BAAL/IRAAL 2006

Representation of refugees and asylum seekers in UK newspapers

Towards a corpus-based analysis

Costas Gabrielatos and Paul Baker
Lancaster University

8 September 2006

Corpus-based methodology in discourse analysis

- Helps recognise the 'big picture'.

- Indicates areas of interest.

- Also allows for closer (qualitative) analysis.

- Examination of a large number of instances of a specific pattern (through concordances).

- Triangulation / checking of intuitions.

- Quantification

(Baker, 2006: 10-21, 175-183)

The Corpus (1996-2005)

Articles: *175,139*

Words: *139,510,037*

Sub-corpora:

❑ *per newspaper*

❑ *broadsheets/tabloids*

❑ *per year*

Business
Evening Standard
Express
Daily Mail + Mail on Sunday
Daily Mirror + Sunday Mirror
Guardian + Observer
Herald
Independent + Independent on Sunday
Liverpool Echo
People
Star
Sun
Telegraph + Sunday Telegraph
Times + Sunday Times

The Query

- refugee OR asylum OR deport* OR immigr* OR emigr* OR migrant OR illegal alien OR illegal entry OR leave to remain AND NOT deportivo AND NOT deportment

- Derived on the basis of two tests:
  - Keywords in the pilot corpus as compared to a general corpus (BNC Sampler)
  - Relevance of key and other intuitively identified terms to the core query terms *(refugee*, asylum seeker*).*

Relevant concepts (1)

- **Collocation**
  - The co-occurrence of two words with frequency above chance (statistically calculated).
  - Collocation is directional.
  - The collocates of a word contribute to its meaning (Nattinger & DeCarrico, 1992: 181-182), or can provide "a semantic analysis of a word" (Sinclair, 1991: 115-116).
  - "Patterns of association – how lexical items tend to co-occur – are built up over large amounts of text and are often unavailable to intuition or conscious awareness. They can convey messages implicitly and even be at odds with an overt statement" (Hunston, 2002: 109).

- **Consistent collocates**
  - Collocates present in at least seven 'year' sub-corpora.
  - MI score favours low-frequency collocates.
  - Extracting collocates from the whole corpus may (and does) include a large number of 'seasonal collocates'.
  - They indicate core elements of meaning / usage / associations.

Relevant concepts (2)

- ❑ **Collocational network**
  - ❑ A network created by a word/lemma and its collocates, or by intercollocating words/lemmas (Phillips, 1989, McEnery, 2006).

- ❑ **Semantic prosody**
  - ❑ The "consistent aura of meaning with which a form is imbued by its collocates" (Louw, 1993: 157).

- ❑ **Semantic preference**
  - ❑ A relation "between a lemma or word form and a set of semantically related words" (Stubbs, 2002: 65).

- ❑ **Topoi**
  - ❑ "Content-related warrants" or "conclusion rules that connect the argument with the conclusion" (Reisigl and Wodak, 2001: 74–76).
  - ❑ "The common-sense reasoning typical for specific issues", or "the most typical elements of the argumentative and persuasive nature of debates on immigration, integration and the multicultural society' (van Dijk, 2000: 97-98).

Preliminary findings from relevant study

❑ *Refugees* commonly described in terms of:

  ❑ provenance

  ❑ temporary residence

  ❑ destination

  ❑ quantity

  ❑ movement

  ❑ plight

❑ *Asylum seekers* frequently linked to:

  ❑ issues of legislation related to immigration

  ❑ problems allegedly stemming from their presence in the host country

  ❑ illegal / unsuccessful / would-be immigrants

(Baker & McEnery, 2005)

Link with other aspects of the project

❑ Previous papers focused on differences between the representation in broadsheets and tabloids.

❑ Overlap in terms of …
   ❑ phraseology surrounding *refugees* and *asylum seekers* (RAS)
   ❑ semantic prosodies (in)directly relating to RAS
   ❑ stance towards RAS

❑ Differences are in terms of degree:
   ❑ positive stance mainly in broadsheets; negative stance more in tabloids (but more balanced than positive).
   ❑ instances of mixed presentation: articles reporting the drowning/suffocating of refugees while trying to enter the UK also contain (in)direct references to their 'illegality'.

❑ A priori B/T distinction may not be always helpful.

❑ Examination of the frequency/use of specific collocations and expressions in individual newspapers can provide a way to quantify 'quality' in the press.

(Gabrielatos, 2006; Gabrielatos & Baker, 2006)

Focus

- Consistent collocates of *refugees, asylum seekers, immigrants, migrants* – particularly those related to *topoi, topics* or *categories* (Reisgl & Wodak, 2001; Sedlak, 2001; van Leuven, 1996; Wodak & van Dijk, 2000).

- Overlap of collocates and topoi/topics/categories of the four terms in focus.

- Intercollocations between the terms.

- Further examination of ways to quantify 'quality' in the press:
  - Relative frequency of *illegal refugee(s)* and *illegal asylum seeker(s)* in the newspapers, and their frequency development over time.

Definitions

- **Refugee**: Someone who "…owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country; or who, not having a nationality and being outside the country of his former habitual residence is unable or, owing to such fear, is unwilling to return to it."

  1951 Convention Relating to the Status of Refugees, Article 1A (2), http://www.refugeecouncil.org.uk/gettinginvolved/campaign/campaigners_pack/conventions_refugees.htm

- **Refugee**: "Someone whose asylum application has been successful and who is allowed to stay in another country having proved they would face persecution back home."

