falling in love'."

> "Owen may look like a refugee from a Hovis ad, but as Sven-Goran Eriksson said: 'He's very cold when he gets a chance and he's very quick.'"

> "Last week I watched some do-it-yourself programme where a couple of refugee presenters from the makeover toolbox showed you how to have a kitsch Christmas."

These three cases are not from newspaper articles which are concerned with actual refugees. Instead, the phrase (*like*) [determiner] *refugee from…* is used to allude to a person's similarity to something else. However, this is a construction which contains an implicitly negative evaluation — the fact that such people are described as refugees at all accesses an existing negative discourse of actual refugees, but it also implies that they are not viewed as possessing the identity they are supposed to have, possibly because they weren't competent at it, or because they look as if they should be something else. So Michael Owen (in the second example) is viewed not so much as a footballer, but a "refugee from a Hovis ad(vert)". Finally, the identity they are supposed to resemble is constructed negatively — so the phrase implies that the person isn't even competent to perform a stigmatised identity properly, and instead is a refugee from it.

## The UNHCR Corpus

An analysis of left hand collocates reveals some similarities and differences in comparison to the discourses of refugees in the news data. For example, attempts to quantify the number of refugees are also frequent in the UNHCR

Corpus — numerals occur 846 times as left hand collocates of *refugees* (28% of all cases of the word's occurrence). Other quantification phrases include *many refugees* (26), *number*(*s*) *of refugees* (37), *more refugees* (21) and *some refugees* (16).

However, compared to the News Corpus, there are many more references to the nationalities of refugees both in terms of quantity and range: e.g. Afghan 176, Iraqi 159, Leonean 125, Liberian 120, Angolan 87, Sudanese 68, Congolese 63, Rwandan 48, Ivorian 41, African 30, Burundian 27, Iranian 27, Eritrean 23, Somali 19, Palestinian 18. There is also more of a sense that the identity of refugee is something which is temporary — terms like *former* (15) and *new* (23) are more common left hand collocates.

A high number of L1 collocates are movement words. Common L1 collocates include *return*, *returning*, *transfer*, *move*, *displaced* and *influx*. We will examine two of these words, *displaced* and *influx*, in more detail to see how they contribute towards constructions of refugees in the UNHCR Corpus.

The word *displaced* occurs 418 times in the UNHCR Corpus. It occurs most often in an adjectival form (*displaced refugees*), but can also be a verb (*others were displaced more recently*). *Displaced* is often used in agentless sentences, obscuring information about who or what has displaced the refugees. In only a few cases is some form of agency attributed — *displaced by* only occurs 16 times, and when it does occur, the agent is never a person or group but a reference to *conflict*, *war*, *attack* or *fighting* (see Table 7). The term *fighting*,

**Table 7.** Concordance of *displaced by* in the UNHCR Corpus.

| | |
|---|---|
| reportedly arrived in this area since April, displaced by | an ongoing **conflict** in the Darfur regi |
| ing reports that hundreds of people were displaced by | **armed clashes** near the border with |
| Tens of thousands displaced by | **fighting** near Monrovia |
| UNHCR concerned about hundreds displaced by | cross-border **attack** in Panama |
| Feature: displaced by | **drought**, Afghans now rush back to |
| 000 to Congolese nationals who had been displaced by | ethnic **fighting** in the Bunia area of It |
| We hope that refugees and other people displaced by | **fighting** in the area will now attempt t |
| nd other West African nationals had been displaced by | **fighting** between government and MO |
| i continues to transfer Congolese refugees displaced by | October's **conflicts** from border sites |
| stance to hundreds of thousands of people displaced by | recent **fighting** in the strife-torn count |
| in Modena, Italy, to raise funds for people displaced by | the Iraq **conflict**. |
| HCR has been trying to persuade refugees displaced by | the recent **fighting** to go back to their |
| , where they had sought refuge after being displaced by | the **fighting** in recent weeks. |
| from 40,000 to 60,000 people — has been displaced by | the **fighting** between the Tutsi-led ar |
| ely south of Grand Geddeh — before being displaced by | the latest **fighting**, and moving on to f |
| own is a prime example of how Angolans displaced by | the 27-year civil **war** have been anxio |

which occurs in nine of these cases, is often framed as a noun rather than a verb, e.g. *the fighting*, *the latest fighting*, suggesting that it is the fighting which takes the place of a human agent, obscuring the apportionment of blame to a single person or group.

In the BNC, *displaced* mainly collocates with words relating to refugees (*persons*, *people*, *war*, *fighting*, *population*, *political*), but also has another set of collocates which relate it to science or ecology: "Ganglion cells representing the fovea itself are laterally displaced in the retina", "One tried to use Archimedes' flotation principle (weight of water displaced = weight of body)." *Displaced* occurs most often in the Natural and Pure Sciences sub-section of the BNC, and is second most often found in the World Affairs sub-section (see Table 8), suggesting that it is a term associated mainly with scientific forms of discourse.

