

Siena English Language and Linguistics Seminars (SELLS),  
University of Siena, 9 November 2009

**Corpus-based methodology  
and critical discourse studies  
Context, content, computation**

Costas Gabrielatos  
Lancaster University

# Main criticisms of CL

**CL does not take account of the relevant context**

**CL does not examine sufficient amount of (co-)text**  
**lists of words (frequency, keywords, collocates)**  
**short concordance lines**

**Also**

**Nature / definition of CL**

**Objectivity – Subjectivity**

**Replicability**

(see also Marchi & Taylor, forthcoming; Partington, 2009; Taylor, 2008)

# Projects

## Current:

*The representation of Islam and Muslims in the UK press, 1998-2008.* ESRC. Sept. 2009 – Aug. 2010.

- PI: Paul Baker; CI: Tony McEnery; RA: Costas Gabrielatos.
- Corpus: 200,000 articles; 140 million words (and counting)

## Completed:

*Discourses of refugees and asylum seekers in the UK press, 1996-2006.* ESRC. Oct. 2005 – March 2007.

- PI: Paul Baker; CIs: Tony McEnery, Ruth Wodak; RAs: Costas Gabrielatos (CL), Majid KhosraviNik (CDA), Michal Krzyzanowski (CDA).
- Corpus: 175,000 articles; 140 million words
- <http://ucrel.lancs.ac.uk/projects/rasim/>

**CL does not take account of the relevant context**

**CL researchers have no less access to sources of relevant contextual information than CDS researchers.**

**A non-linguistic quantitative analysis of a corpus reveals patterns which ...**

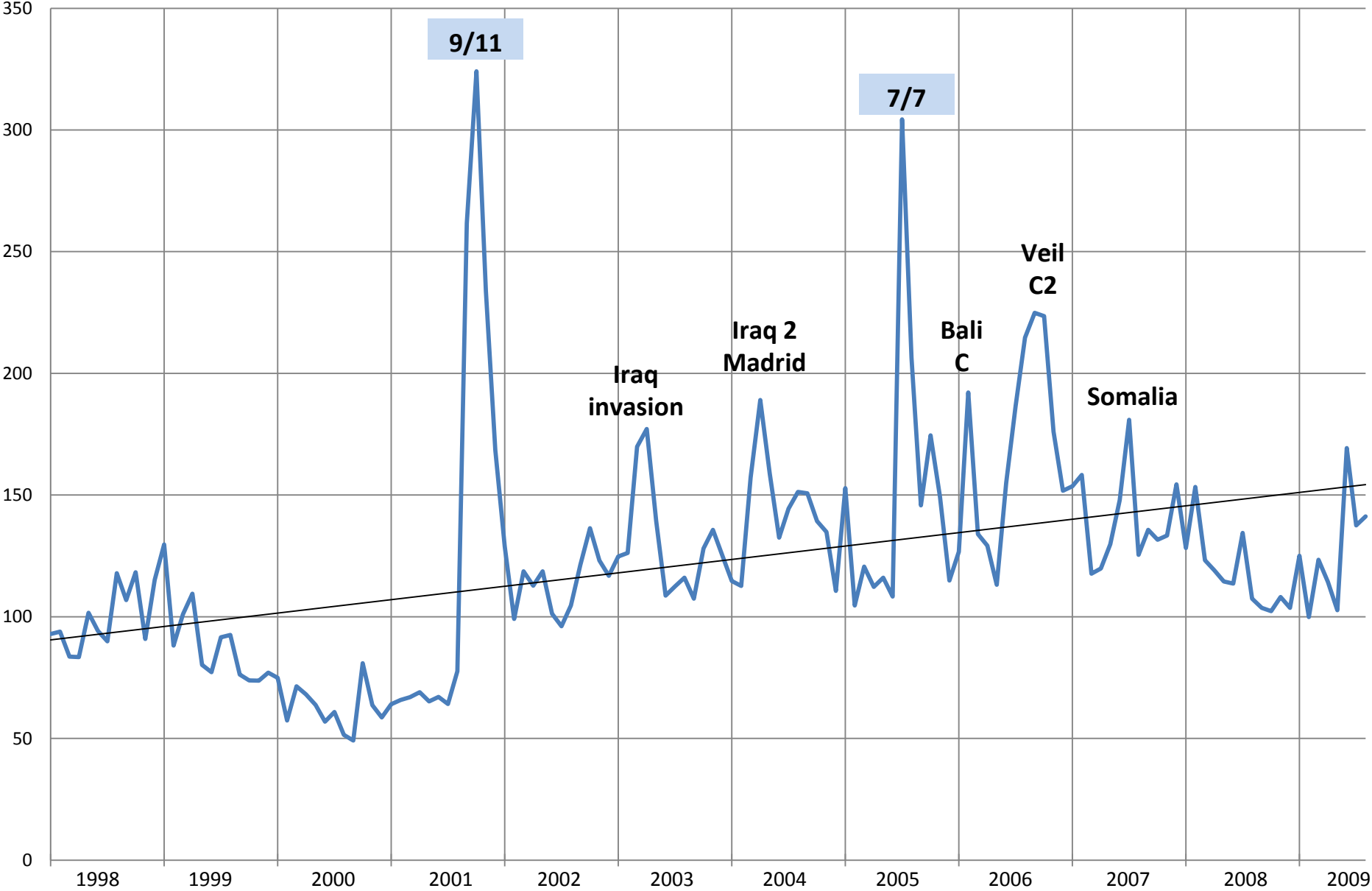
- ... pinpoint periods/sources/texts that can be usefully examined in detail.**
- ... uncover helpful contextual elements.**

# Revealing contextual elements 1: Reaction to trigger events

## Islam Corpus Query

- Alah OR Allah OR ayatolah OR ayatollah OR burka\* OR burqa\* OR chador\* OR fatwa\* OR hejab\* OR imam\* OR islam\* OR Koran OR Mecca OR Medina OR Mohammedan\* OR Moslem\* OR Muslim\* OR mosque OR mufti\* OR mujaheddin\* OR mujahedin\* OR mullah\* OR muslim\* OR Prophet Mohammed OR Q'uran OR rupoush OR rupush OR sharia OR shari'a OR shia! OR shi-ite\* OR Shi'ite\* OR sunni\* OR the Prophet OR wahabi OR yashmak\* AND NOT Islamabad AND NOT shiatsu AND NOT sunnily
- **Graph depicting number of articles per month**
- **Establishing events coinciding with / triggering spikes.**
- **Extent of change in number of articles due to trigger events.**
  - **UK press**
  - **Individual newspapers**

# National UK newspapers: average number of articles /month



Is the general picture  
representative of  
all newspapers?

# Yes

- Four spikes shared by at least two-thirds of the 12 newspapers:
  - 9/11, 7/7: 12
  - Veil + Cartoons 2: 11
  - Bali bombings + Cartoons: 9
- All newspapers but one show an upward trend.

# No

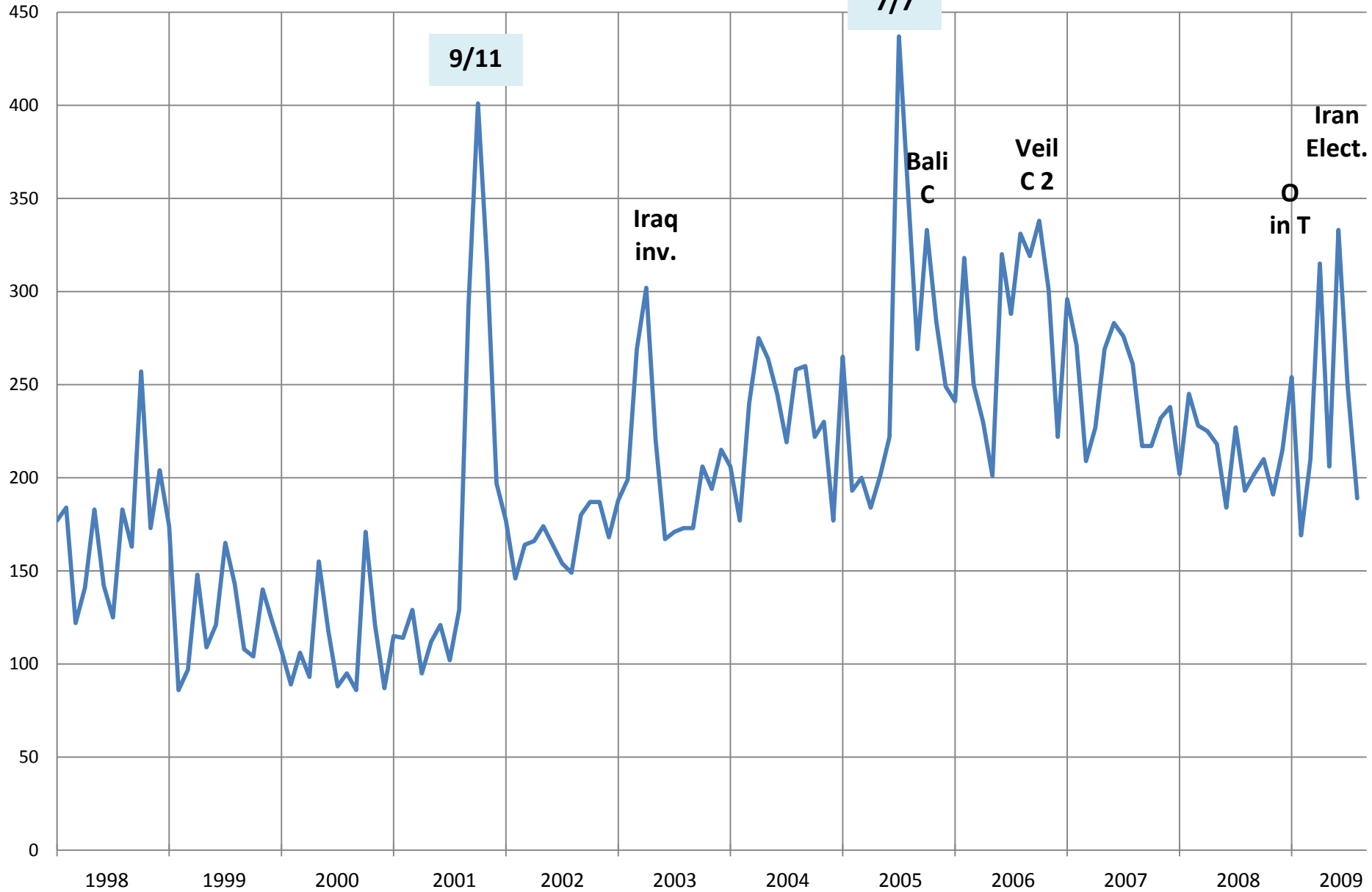
- 19 spikes collectively – only 5 shared by more than half!
- Different relative importance of primary and/or secondary spikes
- Five groups in terms of primary spikes:
  - 9/11 & 7/7
  - 9/11
  - 7/7 + other
  - Other + 9/11 & 7/7
  - Other