- **Asylum seeker**: "Someone who has fled persecution in their homeland, has arrived in another country, made themselves known to the authorities and exercised the legal right to apply for asylum."

[Refugee Council Online, http://www.refugeecouncil.org.uk/practice/basics/truth.htm]

Consistent collocates

- Total word forms: 152
    - *4%-9%* of the total collocates for each term.

- Content words: 135

- Categorisable: 116 (*86%*)

    - The vast majority of consistent content collocates can be clustered in only 8 categories.

Provenance / destination / temp. residence

refugees	asylum seekers	immigrants	migrants
		Africa	
		America	
			Australia
Britain	**Britain**	**Britain**	**Britain**
		China	China
		Eastern	Eastern
		EU	
		Europe	Europe
		France	
Germany			
		India	India
Iran			
	Ireland	Ireland	
		Italy	
Lebanon			
Pakistan			
		Russia	
		Soviet	
Turkey		Turkey	
	UK	UK	

Plight

refugees	asylum seekers	immigrants	migrants
aid			
desperate			
displaced			
driven			
escape			
fear			
fighting			
fled	fled	fled	
flee			
fleeing	fleeing		fleeing
forced	forced		
homeless			
homes			
leave			
persecution			
torn			

Movement / Entry

refugees	asylum seekers	immigrants	migrants
across	across	across	
arrive	arrive	arrive	
arrived	**arrived**	**arrived**	**arrived**
arriving	arriving	arriving	
border			
borders			
		came	
come	**come**	**come**	**come**
coming	**coming**	**coming**	**coming**
cross			
crossing			
	Dover		
enter	enter	enter	
	entered		
		entering	
trying	trying	trying	
	waiting		

Number

refugees	asylum seekers	immigrants	migrants
flood flooding pour poured pouring streaming	pouring	flooding	

Residence

refugees	asylum seekers	immigrants	migrants
allowed	accommodation allowed	allowed	
asylum			asylum
camp			
camps			
granted	granted		
	housed		
	housing		
	live		
	living	living	
seek			
seeking		seeking	seeking
settle		settle	
settled		settled	
shelter			
sheltering			
stay	stay	stay	
	temporary		
waiting			

Return / repatriation

refugees	asylum seekers	immigrants	migrants
back refused return returned returning	refused sent		

Legality

refugees	asylum seekers	immigrants	migrants
	bogus detention held illegal	arrested bogus caught deported detained held hiding illegal legal smuggled	genuine legal

Economic burden/threat

refugees	asylum seekers	immigrants	migrants
economic	benefits claim claims claiming economic entitled receive	jobs working	jobs work

Shared consistent collocates: proportion

↓	refugees	asylum seekers	immigrants	migrants
refugees		34%	33%	18%
asylum seekers	47%		46%	26%
immigrants	34%	40%		39%
migrants	38%	38%	79%	

Shared consistent collocates: categories

↓	refugees	asylum seekers	immigrants	migrants
refugees		ENTRY (9/13) NUMBER (1/1) ECON. BURDEN (1/1) RETURN (1/2)	ENTRY (8/12) RESIDENCE (5/13)	ENTRY (5/12)
asylum seekers	ENTRY (10/12) PLIGHT (3/3) NUMBER (1/1) RETURN (1/2)		ENTRY (9/11) LEGALITY (4/4) PDTR (3/3) RESIDENCE (4/9)	ENTRY (5/10)
immigrants	ENTRY (10/11) RESIDENCE (5/6) PLIGHT (1/1) NUMBER (1/1)	ENTRY (9/11) PLIGHT (1/1) RESIDENCE (3/6) LEGALITY (4/10)		ENTRY (6/10) ECON. THREAT (2/2)
migrants	ENTRY (3/3) RESIDENCE (2/2) PLIGHT (1/1) PDTR (3/6)	PLIGHT (1/1)	PDTR (6/7) ENTRY (3/3) RESIDENCE (2/2) ECON. THREAT (2/2) LEGALITY (1/2)	