*Influx*, which occurs 65 times in the UNHCR Corpus, also collocates strongly in the BNC with *refugees* (its strongest collocate), as well as *immigrants*, *tourists*, *visitors*, *foreign*, *workers* and *people* — suggesting that it is often used when discussing human movements on a large scale. However, it also collocates with the terms *calcium* and *sediment*. As with *displaced*, *influx* occurs most often in the Natural and Pure Sciences subsection of the BNC, with World Affairs coming second (Table 8). Other collocates of *influx* in the BNC are *sudden* and *cope* — suggesting that an influx can be unexpected and problematic: "Ian Black in Tel Aviv asks whether Israel, willing but hardly prepared for the influx, can cope" (example taken from the BNC). The fact that *displaced* and *influx* are used in scientific texts, often to refer to the movement of liquids or minerals, suggests that they are linked to the metaphors of refugees as water used in the News Corpus. However, a word like *displaced* has less emotive value than words like *streaming*, *flood* and *overflowing* — the latter words suggest

**Table 8.** Relative frequencies (per million words) of *displaced* and *influx* in genres of the BNC.

| Genre | displaced | influx |
|---|---|---|
| Natural and pure sciences | 17.18 | 12.16 |
| World affairs | 9.69 | 9.34 |
| Belief and thought | 8.98 | 2.33 |
| Social science | 8.2 | 3.88 |
| Applied science | 7.18 | 2.67 |
| Arts | 6.59 | 1.69 |
| Commerce and finance | 5.79 | 3.86 |
| Leisure | 2.95 | 3.78 |
| Imaginative | 2.2 | 0.61 |

rapid movement that is out of control or unwanted. Therefore, it could be argued that constructing refugees in terms of water metaphors is not intrinsically negative in itself. The value judgement is dependant on how the water is then described. Interestingly, the UNHCR Corpus is not without a few of the more emotive water metaphors. The phrase *flow of* occurs five times in relation to refugees, *outflow* occurs twice, *flood of* occurs twice, *stream* occurs four times, *swell* occurs once and *wave*(*s*) *of* occurs five times. Refaie (2001: 366) argues that the water metaphor has become 'naturalised' in that it has become the accepted way for a large number of politicians, journalists and editors to talk about refugees and is rarely challenged. Its presence in the UNHCR texts supports this finding.

HOST as a verb or noun occurs 227 times in the UNHCR texts, in relation to places, camps or countries where refugees are staying, e.g. "the High Commissioner said that it continues to host the largest number of refugees in Africa" (see Table 9). In the BNC, *host* collocates strongly with terms like *dinner*, *Olympics*, *conference*, *reception* and *festival* and is used in contexts which suggest a positive and friendly form of exchange. However, *host* has a secondary meaning linked to science — particularly involving the collocate *parasite*: "Genetic changes in both host and parasite were responsible for the evolved differences in virulence" (example from the BNC). So while the main discoursal meaning of *host* in relation to refugees is positively loaded, there is an underlying negative meaning potentially associated with the word.

Many of the left hand collocates of *refugee*(*s*) in the UNHCR data are related to official attempts to help: *solutions* (occurs 25 times in a window up to five words to the left of *refugee*(*s*)), also *assistance* (46), *aid* (32), *needs* (23), *programmes* (18), *support* (18), *protect* (13), *helping* (15) and *efforts* (14). The

**Table 9.** Sample concordance of HOST in the UNHCR Corpus.

| | |
|---|---|
| Iran and Pakistan have hosted | millions of Afghan refugees since 1980 |
| hen the latest war in Iraq started, Iran was hosting | 204,000 Iraqi refugees who had been r |
| The agency said it had asked countries hosting | Iraqis to extend by at least one month a |
| jubilant former refugees. Kyrgyzstan hosts | some 7,000 refugees, including more th |
| ent chose Mohammad Kheil, a camp now hosting | 37,000 refugees that was opened for t |
| presented by large refugee populations in host | countries,&quot; said Diagne. |
| r, at the onset of last year. Tanzania still hosted | an estimated 24,000 Rwandans, in addit |
| Lubbers leads Eritreans home; thanks hosts | in Sudan. UNHCR News Stories |
| Bandajuma camp hosts | some 6,500 refugees, of whom about 3, |
| s to assist the local population in refugee- hosting | areas, both to compensate for under-de |

phrase *World Refugee Day* occurs 38 times in the UNHCR Corpus, whereas *refugee agency* occurs 690 times.

Discourses of refugees as tragic or nuisances are not as frequent as in the News Corpus, although the 'tragic' discourse still exists. The word *fled* occurs 228 times, while more emotive words such as *plight*, *tragedy* and *despair* occur 23, 6 and 3 times respectively. Refugees are described as *vulnerable* 10 times.

Words which construct refugees as packages such as *deliver*, *smuggle*, *carry* and *transport* also have relatively low frequencies in the UNHCR Corpus. Instead, we find a number of terms involved with returning refugees. The phrase *returning refugees* occurs 26 times while *return of refugees* occurs 9 times. In addition, the terms *reintegration* and *repatriation* occur 83 and 471 times in the corpus respectively.

## Asylum seekers

### The UNHCR Corpus

In a number of ways, the representation of asylum seekers is similar to that of refugees in the UNHCR Corpus. Both subjects are sometimes linked together with the conjunction *and*. Of the 202 cases of *asylum seekers* in the corpus, 33 of them are specifically linked in this way to refugees — of which 27 feature the word *refugee* first.