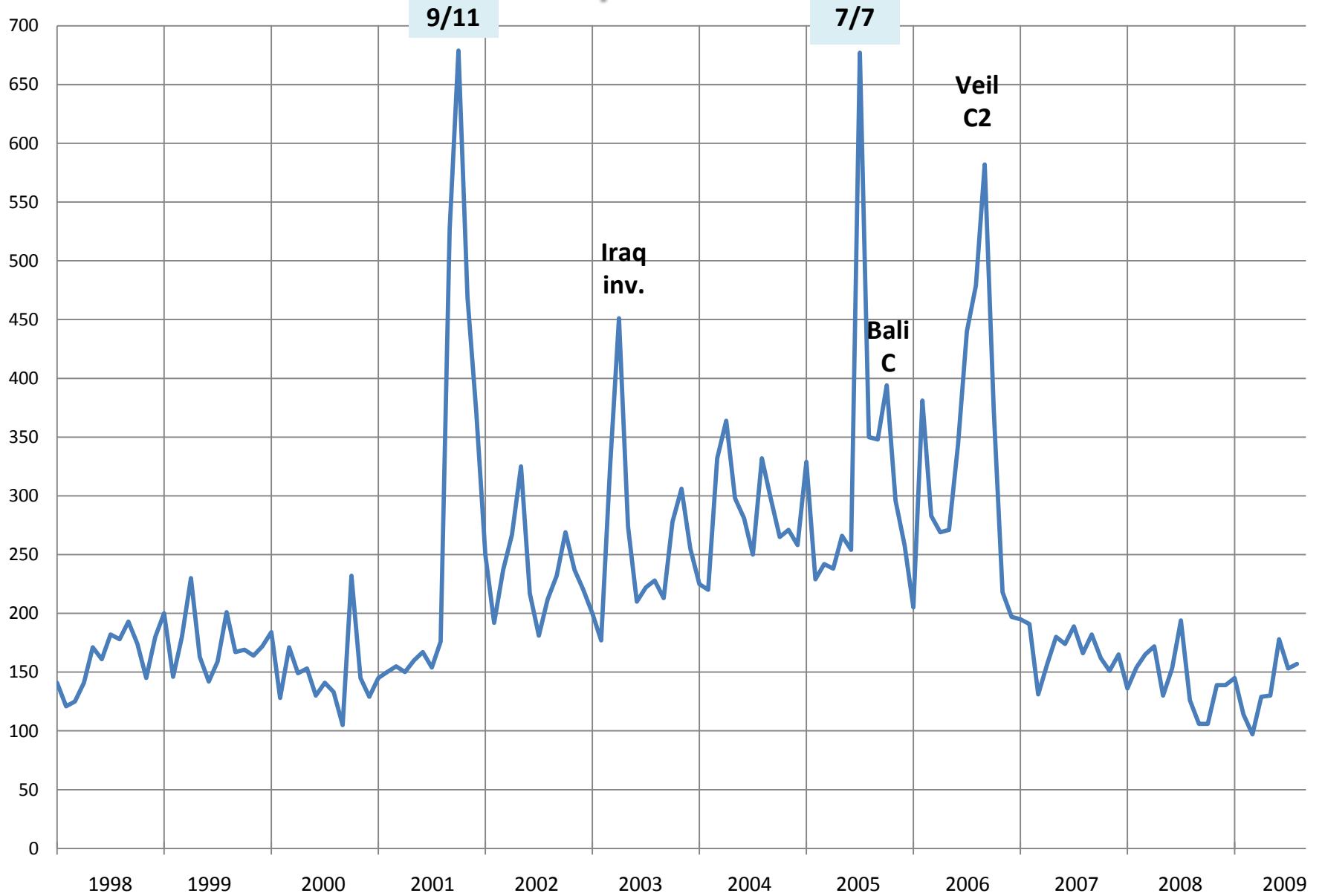
9/11 & 7/7

3 newspapers

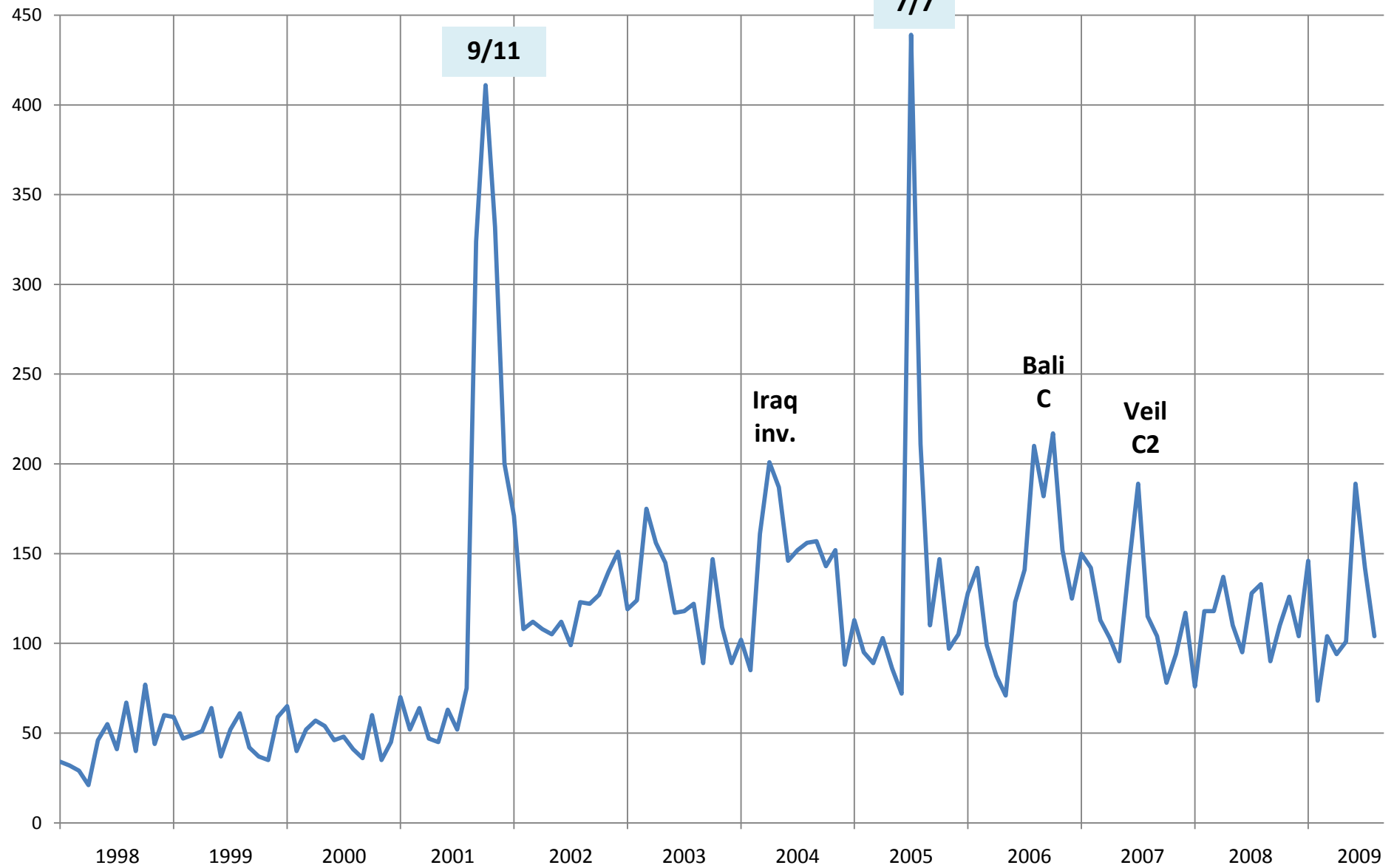
# Guardian



# *Independent*



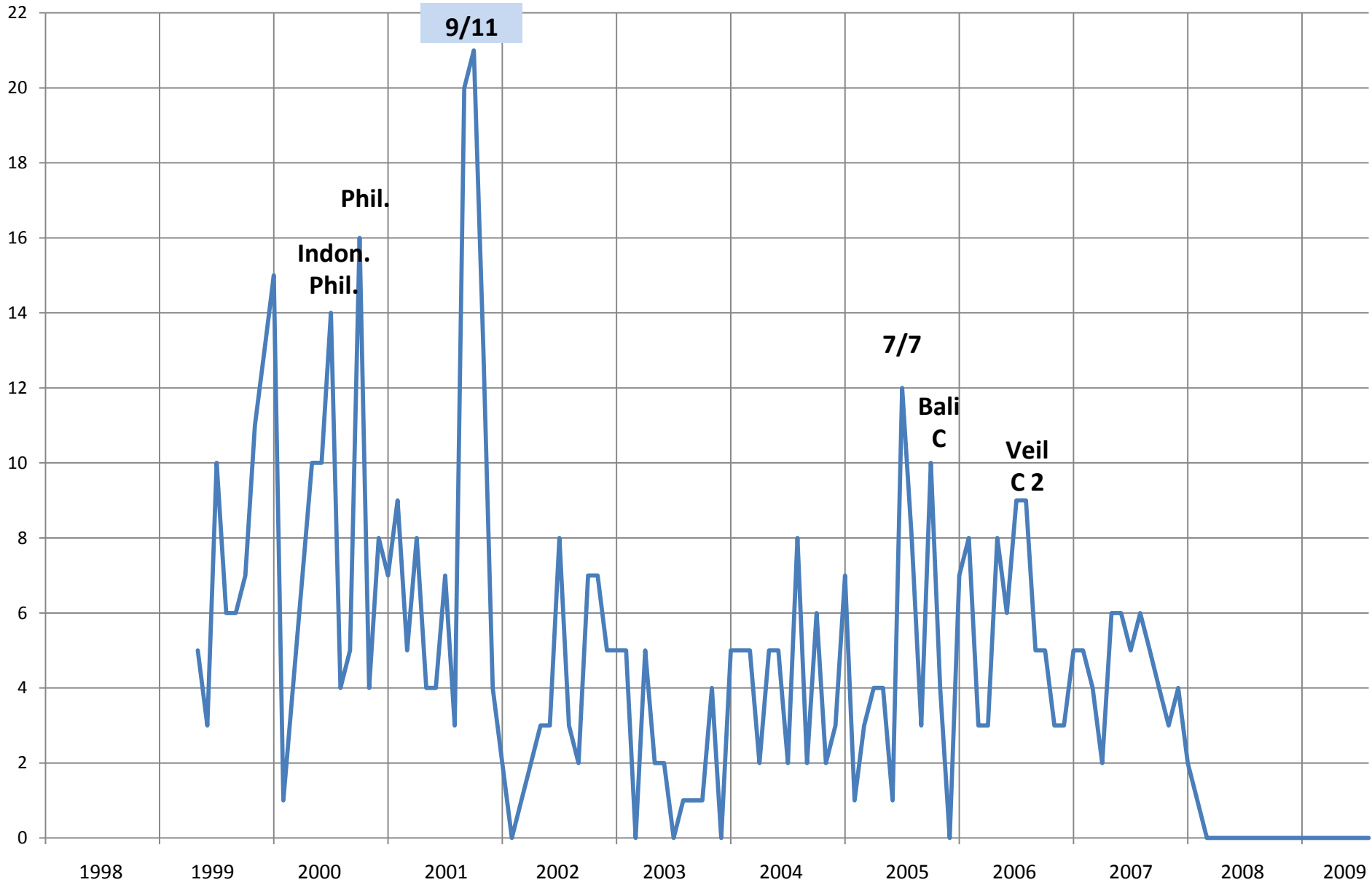
# Mirror



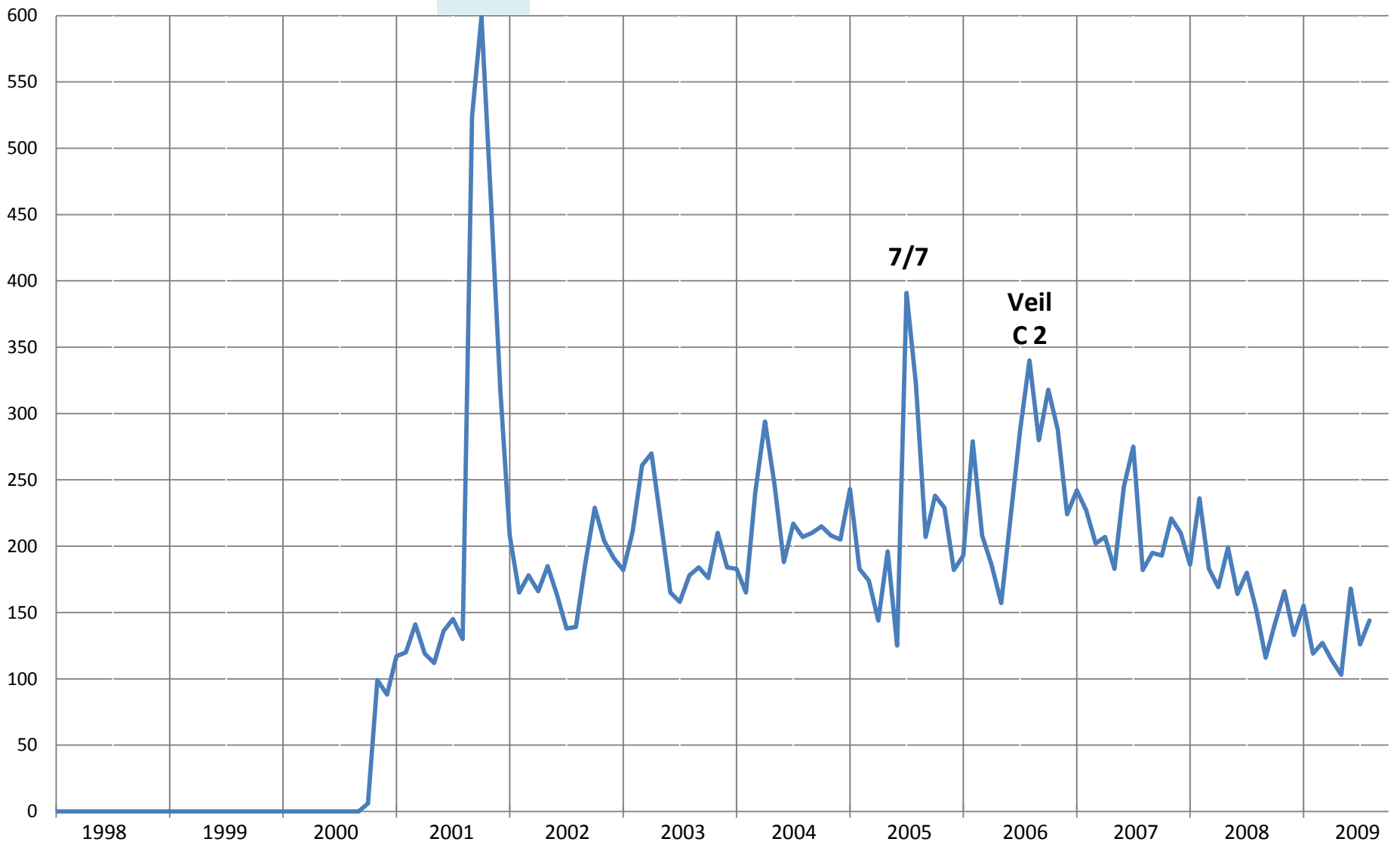
9/11

4 newspapers