Intercollocations
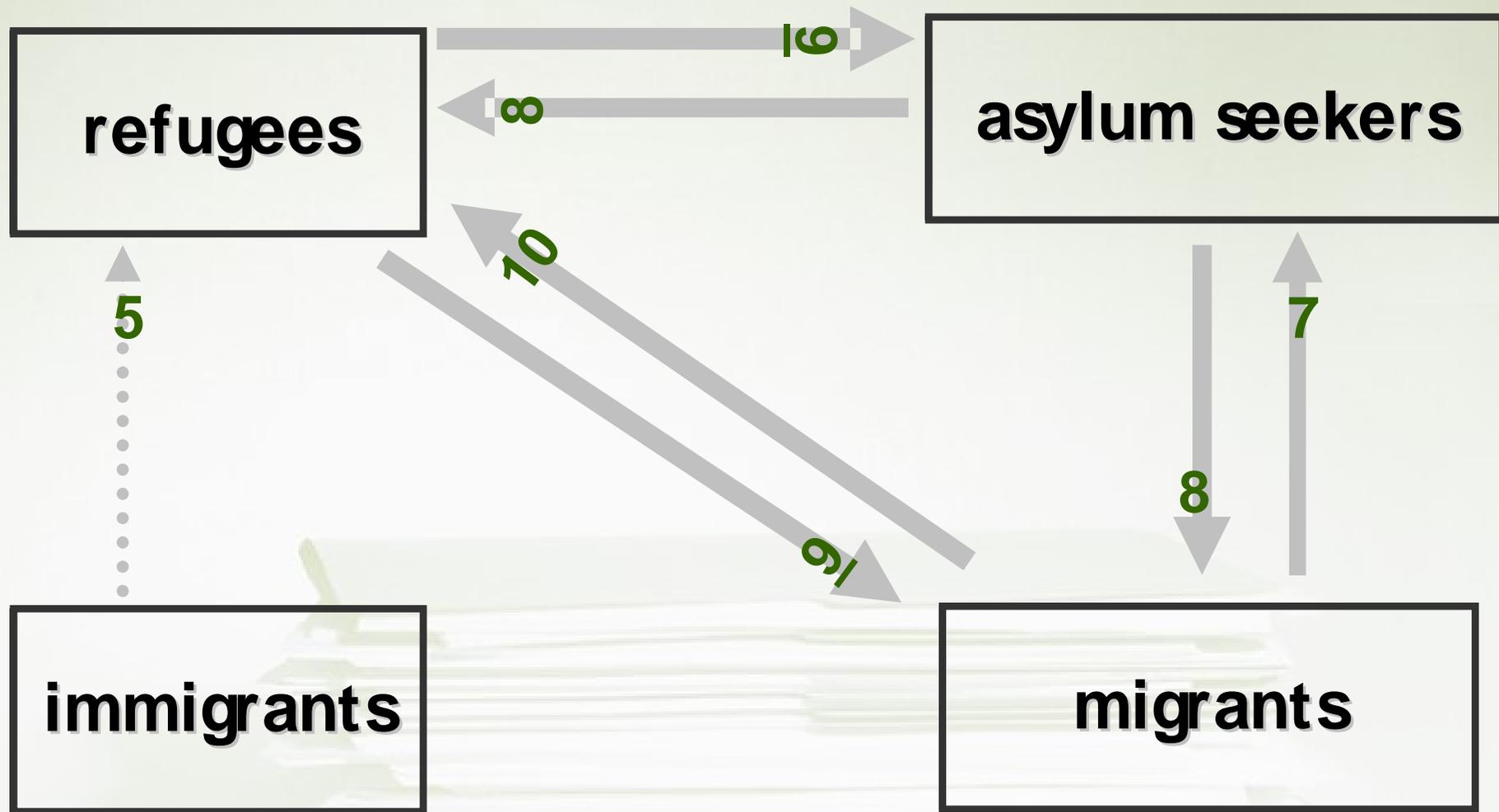
Distorting/misleading phraseology