As with refugees, two of the strongest ways that asylum seekers are referred to are in terms of their nationality (e.g. Afghan 4, African 1, Acehnese 3, Angolan, Arab 1, Colombian 1, Burundian 1, Congolese 1, Iraqi 17, North Korean 3, Liberian 2, Russian 1) and their numbers. There are 27 cases of exact numbers of asylum seekers being given in an article, but also many cases where numbers are left vague e.g. *thousands of* (3), *majority of* (3), n*umber(s) of* (12). There are also several cases of quantification which relate to changes in numbers, e.g. "largest number", "a mere 13 extra", "fewer people applied for", "a 32 per cent drop", "downward trend", "increase in". On the whole, this information is presented without the use of metaphor or subjective language, although there are a few cases where this appears to be the case, e.g.: "10,000 people who thronged a transit centre in a bid to apply for…"

The strongest collocate of *asylum* is *seekers*, of the 432 cases of *asylum*, 202 of them occur in the phrase *asylum seekers*. However, we also find the singular form: *asylum seeker* (23) and a number of verb forms: *seeking asylum* (10), *seek asylum* (6) and *sought asylum* (7). SEEK collocates in the BNC

with *advice*, *redress*, *asylum*, *refuge*, *solace*, *revenge*, *permission*, and *re-election*. On the whole, SEEK has a discourse prosody for positive events or situations. On the other hand, *asylum* collocates in the BNC with *seekers*, *seeker*, *seeking*, *sought*, *immigration* but also *mental*, *lunatic* (as in a lunatic asylum), *bogus* (as in *bogus asylum seekers*), *genuine*, *applications*, *applicants* and *granted*. Therefore, although the strongest collocational meaning of *asylum* is to do with its current meaning, it also has another much more negative meaning associated with mental illness and incarceration. The pairing of *asylum* with a more positive term *seeker* is perhaps one way of removing the existing negative stigma associated with the word.

However, *asylum* also collocates in the UNHCR Corpus with a group of words which are linked to processes rather than people: *system(s)* 16, *procedure(s)* 16, *policy(s)* 5, *directive* 1, *process(ing)* 1, *law* 6, *legislation* 4.

As well as SEEK, the verb APPLY has a strong collocational relationship with *asylum*. The phrase *apply/applied/applying for asylum* occurs 8 times in the corpus, and as noted above, the nominalised *asylum applications* occurs 15 times, while *asylum applicants* occurs 3 times. In the BNC, APPLY as a verb collocates strongly with *rule*, *principle*, *criteria*, *regulations*, *considerations* and *provisions* — words which suggest bureaucratic processes. The phrase *claiming/claimed asylum* occurs 6 times, along with 15 cases of *asylum claims*. CLAIM in the BNC collocates with *damages*, *compensation*, *falsely*, *entitled*, *dismissal*, *benefit*, *repayment* and *unfairly*. The term *claim* is much more strongly linked to the rewards of claiming and whether they are viewed as deserving or not. In the UN corpus, claims can be *lodged* (by asylum seekers or people working on their behalf — this occurs 5 times), and they can be *assessed* (occurring 4 times). LODGE as a verb collocates in the BNC with *objections*, *complaint*, *appeal* and *protest*, suggesting a somewhat negative semantic prosody.

The question of whether asylum seekers are 'deserving' is something which arises in the corpus. For example, *the right* to asylum or to seek it is referred to 4 times, while the terms *genuine* and *bona fide* occur 5 and 1 times respectively with *asylum*. The term *access to* occurs twice with *asylum*. In addition, the phrase *reasonable to seek asylum* occurs in the UNHCR corpus, while *want to seek asylum* occurs twice.

Linked to the issue of 'deserving' are terms which describe what happens to asylum seekers who are not found deserving. The term *rejected* occurs 18 times in the UN corpus (Table 10) and is linked to asylum seekers in 15 of these cases. One interesting facet of the term *rejected* is that these constructions very rarely explicitly state who is responsible for the rejecting. — the term *rejected*

**Table 10.**  Concordance of *rejected*.

| |
|---|
| y would continue enjoying refugee status. If rejected , they could appeal, and failing that, they |
| bidjan. But over the weekend, local residents rejected the idea of relocation on their soil, citing |
| te d'Ivoire in order to overthrow him. UNHCR rejected the allegation as groundless. The a |
| as the right to appeal if his or her application is rejected . Some Eritreans say they ca |
| possibility that their claims may ultimately be rejected makes local integration a secondary con |
| ernments that they suspend the forced return of rejected Liberian asylum seekers for at least six m |
| Halt return of all rejected Iraqi asylum seekers, advises UNHCR |
| ments to temporarily halt any forced returns of rejected Iraqi asylum seekers. |
| day asked governments to refrain from sending rejected Iraqi asylum seekers back home against t |
| one month a moratorium on forcible returns of rejected Iraqi asylum seekers to their country. |
| for all Iraqis abroad, and hold deportations of rejected Iraqi asylum seekers for an initial three m |
| lso appealed for governments not to force any rejected Iraqi asylum seeker to return now to a c |
| tend its moratorium on decisions and return of rejected cases until February 2004. |
| ve a ban on forced returns to Iraq, including of rejected asylum seekers. UNHCR also advis |
| ssing of Iraqi asylum claims against returning rejected asylum seekers who had suffered persecu |
| in camps like Um Gargour. Those who were rejected , as well as those who registered for volu |
| eatings, banditry and death threats. But she rejected any notion that her life was one of sacrifi |
| anage to get into the asylum procedure but are rejected after their first application, the directive d |

*by* never occurs in the corpus. Just over half of the occurrences of *rejected* occur as adjectives (9 appear as part of the phrase *rejected (modifiers) asylum seekers*, while there is one occurrence of *rejected asylum cases* — many of these cases specifically refer to Iraqi asylum seekers). There are a smaller numbers of references (5) to rejected asylum seekers where *rejected* is a verb: e.g. "asylum seekers who manage to get into the asylum procedure but are rejected after their first application".

The term *forced return/s* occurs 14 times in the corpus, while *forcible returns* and *forcibly returned* appear once each (examples relating to asylum seekers are shown in Table 11).