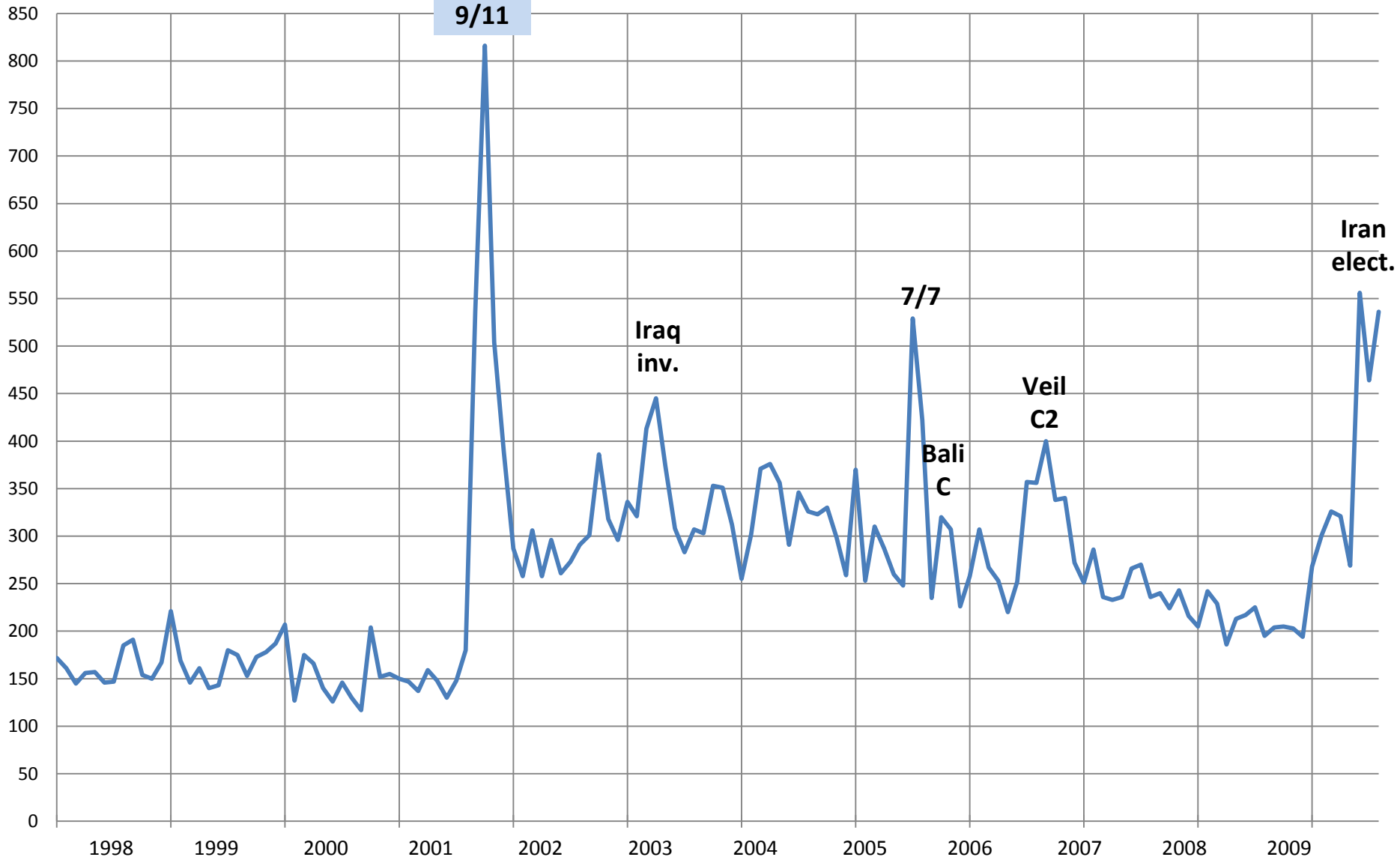
# *Business*



# Telegraph

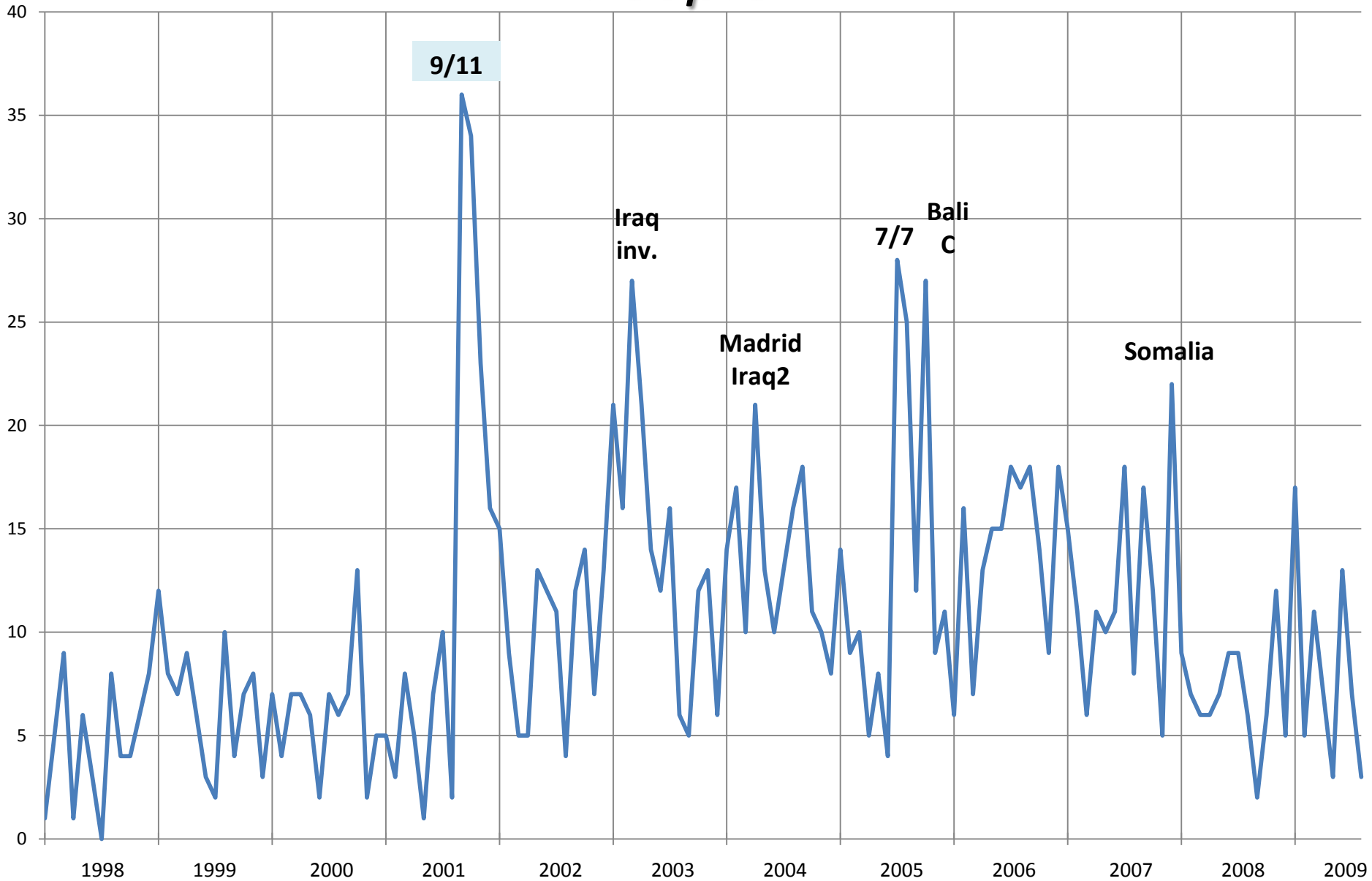


# Times





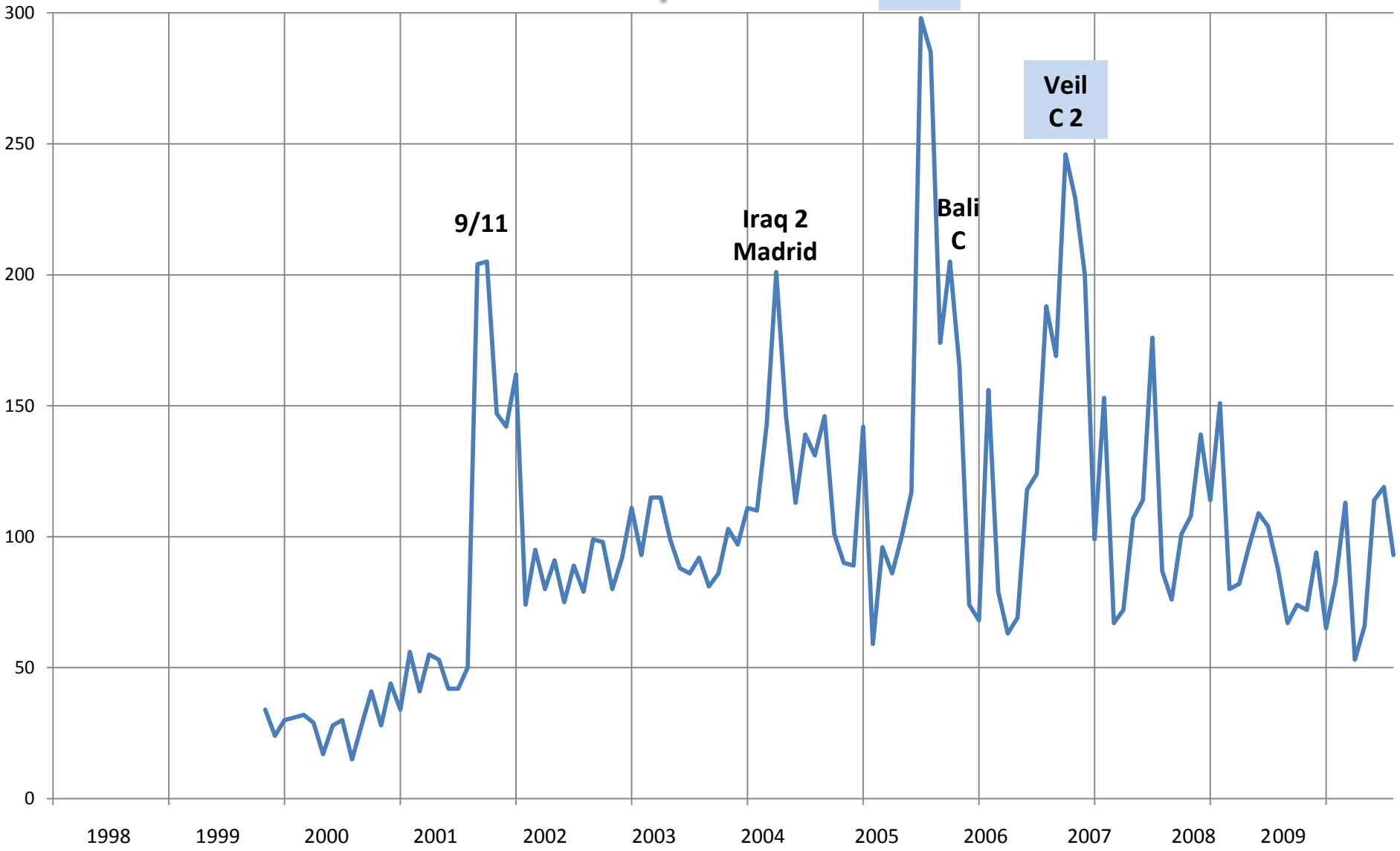
# People



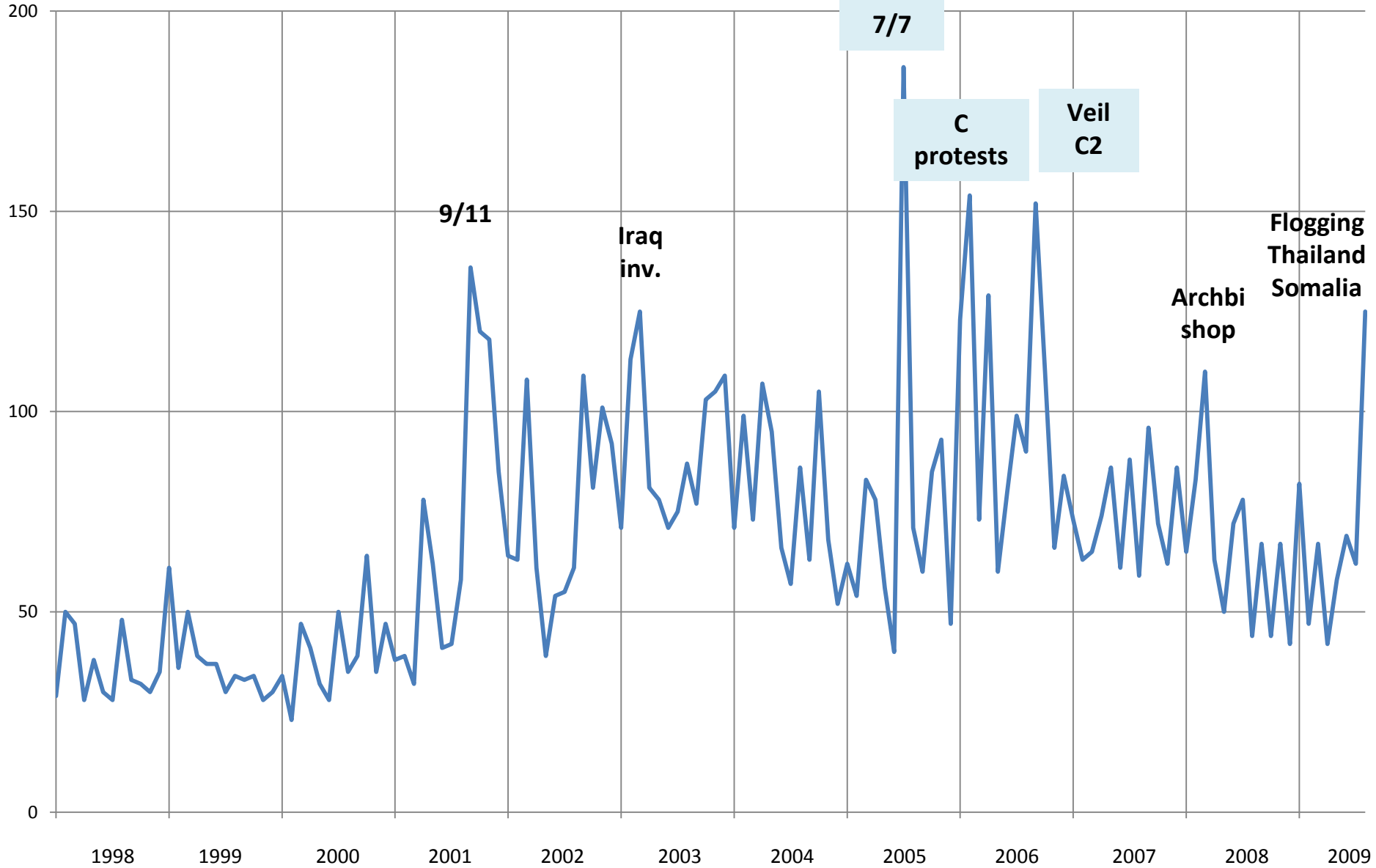
*7/7*  
+  
other

2 newspapers