"The Commission is concerned that editors should ensure that their journalists covering these issues are mindful of the problems that can occur and take care to avoid misleading or distorted terminology. By way of example, **as an "asylum seeker" is someone currently seeking refugee status or humanitarian protection, there can be no such thing in law as an "illegal asylum seeker".** A "refugee" is someone who has fled their country in fear of their life, and may have been granted asylum under the 1951 Refugee Convention or someone who otherwise qualifies for Humanitarian Protection, Discretionary Leave or has been granted Exceptional Leave to Remain in the country. An asylum seeker can only become an "illegal immigrant" if he or she remains in the UK after having failed to respond to a removal notice." [Our emphasis]
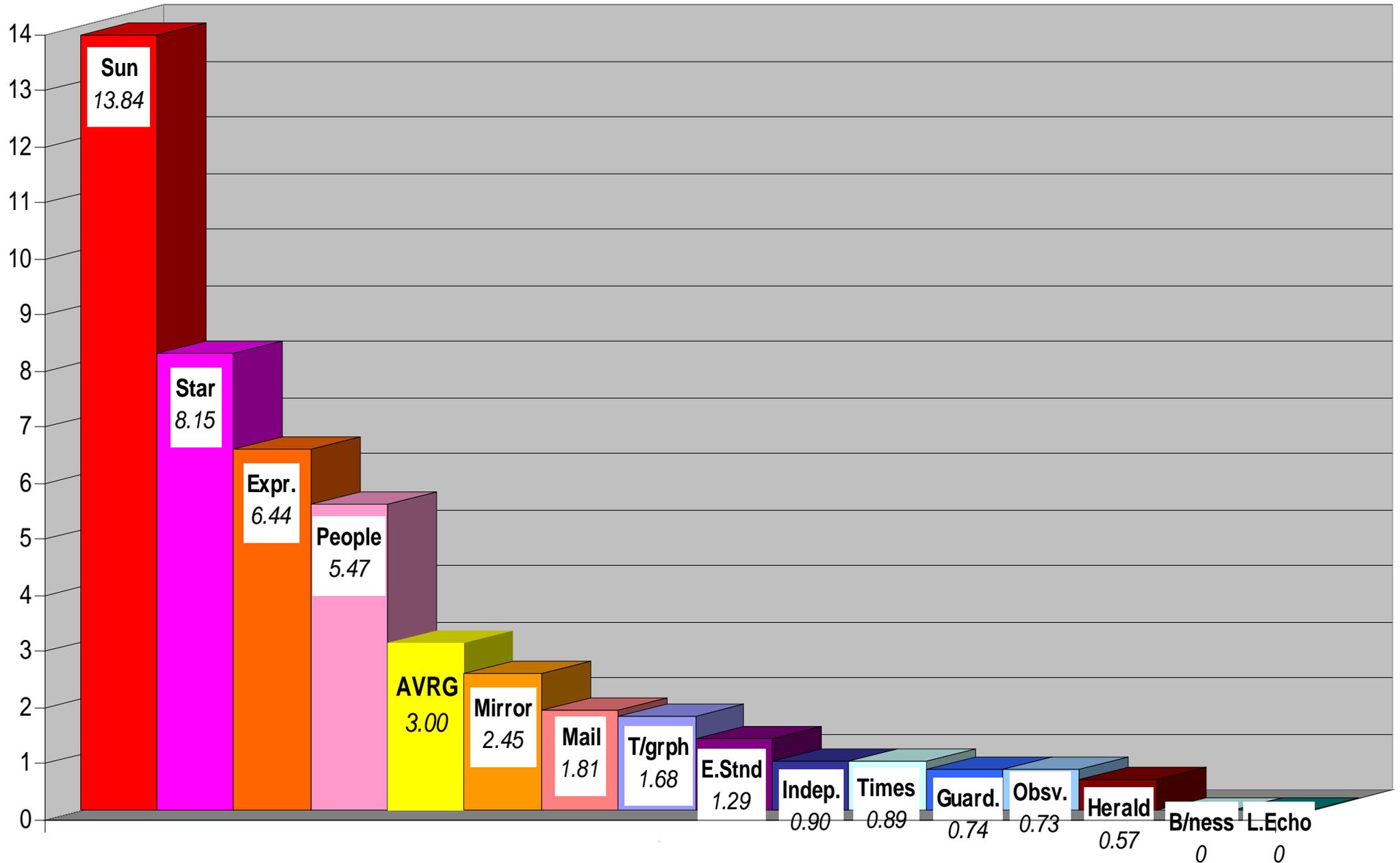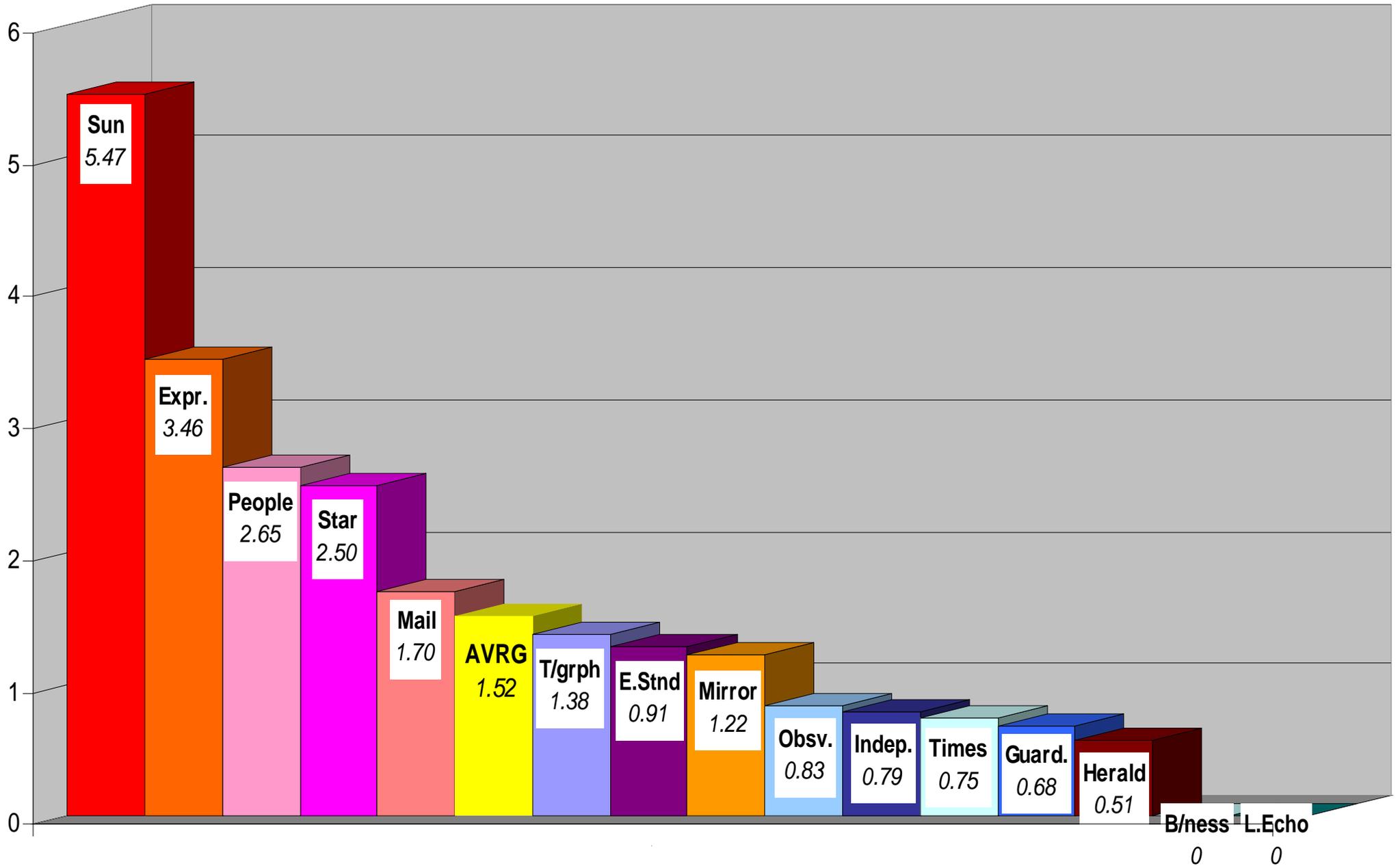
Press Complaints Commission, 23 October 2003
http://www.pcc.org.uk/news/index.html?article=OTE=

illegal asylum seekers / refugees

❑ Frequency in newspaper-specific sub-corpora

❑ Frequency per million words and per 1000 articles:

  ❑ Tabloid articles tend to be shorter than broadsheet ones, which tends to overestimate tabloid frequencies and underestimate broadsheet ones.
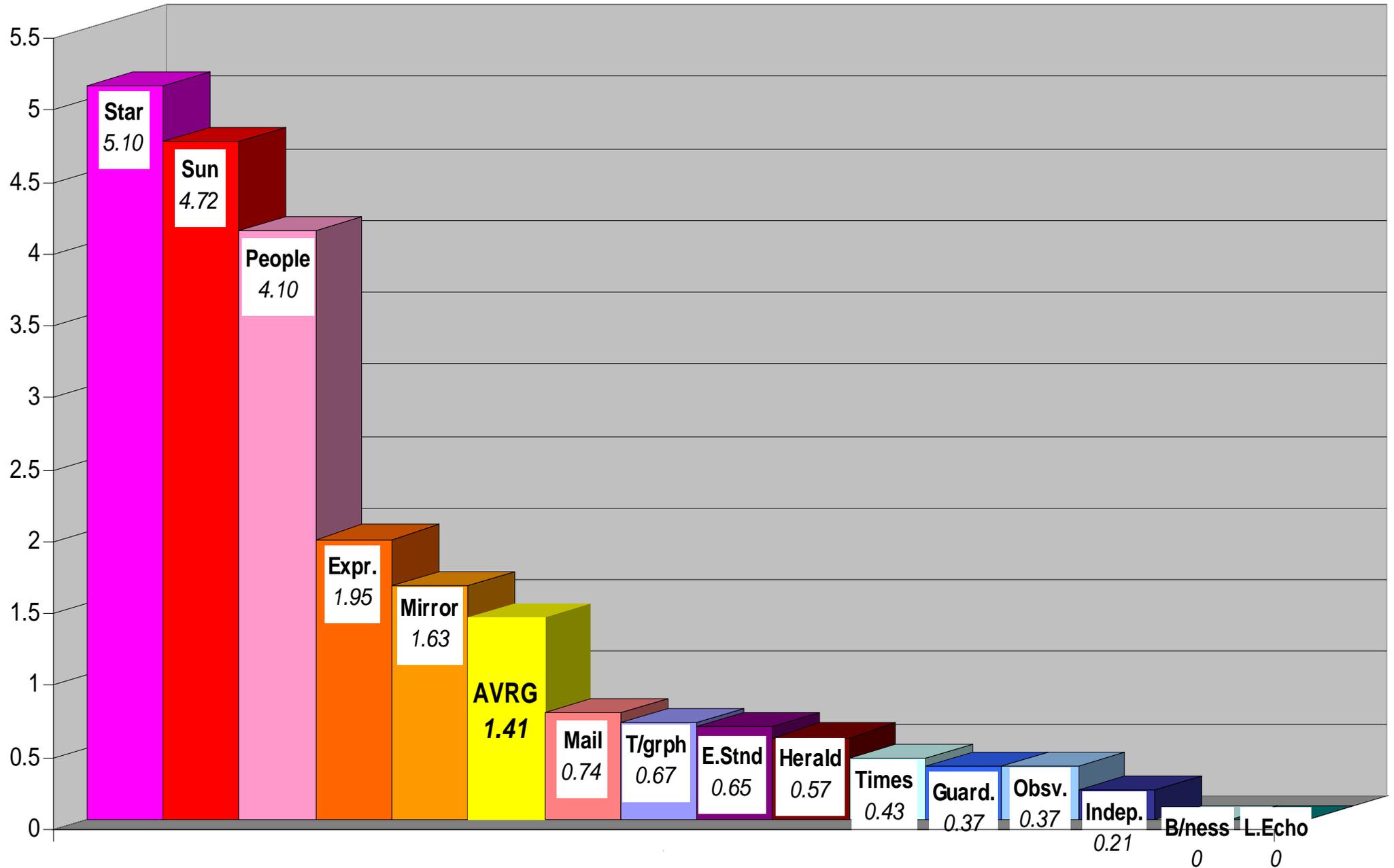
❑ Frequency development over time