Note here that agency is again left unclear. It is not said who is forcing the returns, and in six of these cases, the term *forced return* is not linked specifically in the same sentence to a person or group of people. Instead, a *forced return* stands in as a label for a type of person or group, a form of impersonalisation (van Leeuwen 1996).

The term DEPORT occurs seven times in reference to asylum seekers, in three cases agency is made explicit: *Malaysia deports asylum seekers… Malaysian authorities today deported seven asylum seekers to the strife-torn… R's note appealed to the Indonesian authorities not to deport the likely asylum seekers,*

**Table 11.**  Concordance of *forced returns* and similar phrases.

| | |
|---|---|
| 'compelling reasons' against forced returns | UNHCR first made |
| e agency and despite recent assurances against forced returns | According to the U |
| e agency has advised states to preserve a ban on forced returns | to Iraq, including of rejected asyl |
| point, advising governments to continue a ban on forced returns | to Iraq until further notice. |
| NHCR had already urged asylum countries to halt forced returns | of Iraqis. Now the agency' |
| ier appeal for governments to temporarily halt any forced returns | of rejected Iraqi asylum seekers. |
| Rwandan refugees fear forced returns | from Tanzania. UNHCR |
| rs, citing the principle of non-refoulement or non- forced return | . The incident started |
| s overturned on appeal. So the possible forced return | of such a wide range of people be |
| alaysian government for a moratorium against the forced return | of asylum seekers to Aceh. |
| ommended to governments that they suspend the forced return | of rejected Liberian asylum seeke |
| Colombians should not be forcibly returned | from Panama, says UNHCR |
| xtend by at least one month a moratorium on forcible returns | of rejected Iraqi asylum seek |

*citing the principle…* In other cases, DEPORT occurs as a noun: e.g. *deportations of (rejected Iraqi) asylum seekers* (3), *access to the deportees.*

As with refugees, asylum seekers are also constructed in the UN texts in terms of being vulnerable to danger. *Vulnerable* occurs 76 times in the UN corpus, although there are only two cases of explicit referral to "vulnerable asylum seekers". In addition, the term *vulnerability of refugees and asylum seekers* occurs once. The terms *the issue of* and *the plight of* in relation to asylum seekers also construct them as a problematised group. The verb lemma FACE is often used in order to demonstrate that asylum seekers are in a difficult position (Table 12).

**Table 12.**  Concordance of the lemma FACE.

| | | |
|---|---|---|
| als with asylum-seeker certificates continue to | face | serious problems securing residency per |
| sian Federation. But even then, they continue to | face | serious problems securing residence per |
| erious concern for asylum seekers. They | face | problems getting such permits because, un |
| ned to their country of origin, where they could | face | persecution. UNHCR's note appealed t |
| seekers back to their countries where they could | face | persecution. UNHCR offered to assess |
| Feature: Asylum seekers in Moscow | face | major challenges despite UNHCR efforts |
| Another problem asylum seekers | face | is the attitudes of the local population. |
| ugee rights and the difficult circumstances they | face | in seeking asylum. A special UNHCR day |
| e their homes is enhanced by the problems they | face | face in adjusting to a new, often hostile ho |
| Majid's scepticism reflects the dilemma | faced | by many asylum seekers in Russia. |
| This helps to reduce the isolation and stress | faced | asylum seekers, strengthen grassroots soc |
| be displayed to show the trials and tribulations | faced | by asylum seekers. |
| the most serious issue facing | | asylum seekers is the lack of official docu |

FACE as a verb in the BNC shows a negative discourse prosody, with strong collocates being *problems*, *dilemma*, *challenges*, *difficulties*, *charges* and *eviction*.

In addition, the issues that asylum seekers encounter are phrased as being multiple either by using plurals, e.g. *problems*; *difficult circumstances*; *serious problems* or by referring to more than one form of problem, e.g. *isolation and stress faced by*; *trials and tribulations faced by*; *despair and hopelessness*. In the bottom line of the above concordance, although the problem is in the singular: "the…. issue facing asylum seekers", the fact that it is prefaced by "the most serious issue" suggests that there are other serious issues which are not being discussed here.

Finally, asylum seekers are constructed as needing help and assistance — a discourse which acknowledges the fact that they face multiple difficulties by using verbs such as *assist*, *enable* and *help*: e.g. *to help asylum seekers adjust to these difficulties*, *assisting asylum seekers* (2), *to better assist asylum seekers*, *to help*, *enable asylum seekers to*. In addition to this, the term *promote* is occasionally used in reference to asylum seekers, although this tends to occur in relationship to non asylum-seeker's attitudes — e.g. *To raise public awareness and promote acceptance of asylum seekers in the Russian Federation*, *promote better relations*, *promote intercultural tolerance*.

### *The News Corpus*

The News Corpus contains 359 cases of the word *asylum* of which 131 are part of the phrase *asylum seekers*. *Asylum seeker* occurs 10 times, whereas other strong immediate right hand collocates are *asylum bill* (16) and *asylum appeals* (12). One way that the news data differs from the UNHCR data is in its description of appeals that have not been successful. In contrast to the UNHCR corpus, the News Corpus favours the term *failed asylum seekers* (33 occurrences) instead of *rejected asylum seekers* (4 occurrences). REJECT in the BNC collocates with nouns such as *proposal*, *idea*, *suggestion* and *argument* suggesting a semantic prosody for intellectual propositions, whereas *failed* as an adjective collocates with *coup*, *attempt*, *marriage*, *uprising*, *bid* and *rescue* — having a wider semantic prosody, more concerned with events rather than ideas. Perhaps the difference between *rejected* and *failed* is that *rejected* is less personal and emotive — implying that the application has been rejected rather than the person. *Failed* implies a more personal failure. There are only two cases of *unsuccessful asylum seekers* in the News Corpus.