# *Express*



# Observer



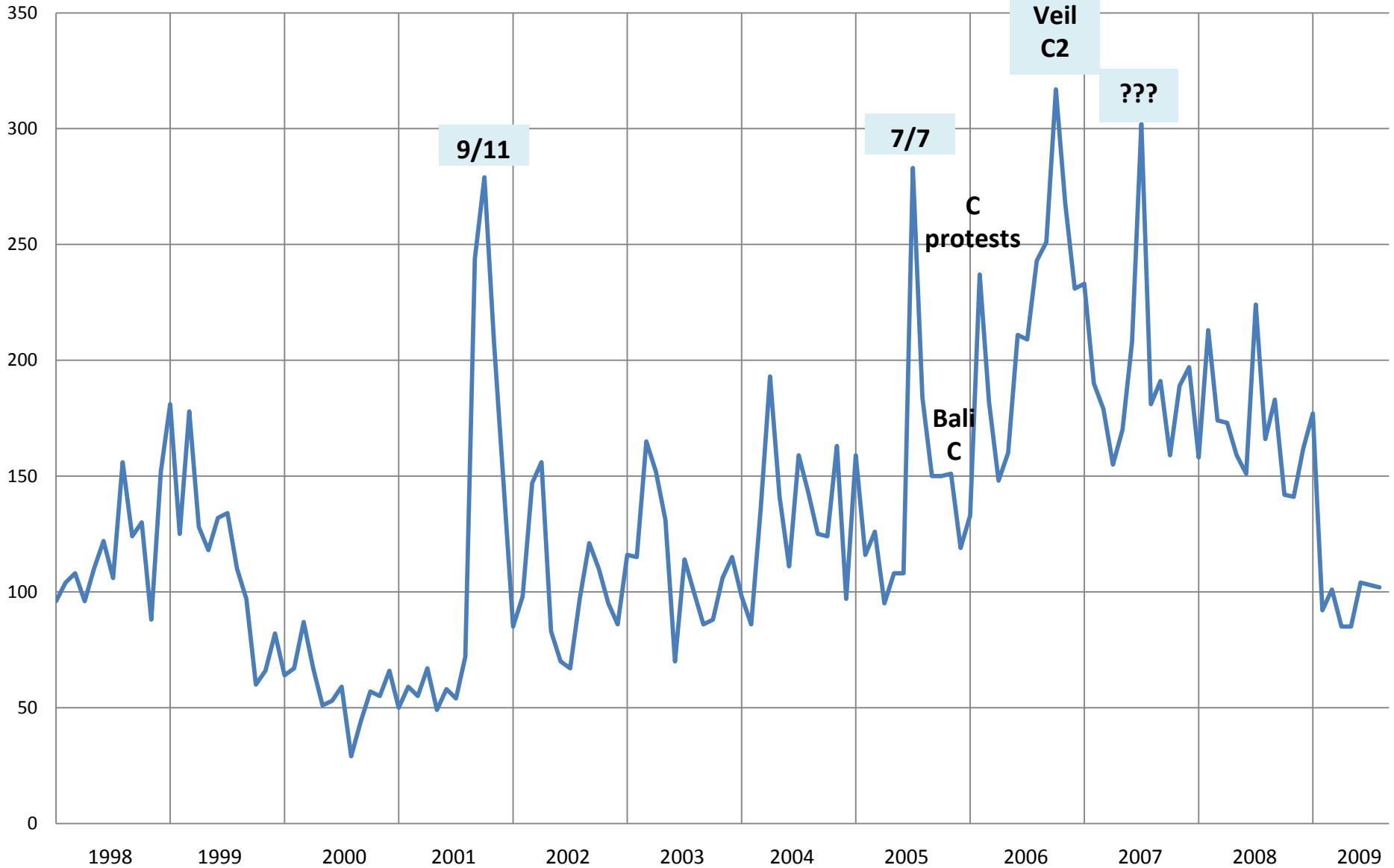
Other

+

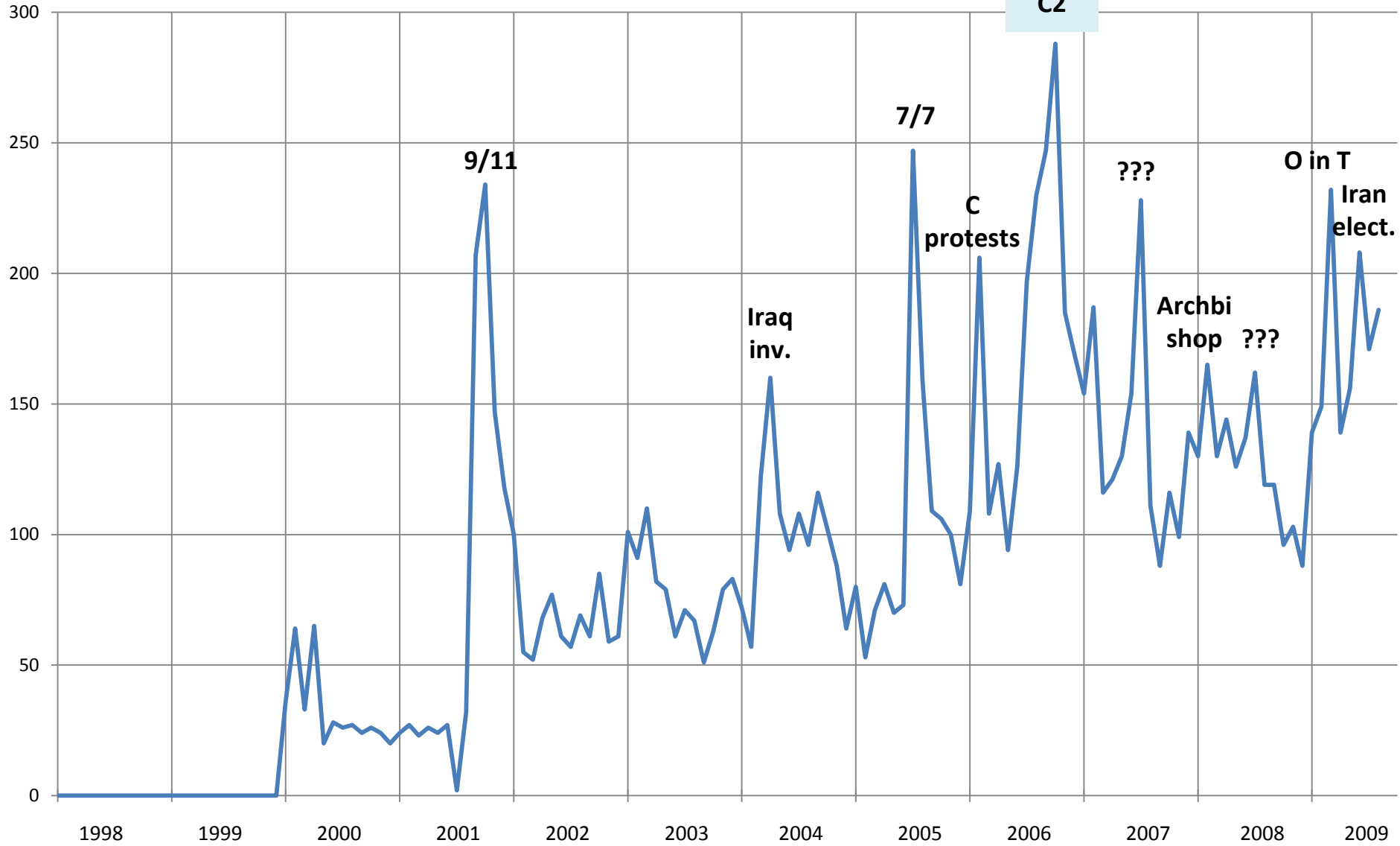
9/11 & 7/7

2 newspapers

# Mail



# Sun

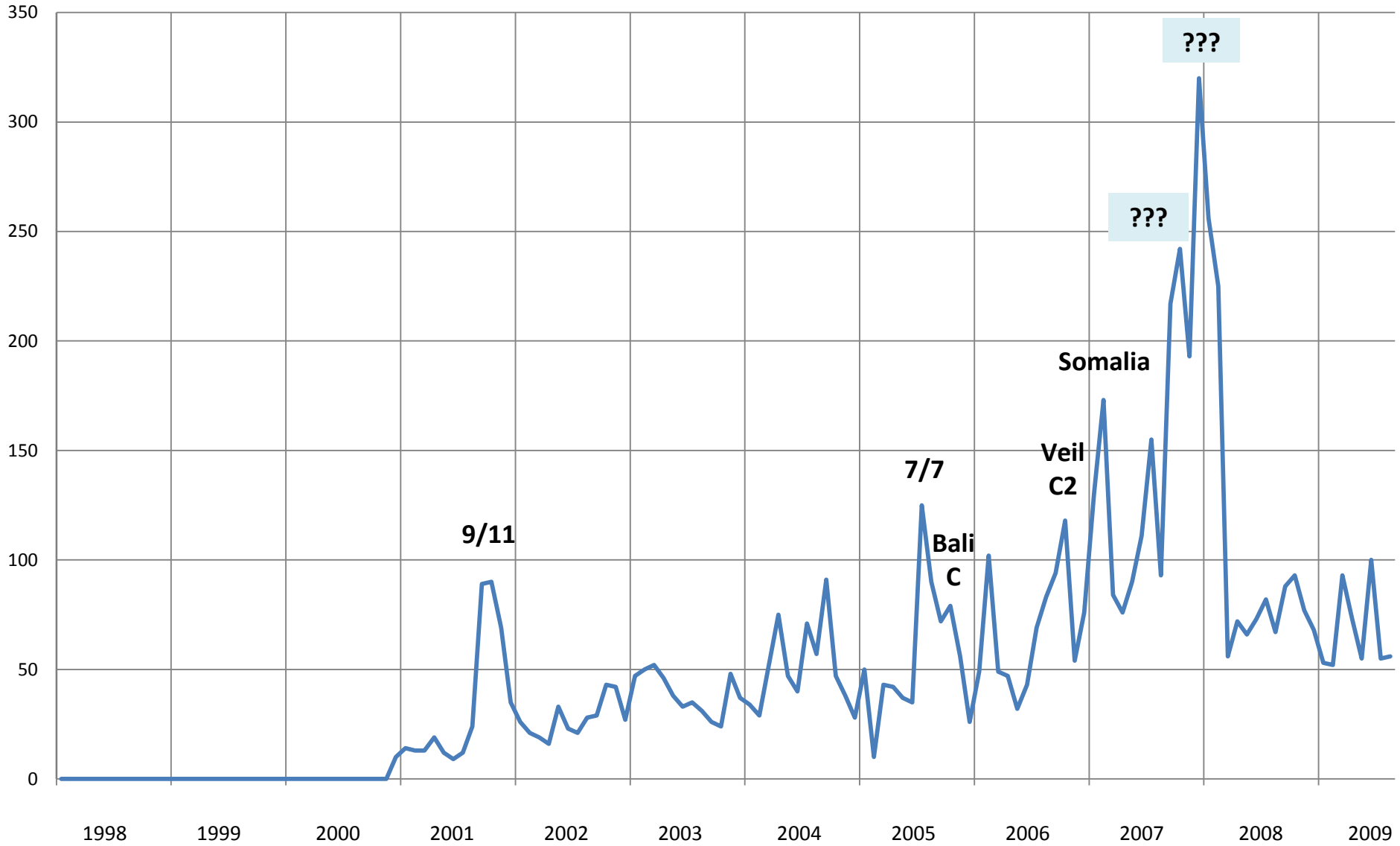


Other

1 newspaper



# Star



Can we measure  
a newspaper's response  
to a given trigger event?

# Change in number of topic-related articles

## Trigger event: 9/11

Diff. % of number of articles

### Reaction

The extent to which the number of articles changed immediately after a trigger event

Average of 12 months pre-9/11

9/11 spike (avg. Sept.-Oct. 2001)