Illegal asylum seekers:
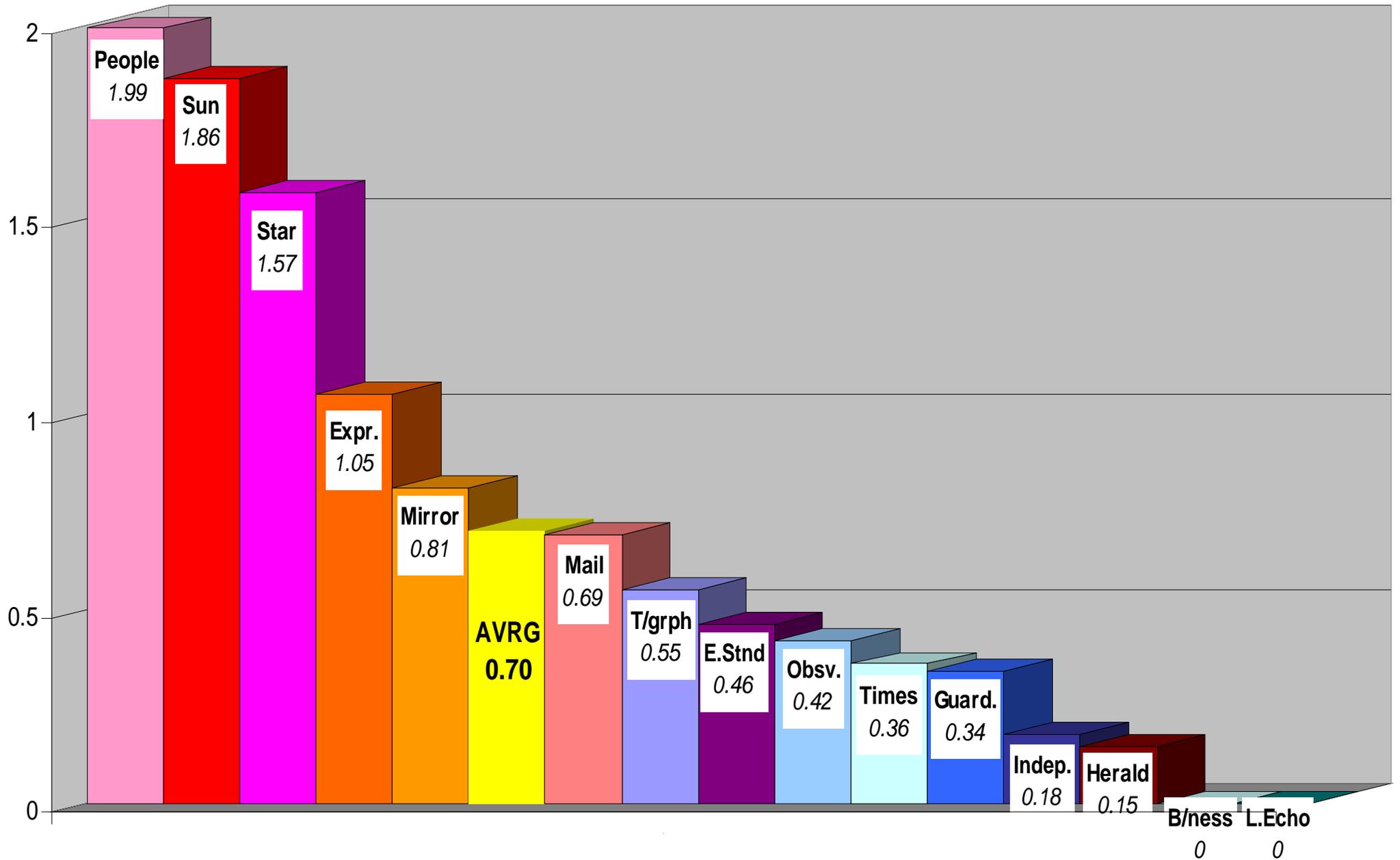## Total frequencies per million words

Sun 13.84
Star 8.15
Expr. 6.44
People 5.47
AVRG 3.00
Mirror 2.45
Mail 1.81
T/grph 1.68
E.Stnd 1.29
Indep. 0.90
Times 0.89
Guard. 0.74
Obsv. 0.73
Herald 0.57
B/ness 0
L.Echo 0

Illegal asylum seekers:
Total frequencies per 1000 articles

Illegal refugees:
Total frequencies per million words

Newspaper	Frequency
Star	5.10
Sun	4.72
People	4.10
Expr.	1.95
Mirror	1.63
AVRG	1.41
Mail	0.74
T/grph	0.67
E.Stnd	0.65
Herald	0.57
Times	0.43
Guard.	0.37
Obsv.	0.37
Indep.	0.21
B/ness	0
L.Echo	0

Illegal refugees:
Total frequencies per 1000 articles

Publication	Frequency
People	1.99
Sun	1.86
Star	1.57
Expr.	1.05
Mirror	0.81
AVRG	0.70
Mail	0.69
T/grph	0.55
E.Stnd	0.46
Obsv.	0.42
Times	0.36
Guard.	0.34
Indep.	0.18
Herald	0.15
B/ness	0
L.Echo	0

Illegal refugees/asylum seekers:
Total frequencies per million words

Source	Value
Sun	18.56
Star	13.25
People	9.57
Expr.	8.39
AVRG	4.40
Mirror	4.08
Mail	2.55
T/grph	2.35
E.Stnd	1.94
Times	1.32
Guard.	1.11
Indep.	1.11
Obsv.	1.10
Herald	0.73
B/ness	0
L.Echo	0

Illegal refugees/asylum seekers:
Total frequencies per 1000 articles

- Sun 7.33
- People 4.64
- Expr. 4.51
- Star 4.07
- Mail 2.39
- AVRG 2.22
- Mirror 2.03
- T/grph 1.93
- E.Stnd 1.37
- Obsv. 1.25
- Times 1.11
- Guard. 1.02
- Indep. 0.97
- Herald 0.66
- B/ness 0
- L.Echo 0