The News Corpus also has a slightly wider range of terms for the 'honesty' of asylum seekers: *genuine* 2, *illegal* 2, *obviously fraudulent* 1, *false*, 2, *bogus* 3. As Kaye (1998: 171) found, such expressions are frequent in newspaper reports of asylum seekers, used in 58% of articles that were analysed. While Refaie (2001: 368) argues that when metaphors are constantly repeated they can become naturalised, similarly, Hunston (2002: 119) points out that frequent collocations can eventually become so entrenched that they can prime one another. For example, the frequent use of the collocates *illegal* and *immigrant* in the phrase *illegal immigrant* may lead people to view all immigrations as illegal. The same process may be true with discourse prosody — if the term *asylum seeker* is perpetually paired with a range of concepts that express falsity, then eventually people may be primed to think of this concept whenever they encounter the term *asylum seeker*.

The newspapers also use constructions which report quantities of asylum seekers: e.g. *number(s) of* (10), although this is generally a less prevalent feature than the UNHCR texts. The News Corpus surprisingly tends not to talk about specific asylum seekers e.g. by naming different nationalities (this only happens in 13 cases). There are fewer uses of the conjunction *and* in the News Corpus, linking asylum seekers with other groups, e.g. *immigration and asylum* occurs eleven times (although these are referred to as issues rather than people). There are seven cases of asylum seekers being linked to other groups via *and*, although none of these are refugees as in the UNHCR Corpus. Instead, asylum seekers are linked to *illegal/unsuccessful immigrants* (4), *would-be immigrants* (1) and *terrorists* (2).

What the News Corpus does contain, though, are references to, descriptions of and evaluations of legislation affecting asylum seekers. For example, the term *asylum bill* appears 17 times, whereas *asylum and immigration bill* appears 5 times. There are references to the bill in terms of words like *process* (7), *plan* (11), *mechanism* (2), *legislation* (5), *policy* (10), *procedure* (2) and *reform* (18).

There are also commentaries on concerns people have about asylum seekers and over current and future legislation, shown by collocational words and phrases such as *fears over*, *outcry over*, *rebel over* and *expressed opposition to*. Although the nature of *fears* is not always from the same political perspective, the term is another word which is often agentless (Table 13).

In this table, in six cases the word *fears* is given a direct agent, but there are also cases where the owner(s) of the fears are not identified: e.g. *The figures reinforce fears that the Government is failing to fund the asylum process enough.*

**Table 13.**  Concordance of *fears*.

| | | |
|---|---|---|
| h the need for independence." He hoped his | fears | were unjustified, but he found it worrying |
| new EU citizens' right to work here — prompting | fears | that it will become extremely attractive f |
| urt with Supreme Courts in other countries and | fears | that the British model will be "second clas |
| ecisions in the first place. The figures reinforce | fears | that the Government is failing to fund the |
| ace. Events overseas have conspired to fuel the | fears | Right-wing commentators point to the N |
| Plea to stem rural | fears | over asylum: Survey reveals climate of m |
| ampdowns in even the most liberal of countries. | fears | of a huge influx of immigrants, social we |
| the EU, and many others are doing so illegally. | Fears | of a big new migration wave from central |
| Last week the Daily Mail also raised | fears | of a wave of HIV-infected migrants arr |
| s will be allowed to work in the UK, prompting | fears | of "benefit tourism". |
| mmunities to those seeking refugee status. The | fears | of local people and refugees themselves |
| onal Affairs; Social Welfare & Social Services | fears | of migration from east to west ON ONE s |
| Blunkett is right on asylum | fears | . |
| holly-elected Lords. The real argument: Woolf | fears | MPs will follow the American example, |
| s been told there is no evidence to support such | fears | Her officials say there is nothing to sugg |
| | Fears | are mounting for the three asylum seekers |
| Elsewhere in the group, | fears | are rising that the merger of Group 4 and |
| Playing on public | fears | after the September 11th terrorist attacks |

In some cases, the agent is made vague e.g. *Plan to stem rural fears over asylum*. Who is responsible for rural fears? Rural people? The answer to the question is left open.

There are also descriptions of legislation to make it more difficult for people to successfully claim asylum, described as *curbs* (2) and a *crackdown* (3). Policy is described as becoming more *stringent* (3), and the word *limit* occurs as a collocate of *asylum seekers* (Table 14).

**Table 14.**  Concordance of the verb lemma LIMIT.

| | |
|---|---|
| volving its members in Calais, whose task is to limit | the number of asylum seekers entering B |
| als to higher courts, the new legislation would limit | unsuccessful asylum applicants to a single |
| n it comes to reform of the asylum system and limiting | the right of asylum seekers to appeal |
| is "two-tier" tribunal and high court process to limit | the opportunity for failed asylum seekers |
| s part of an asylum crackdown which includes limiting | appeals for failed asylum seekers, ending |

*Limit* as a verb in the BNC collocates with nouns such as *earnings*, *powers*, *cash*, *account* and *shares*, suggesting a semantic prosody linked to finance.

Other phrases relating to asylum seekers in the News Corpus include: *tightens the screws*, *getting tougher on*, *oust*, *controls on*, *refusals of*, *to lay into* and *remove the right of*. Asylum policy is also described critically: *creaking asy-*

*lum system*, *decrepit asylum system*, *controversial asylum legislation* (2), *vexed asylum issue*, *asylum shambles*, *asylum crisis*, *asylum problem* and an *issue* (5).