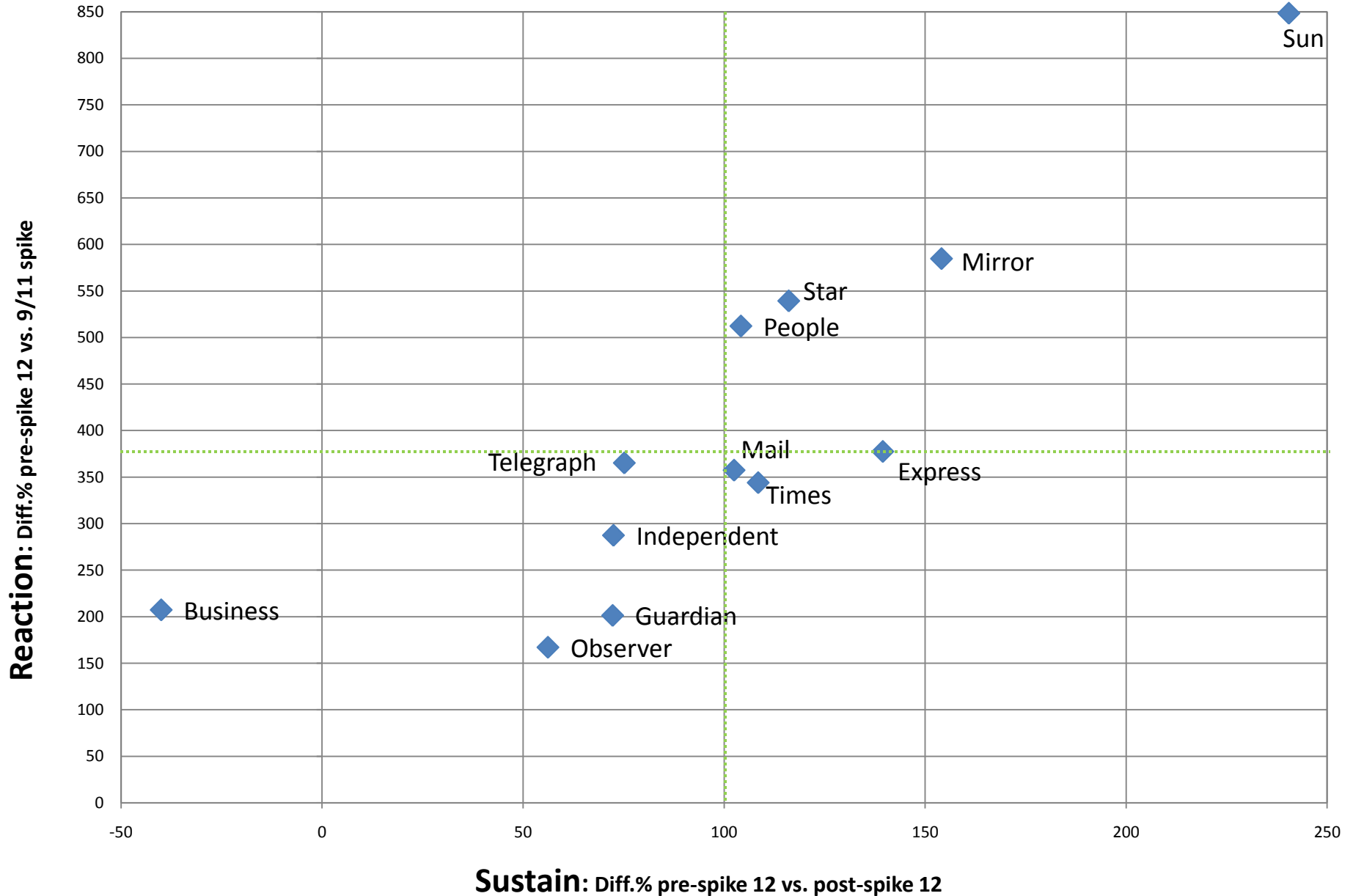
### Sustain

The extent to which the change was sustained a year after the trigger event

Average of 12 months pre-9/11

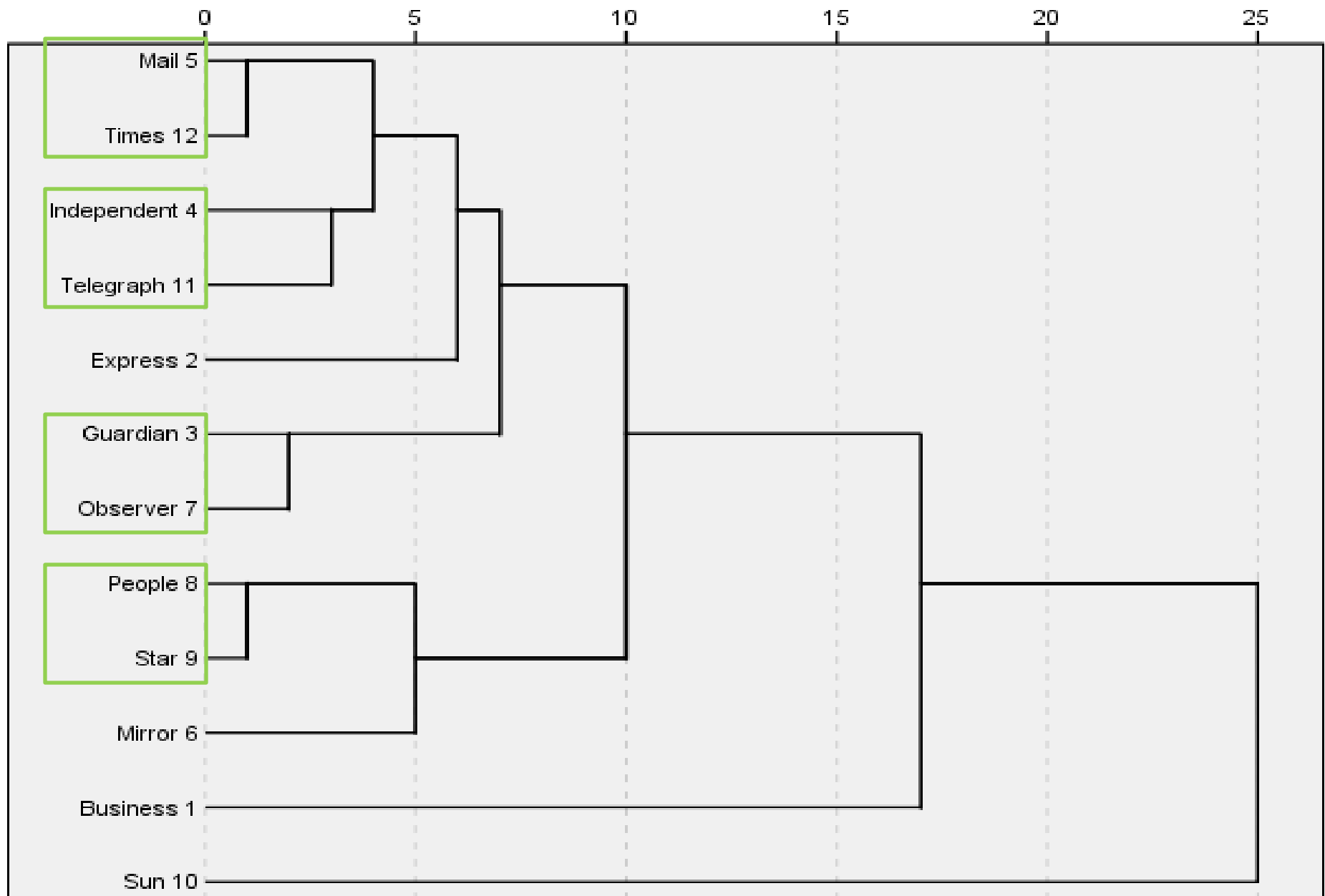
Average of 12 months post-9/11

# 9/11 and Islam/Muslims: Reaction and Sustain

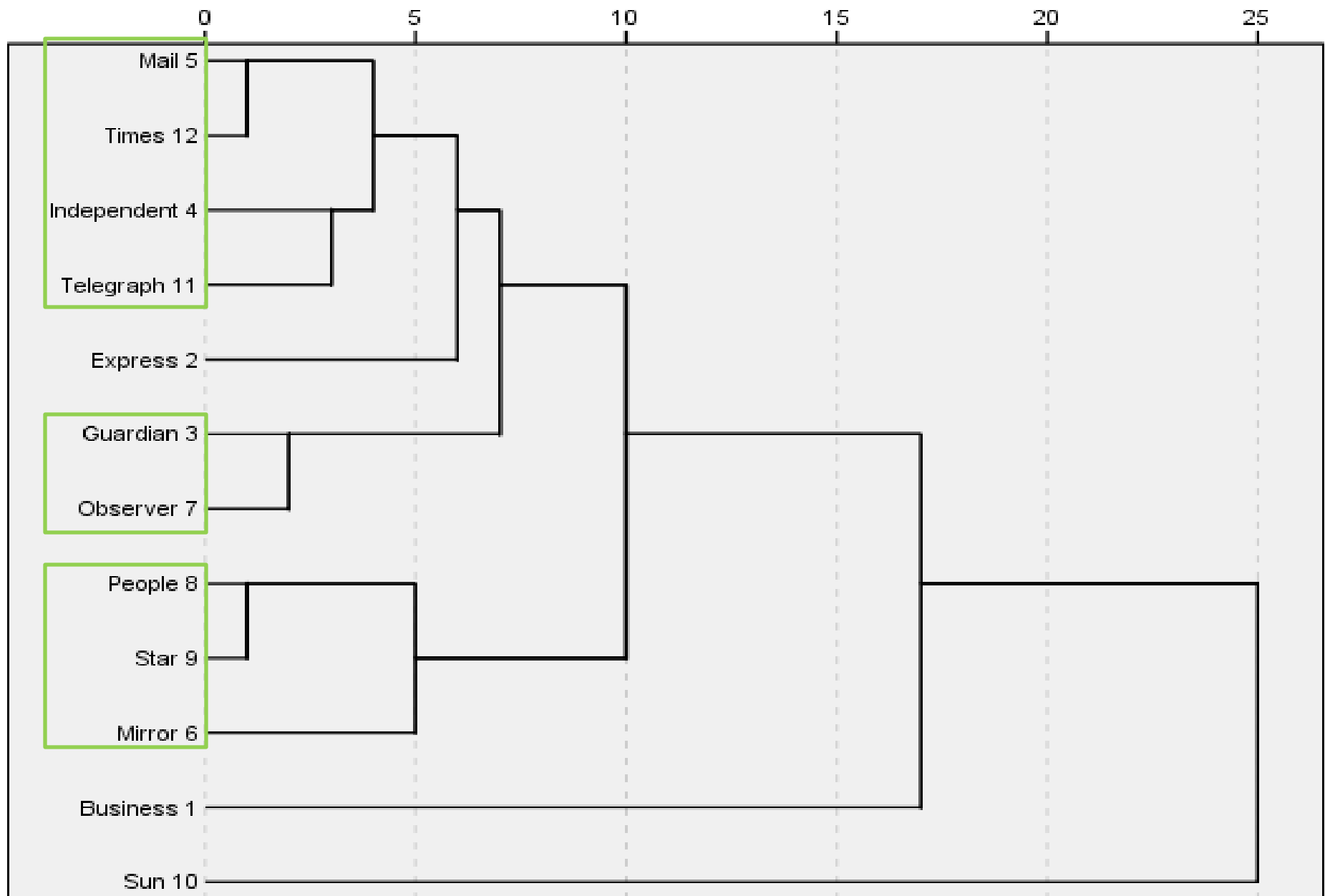


What about the  
broadsheets-tabloids  
distinction?