Illegal asylum seekers:
Annual frequencies per million words
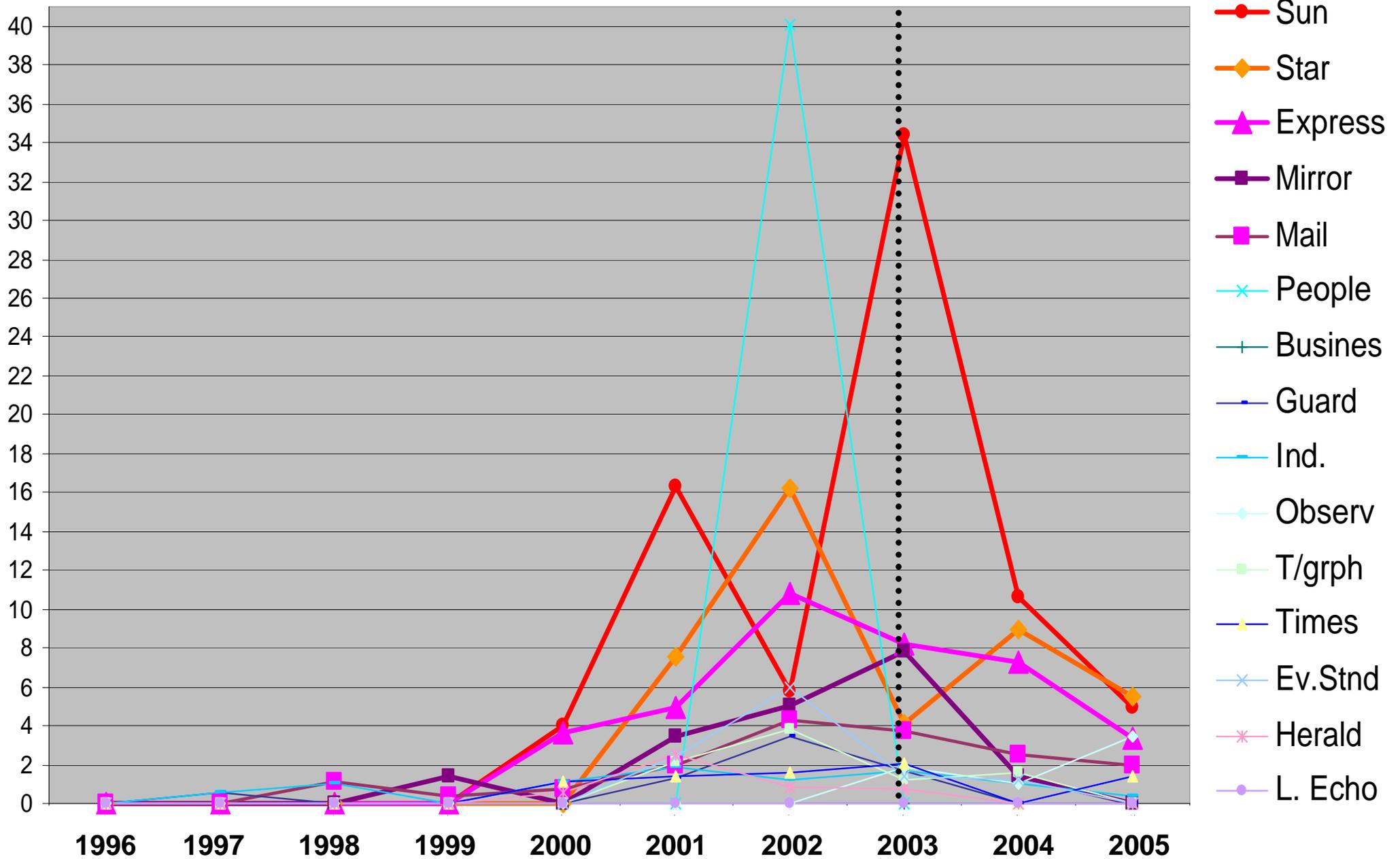
Illegal refugees:
Annual frequencies per million words
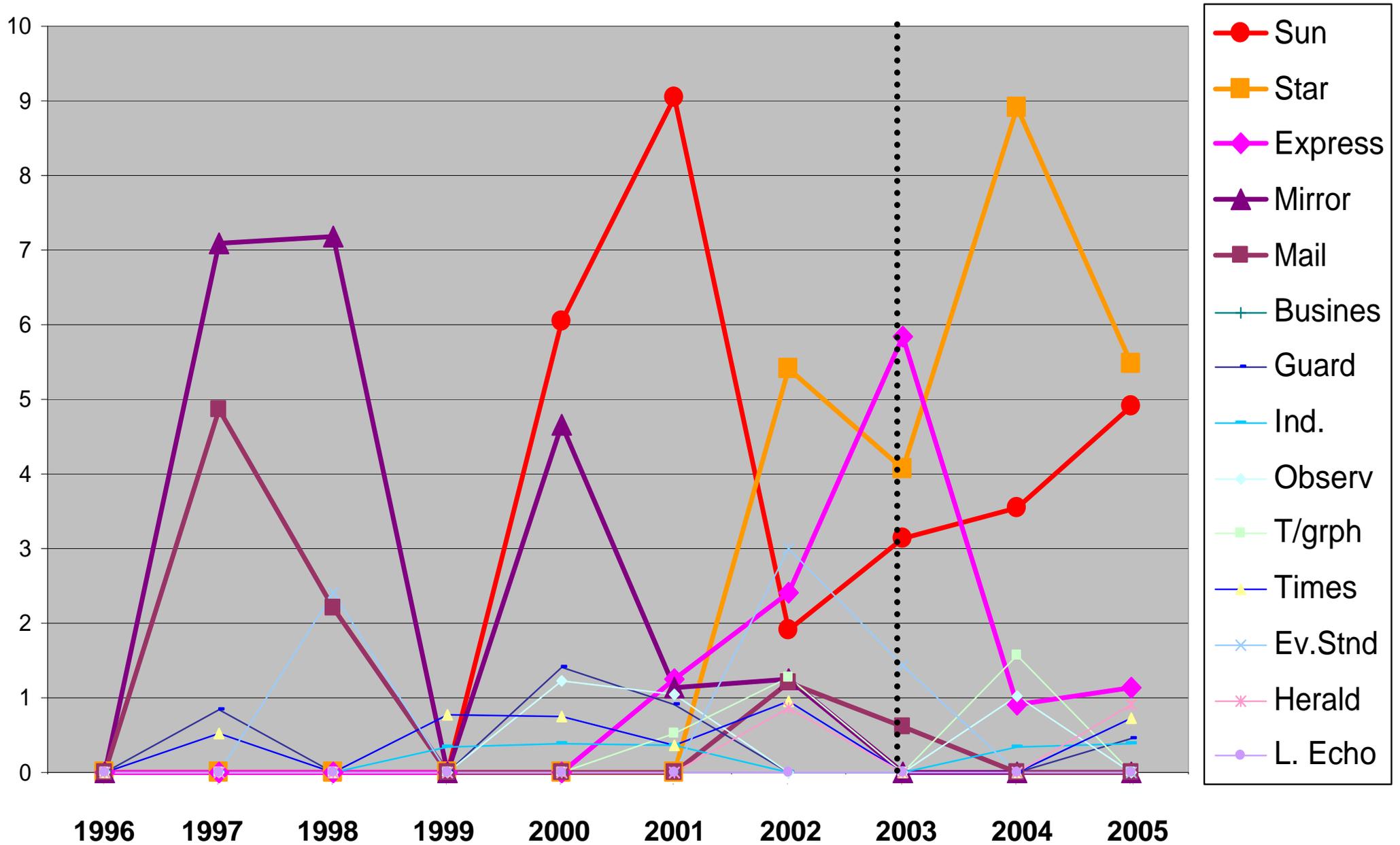
Tentative conclusions