Various aspects of asylum are also described in terms of a number of metaphors which reference violence or war. For example, the process of seeking asylum is seen to be a minefield: *The sensitive issue of immigration — and asylum in particular — is proving to be a minefield for ministers to negotiate*. In addition, it is viewed as a fight (2 cases): *Lawyers are to receive less money to fight asylum cases*; *families who exhaust their legal fight for asylum will be denied welfare payments*… Legislation itself is viewed as a battle: *The new battle over immigration and asylum*… while the system is said to *suffer a hammer blow today when immigration officers begin industrial action*.

The term *blow* occurs in another article entitled: *Asylum blow as immigration officers start a work-to-rule*. Another violence metaphor is *attack*, which occurs in almost all cases as part of a news story relating to Lord Woolf's speech on the Asylum Bill (Table 15).

**Table 15.** Concordance of the verb lemma ATTACK.

| | |
|---|---|
| accept the constitutional reforms. The asylum attack | from the Lord Chief Justice is likely to pl |
| enior judge last night launched a twin-pronged attack | on the Government's constitutional and as |
| m seekers, our unwritten constitution is under attack | This is all done in the name of "third wa |
| ies of Home Secretary David Blunkett. But his attack | on the Asylum Bill carried the unprecede |
| S most senior judge launched a devastating attack | on the Government's latest asylum crackd |
| Lord Chief Justice leads double attack | on supreme court and asylum TWO of B |
| h MPs over asylum reform after the swingeing attack | from the lord chief justice in which he sai |
| re at the University of Cambridge, Lord Woolf attacked | part of the Asylum and Immigration Bill, |
| LORD WOOLF, the Lord Chief Justice, attacked | the Government's plans for asylum law ch |

How does the News Corpus compare to the UNHCR Corpus in the ways that it describes what happens to asylum seekers after their claim is rejected? The verb lemma ESCORT occurs five times (see Table 16). Note that in three of these cases, the verb occurs with the word *refuse* or *refusal*, whereas in the final concordance line, there is the negative construction "not being escorted". These cases refer to newspaper stories that describe how an overtime ban or staff shortages at the Immigration Service Union have led to asylum seekers not being escorted to airports.

ESCORT as a verb collocates in the BNC with a number of words relating to war: *destroyers*, *civilian*, *troops*, *officers*, *armed* or crime: *police*, *prisoners*. DE-PORT occurs thirty-three times in relation to asylum seekers, of which 18 occur as the noun *deportation(s)* and the remaining 15 are verbs. As with a number

**Table 16.**  Concordance of the verb lemma ESCORT.

| | |
|---|---|
| eekers being lost, if officers refuse to carry out escort | duties when private contractors are unav |
| l asylum seekers facing deportation should be escorted | to airports by private contractors. How |
| claimants a week. And they will refuse to escort | failed asylum seekers on to flights out of t |
| the second plank of the action — the refusal to escort | asylum seekers to deportation flights — wi |
| arch 1) about failed asylum seekers not being escorted | to airports due to staff shortages in the Im |

of other verbs, it is difficult to determine agency with the lemma DEPORT. The nominalised forms do not require an agent: e.g. *It is one of the biggest deportations in modern European history…* although the verb forms often do not show agency either: *It is likely that the Piraees will be among the 3,000 who are due to be deported this summer*. Only six cases out of 33 of DEPORT reveal who is the person or group responsible for deporting.


## Conclusion

As Jaworski and Coupland (1999: 9) point out, "most texts are not 'pure' reflections of single discourses." This point is even more relevant when considering collections of texts. The newspaper articles in the News Corpus were written by a variety of authors (and subjected to editing processes), and for a range of newspapers with different political standpoints that may shift over time. The UNHCR Corpus, although much larger, contains a more coherent discourse of refugees and asylum seekers, largely because it is an organisation set up to deal with refugees and is therefore more careful to present a more consistent worldview. The UNHCR focuses more on returning refugees to their countries of origin, and while it highlights a number of problems that refugees encounter (for example *conflict* and *HIV*), it does not normally represent refugees as helpless victims or nuisances, although indirectly, by referring to their needs and ways of helping them, other conclusions may be inferred. The UNHCR's view of refugees is as a global issue, rather than focussing on a single group at any given time. It also focuses on its own actions in relation to helping refugees.

In contrast, the News Corpus is more concerned with the impact of refugees on the UK rather than taking a global perspective. Attitudes are presented in a more ambivalent and complex way, with refugees constructed as tragic victims, an out-of-control mass, pests or potential invaders. Metaphors of water or packages serve to dehumanise refugees further. The News Corpus also refers to official attempts to help refugees, but this is simply one discourse type among many which are present. The fact that the term *refugee* is used in metaphorical

constructions "like some refugee from the Ricky Lake Show", reveals the negative connotation embedded within the word. While we would perhaps expect to find water metaphors in the News Corpus as this is consistent with the work of others like Refaie, the presence of such metaphors in the UNHCR Corpus is more surprising, showing the power of naturalised metaphors — it becomes very difficult to talk about a subject except in terms of the dominant discourse surrounding it.

Ironically, although the term *refugee* carries a mainly negative discourse prosody, it is an identity which is seen as preferable to (or even the goal identity of) *asylum seekers*, who are viewed in the newspapers with considerably more suspicion. The use of terms like *failed asylum seekers*, *fears over asylum*, *crackdown*, *bogus* etc. suggest that some British newspapers are engaged in the maintenance of a moral panic (Cohen 1972) over asylum, which has been focussed around legislation. Both the News and UNHCR corpora referred to the process of asylum seeking by using language which obscured agency, particularly in the cases of who makes decisions and who deals with the consequences when an asylum seeker's application is rejected. Nominalised and impersonalised phrases like *rejected cases* and *forced returns* serve to highlight an invisible bureaucracy at work.