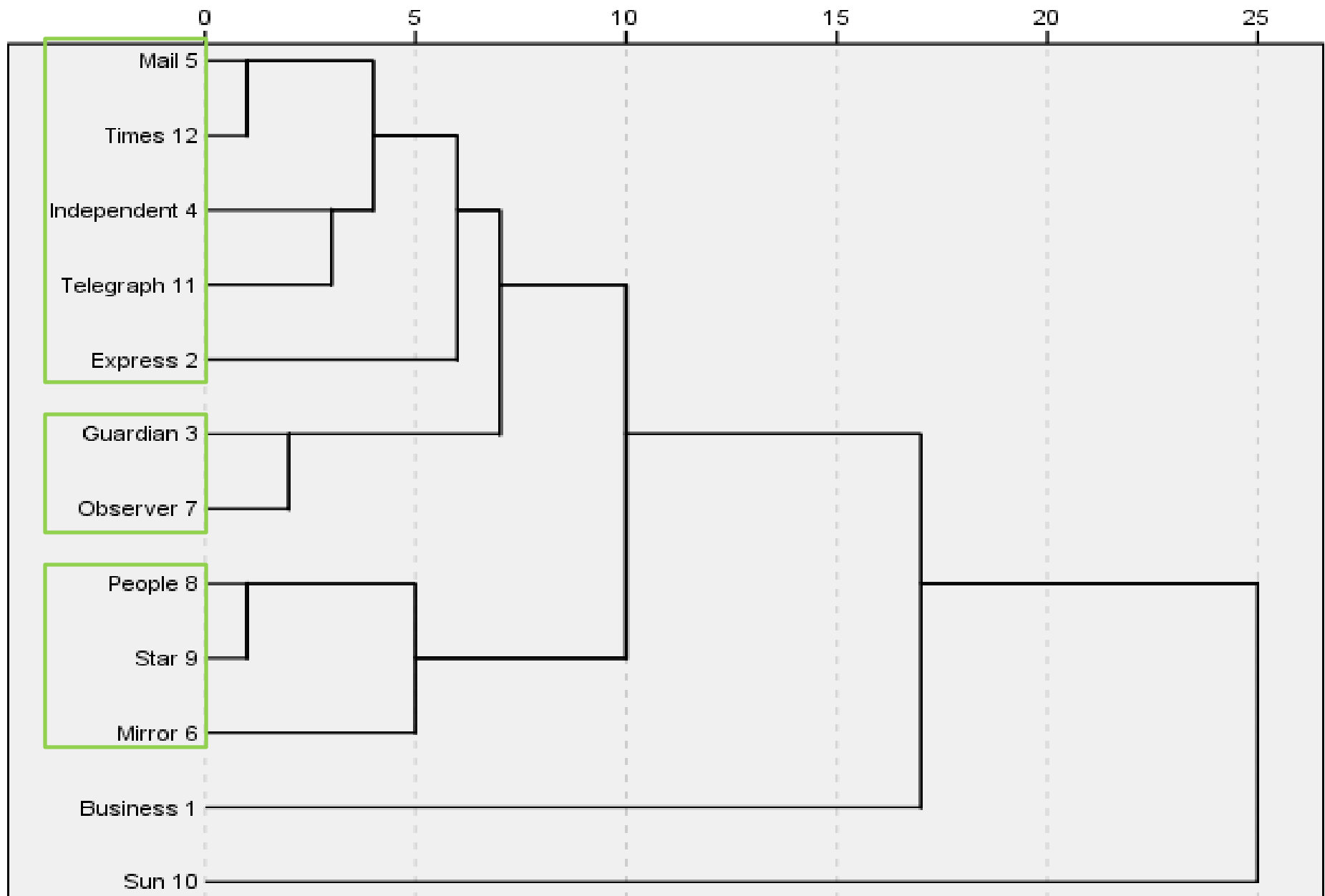
# 9/11, reaction and sustain: clusters



# 9/11, reaction and sustain: clusters

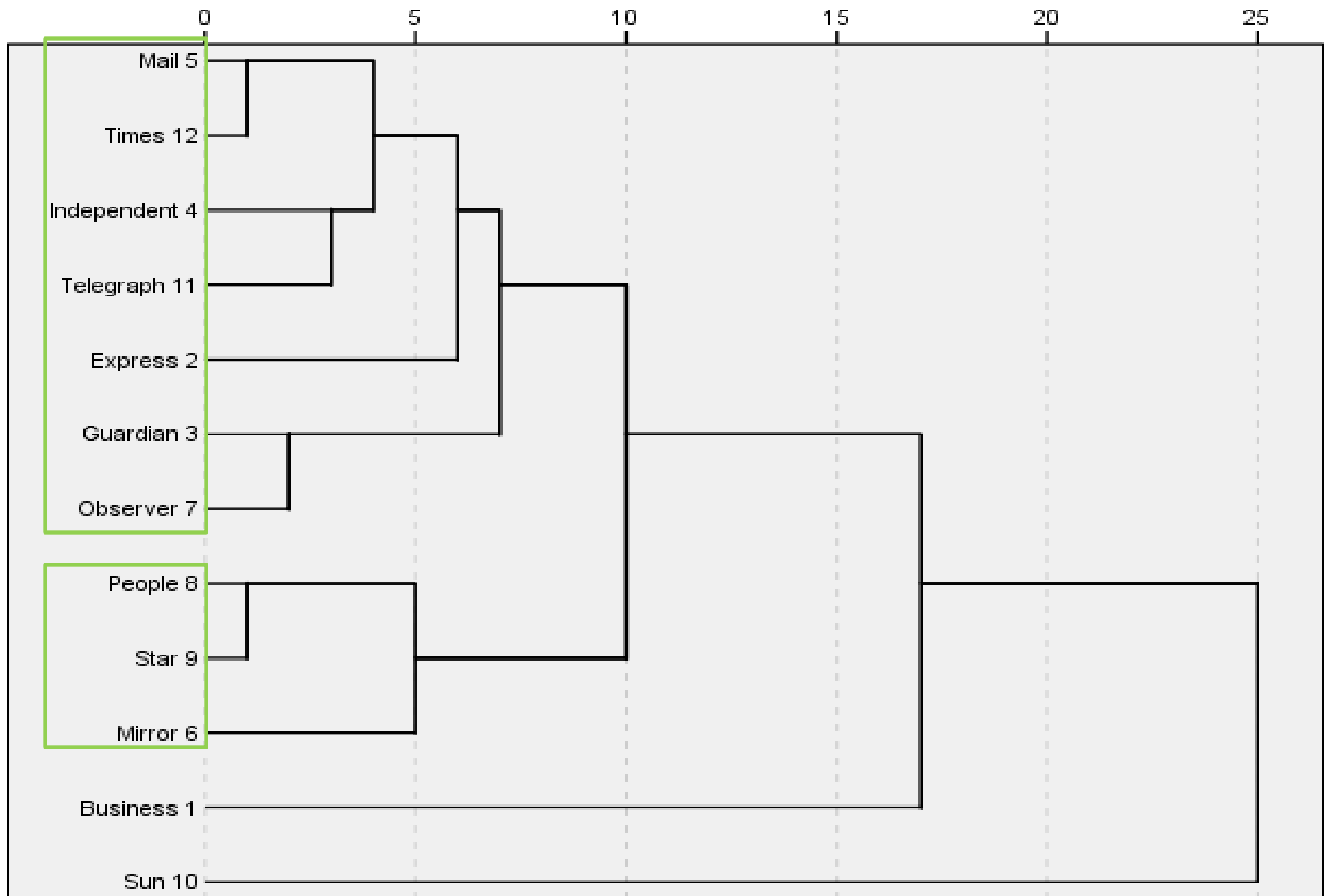


# 9/11, reaction and sustain: clusters

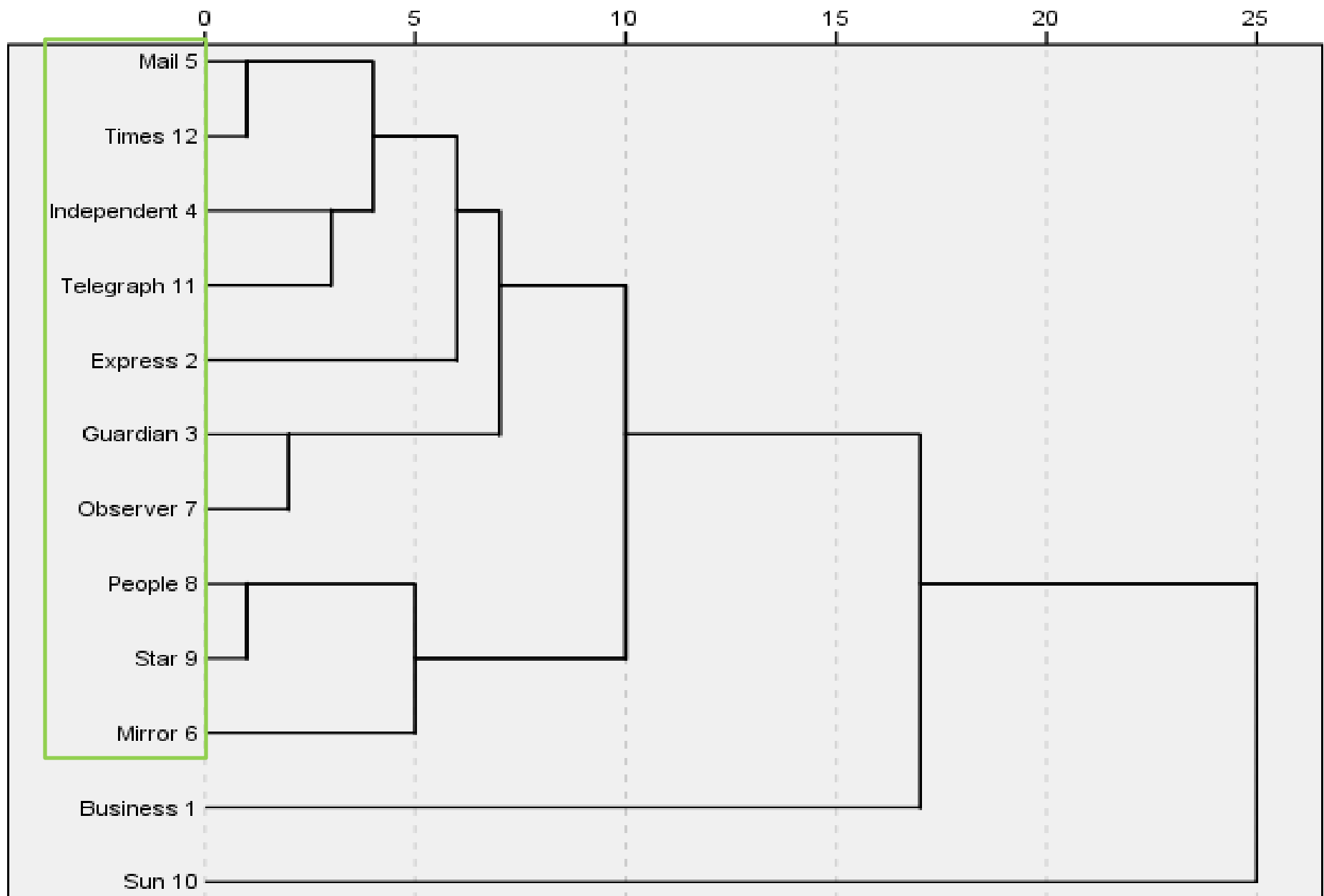




# 9/11, reaction and sustain: clusters



# 9/11, reaction and sustain: clusters



# Revealing contextual elements 2: Use of loaded terms

## RASIM Corpus 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 department (see Gabrielatos, 2007)

## Nonsensical (= loaded) terms

*illegal/legal*      *refugee\*/asylum seeker\**  