❑ Discourse related to RASIM seems to revolve around a small number of topoi/topics/categories, most of which denote negative stance.

❑ Significant overlap of consistent collocates of RASIM, and strong intercollocations, can be interpreted as an indication of overlap in their usage, and, in some cases, their use as near synonyms.

❑ The quantitative analysis of the phraseology employed in expressing a strong stance (e.g. extreme negative bias) can provide a measure of the 'quality' of a newspaper.

Representation of refugees and asylum seekers in UK newspapers

Costas Gabrielatos and Paul Baker
Lancaster University

References

Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.

Baker, J.P. & McEnery, A.M. (2005). A corpus-based approach to discourses of refugees and asylum seekers. *Journal of Language and Politics,* 4(2), 197-226.

van Dijk, T.A. (1991). *Racism and the Press*. London and New York: Routledge

van Dijk, T.A. (2000). The reality of racism. On analyzing parliamentary debates on immigration. In Zurstiege, G. (ed.). *Festschrift für die Wirklichkeit.* Wiesbaden: Westdeutscher Verlag, 2000, 211-226.

Gabrielatos, C. & Baker, P. (2006). Representation of refugees and asylum seekers in UK newspapers: Towards a corpus-based comparison of the stance of tabloids and broadsheets. Paper given at the First International Conference: Critical Approaches to Discourse Analysis Across Disciplines (CADAAD 2006), University of East Anglia, Norwich, UK, 29-30 June 2006. (Available online: http://eprints.lancs.ac.uk/250).

Gabrielatos, C. (2006). Towards quantifying 'quality' in the press: Comparing the stance of UK broadsheets and tabloids towards refugees, asylum seekers and immigrants. Paper given at the joint meeting of the Corpus Research Group (CRG) and Language, Ideology and Power Research Group (LIP), Department of Linguistics and English Language, Lancaster University, UK, 13 June 2006. (Available online: http://eprints.lancs.ac.uk/231).

Hardt-Mautner, G. (1995). Critical discourse analysis and corpus linguistics. UCREL Technical Papers, 6.

Hunston, S. (20020. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Baker, M., Francis, G. & Tognini-Bonelli, E. (eds.) *Text and Technology: In honour of John Sinclair*, Philadelphia and Amsterdam: John Benjamins, 157-176.

McEnery, T. (2006). *Swearing in English: Bad language, purity and power from 1586 to the present*. London and New York: Routledge.

Nattinger, J. R. & DeCarrico, J. S. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

Orpin, D. (2005). Corpus linguistics and critical discourse analysis: Examining the ideology of sleaze. *International Journal of Corpus Linguistics,* 10(1), 37-61.

Phillips, M. (1989). *Lexical Structure of Text*. Discourse Analysis Monograph no. 12. English Language Research, University of Birmingham.

Reisigl, M. & Wodak, R. (2001). *Discourse and Discrimination.* London: Routledge

Sedlak, M. (2001). You really do make an unrespectable foreign policy. In Wodak, R. & van Dijk, T.A. (eds.) *Racism at the Top: Parliamentary discourses on ethnic issues in six European states*. Drava Verkag, 107-168.

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

Sotillo, S.M. & Wang-Gempp, J. (2004). Using corpus linguistics to investigate class, ideology, and discursive practices in online political discussions: Pedagogical applications of corpora. In Connor, U. & Upton, T.A. (eds.) *Applied corpus linguistics*, 91-122.

Stubbs, M. (2002). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.

ter Wal, J. (2002). Racism and cultural diversity in the mass media: An overview of research and examples of good practice in the EU Member States, 1995-2000. European Research Centre on Migration and Ethnic Relations (ERCOMER).

van Dijk, T. (2000). On the analysis of parliamentary debates on immigration. In Reisigl, M. & Wodak, R. (eds) *The Semiotics of Racism*. Vienna: Passagen Verlag, 85-104.

van Leeuven, Theo. (1996). The representation of social actors. In: Carmen-Rrosa Caldas-Coulthard and Malcolm Coulthard (eds.). *Texts and Practices. Readings in Critical Discourse Analysis*. London: Routledge, 32-70.

Wodak, R. & van Dijk, T.A. (eds.) *Racism at the Top: Parliamentary discourses on ethnic issues in six European states*. Drava Verkag.