Many of the linguistic strategies used to refer to refugees and asylum seekers — such as referring to them as an indistinguishable mass or vague quantity, using metaphors, describing them as bogus or referring to unspecified 'fears' — serve several purposes which are linked to the notion of racist discourse. As van Dijk (1987: 58) describes, there are four topic classes for racist discourses: they are different, they do not adapt, they are involved in negative acts and they threaten our socio-economic interests. Hardt-Mautner (1995:179) points out, "National identity emerges very much as a relational concept, the construction of 'self' being heavily dependent on the construction of 'other'". The racist constructions of refugees and asylum seekers, therefore, not only construct a threat to the status quo and national identity (which incidentally helps to sell newspapers), they also help to construct national identity by articulating what it is not.

However, more encouraging aspects of the corpus data suggest a less prejudiced picture than earlier researchers have found when looking at newspaper data. Stereotypes of refugees as criminal nuisances and constructions of asylum seekers as 'bogus' were still present in the corpora, yet they were relatively rare. Discourses which focussed on the problems encountered by refugees and asylum seekers and/or attempts to help them were relatively more common,

suggesting that in 2003 at least, there was a growing awareness of the need for sensitivity when discussing issues connected to immigration in the UK. As Law et al. (1997:18) found in a recent study, about three quarters of news articles concerned with race contained media frames

> "which seek to expose and criticise racist attitudes, statements, actions and policies, which address the concerns of immigrant and minority ethnic groups and show their contribution to British society, and which embrace an inclusive view of multi-cultural British identity".

A study by Jessika terWal concluded that "the British tabloid press no longer seem to merit the overly racist tag that they were given by studies in the early 1980s." (2002: 407)

A corpus-based approach is therefore useful, in that it helps to give a wider view of the range of possible ways of discussing refugees and asylum seekers. A more qualitative approach to analysis may mean that saliency is perceived as more important than frequency — whereby texts which present shocking or extreme positions are focussed on more than those which are more frequent, yet neutral. While it is important to examine extreme cases, it is also useful to put them into perspective alongside other cases. In addition, corpus data can help us to establish which sorts of language strategies are most frequent or popular. For example, the refugees as water metaphor was found to be much more frequent than other metaphors, such as refugees as illegal packages or as invaders. Rather than simply listing the metaphors which appear in the data then, we are able to get a more accurate sense of which ones are naturalised, and which ones may be particularly salient *because* they are so infrequent.

In addition, the corpus-based approach enables the researcher to arrive at a more complete understanding of the meanings and functions of certain word choices in texts about refugees and asylum seekers. The connotative use of language in critical discourse analysis is one of the most fruitful areas of analysis available to researchers — and by looking at the collocational strength of lexical items in a corpus of general language, we are given an objective sense of the themes and associations that are embedded in words due to their continual pairing with other words. By 'exposing' the hidden collocations of certain words, we can explain that a certain word or phrase contains a hint of bias, but have not been able to specify why.

However, corpus-based research is not without its problems. Taking a wider view may mean that subtleties of language are overlooked. In addition, corpus-based data does not explain itself; the interpretative skills of the researcher are still required. And the context is still paramount; perhaps not all concordance

lines possess equal 'strength' in terms of their influence in forming discourses. The identity and power of the text producer, the number of people who have access to a text and the extent to which a text echoes what is already 'accepted wisdom' — such issues need to be taken into account when carrying out corpus-based analysis.

We also need to consider the relationship between text types, rather than viewing the two corpora as separate entities. While it is the case that the UNHCR helps to play a part in the type and amount of information which is available to newspapers on refugees (via press releases and other publications), newspapers are at liberty to use, manipulate or discard such information at will, and the discourses within them are likely to reach the lives of many more ordinary people. The success of such discourses, among other things, may impact upon voting behaviour, meaning that government policy, to an extent, is influenced by media discourse. The power of such discourses is therefore not to be under-estimated.

## Notes

**1.** *The Guardian* (May 22, 2001) notes that in reporting William Hague's speech on asylum policy May 19th 2001, *The Times* referred to 'asylum seekers' in its first paragraph; 'immigrants' in its second; and 'refugees' in its third. The Guardian says that *refugee* is a word which evokes more immediate sympathy, whereas *asylum seeker* is a colder, more bureaucratic term.

**2.** The newspaper articles were taken from an online collection of UK national newspapers which covered the following: *The Guardian* and *The Observer*; *The Independent and Independent on Sunday*; *The Daily* and *Sunday Telegraph*; *The Times* and *Sunday Times*; *Daily Mail* and *Mail on Sunday*; *Daily Mirror* and *Sunday Mirror*; *The Express* and *Express on Sunday*; *The Economist*; *Evening Standard*; *Sunday People*.

**3.** See http://infoweb.newsbank.com.

**4.** See http://www.unhcr.org.

**5.** A lemma consists of all the possible forms of a word for its grammatical category. For example, the verb lemma SLEEP consists of the verbs *sleep*, *sleeping*, *sleeps* and *slept*.

**6.** Collocates in the BNC are calculated using the log-log score provided by the BNC Web Server at Zurich University.

**7.** As the word *heave* only occurs 206 times in the BNC, it does not have many collocates. In cases like this, it is sometimes useful to widen the search to look at the lemma of a word e.g. *heave*, *heaves*, *heaved*, *heaving*. The verb lemma HEAVE occurs 687 times in the BNC and gives a wider range of collocates to examine. However, care should be taken with this

procedure as sometimes different forms of a lemma can have very distinct collocates e.g. *kick* collocates with *bucket*, but *kicking* does not collocate with *bucket*. It is therefore important to investigate the collocates of lemmas to see whether they are due to idiomatic phrases such as *kick the bucket* and are therefore not true indicators of the collocation between a lemma and another word.

## References

Aston, G. and Burnard, L. 1998. *The BNC Handbook: Exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.

Becker, H. 1972. Whose side are we on? In J. D. Douglas (ed.). *The Relevance of Sociology*. New York: Appleton-Century-Crofts, 99–111.

Burr, V. 1995. *An Introduction to Social Constructionism*. London: Routledge.

Cohen, S. 1972. *Folk Devils and Moral Panics*. London: MacGibbon and Kee Ltd.

Downing, J. D. H. 1980. *The Media Machine*. London: Pluto.

Fairclough, N. 1989. *Language and Power*. London: Longman.

Fairclough, N. 2000. *New Labour, New Language?* London: Routledge.

Fairclough, N. 2003. *Analysing Discourse: Textual Analysis for Social Research*. London: Routledge.

Flowerdew, J. 1997. The discourse of colonial withdrawal: a case study in the creation of mythic discourse. *Discourse & Society* 8, 453–477.

Hardt-Mautner, G. 1995. How does one become a good European: The British Press and European integration. *Discourse & Society* 6:2, 177–205.

Hall, S., Critcher, C., Jefferson, T., Clarke, J. and Roberts, B. 1978. *Policing the Crisis: Mugging, the State, and Law and Order*. London: Macmillan.

Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Kaye, R. 1998. Redefining the refugee: The UK media portrayal of asylum seekers. In: K. Koser and H. Lutz (eds). *The New Migration in Europe: Social Constructions and Social Realities*. London: MacMillan Press, 163–182.

Krishnamurthy, R. 1996. Ethnic, racial and tribal: The language of racism?'. In: C. R. Caldas-Coulthard and M. Coulthard (eds). *Texts and Practices: Readings in Critical Discourse Analysis*. London: Routledge, 129–49.

Law, I, Svennevig, M. and Morrison, D. E. 1997. *Privilege and Silence. 'Race' in the British News during the General Election campaign, 1997. Research Report for the Commission for Racial Equality*. Leeds: University of Leeds Press.

Louw, B. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In: M. Baker, G. Franics and E. Tognini-Boneli (eds). *Text and Technology*. Amsterdam: Benjamins, 157–76.

Parker, I. 1992. *Discourse Dynamics: Critical Analysis for Social and Individual Psychology*. London: Routledge.

Piper, A. 2000. Some people have credit cards and others have giro cheques: 'Individuals' and 'people' as lifelong learners in late modernity. *Discourse & Society* 11, 515–542.

Refaie, E. 2001. Metaphors we discriminate by: Naturalized themes in Austrian newspaper articles about asylum seekers. *Journal of Sociolinguistics* 5:3, 352–371.

Reisigl, M. and Wodak, R. 2001. *Discourse and Discrimination: Rhetorics of Racism and Antisemitism*. London: Routledge.

Scott, M. 1999. *WordSmith Tools Help Manual*. Version 3.0. Mike Scott and Oxford University Press.

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

Stubbs, M. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. London: Blackwell.

TerWal, J. 2002. *Racism and Cultural Diversity in the Mass Media*. Vienna: European Research Center on Migration and Ethnic Relations.

Teubert, W. 2000. A province of a federal superstate, ruled by an unelected bureaucracy: Keywords of the Eurosceptic discourse in Britain. In: A. Musolff, C. Good, P. Points and R. Wittlinger (eds). *Attitudes Towards Europe: Language in the Unification Process*. Aldershot: Ashgate, 45–86.

van Dijk, T. 1987. *Communicating Racism: Ethnic Prejudice in Thought and Talk*. London: Sage.

van Dijk, T. 1991. *Racism and the Press*. London: Routledge.

van Dijk, T. 1996. Discourse, power and access. In: R. Caldas-Coulthard and M. Coulthard. (eds). *Texts and Practices: Readings in Critical Discourse Analysis*. London: Routledge, 84–104.

van Dijk, T. 2001. Critical discourse analysis. In: D. Schiffrin, D. Tannen and H. E. Hamilton. (eds). *The Handbook of Discourse Analysis*. London: Blackwell, 352–371.

van Leeuwen, T. 1996. The representation of social actors. In: C. Caldas-Coulthard and M. Coulthard (eds). *Texts and Practices: Readings in Critical Discourse Analysis*. London: Routledge, 32–70.

Wodak, R. and Meyer, M. 2001. *Methods of Critical Discourse Analysis*. London: Sage.

*Authors' addresses*

Paul Baker and Tony McEnery
Department of English Language and Linguistics
Lancaster University
Lancaster, Lancashire
LA1 4YT

Email: p.baker@lancaster.ac.uk; a.mcenery@lancaster.ac.uk

*About the authors*

**Paul Baker** is a lecturer in the Department of English Language and Linguistics at Lancaster University. His previous research includes *Polari: The Lost Language of Gay Men* (2002) and *Public Discourses of Gay Men* (2005), both published by Routledge.

**Tony McEnery** is a professor in the Department of English Language and Linguistics at Lancaster University. His previous research includes *Corpus Linguistics* (2001, with Andrew Wilson) published by Edinburgh University Press and *Swearing in English* (in press), published by Routledge.