*bogus/genuine*    *immigrant\*/migrant\**

[definitions](#)

- **Which newspapers use them?**
- **How frequently?**
  - Per million words
  - Per thousand articles(see also Baker et al., 2008)

# Definitions

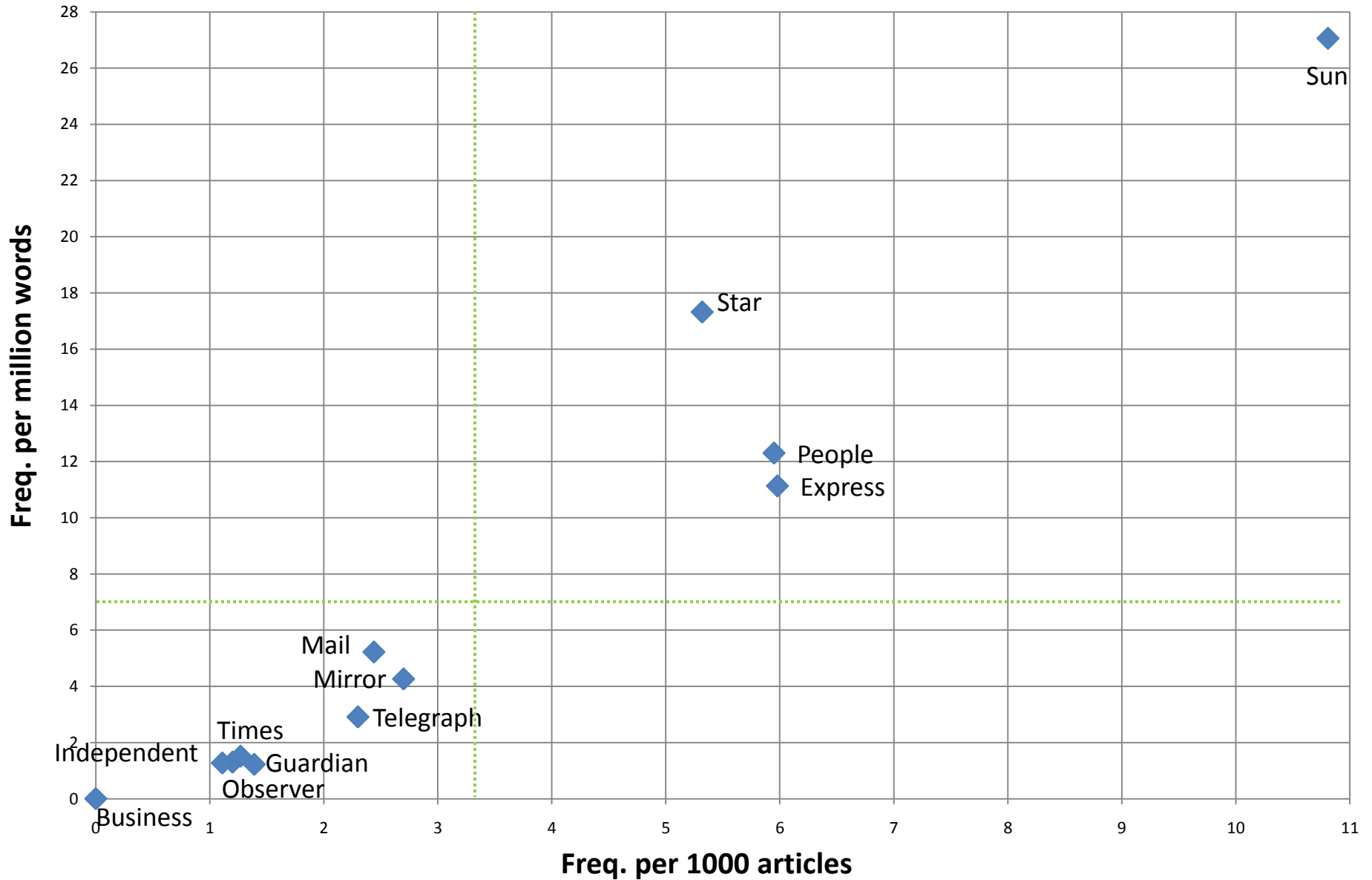
	Longman Dictionary of Contemporary English (2003)	Refugee Council
<i>refugee</i>	Someone who has been <b>forced</b> to leave their country, especially during a war, or for political or religious reasons.	Someone whose <b>asylum application has been successful</b> and who is allowed to stay in another country having proved they would face persecution back home.
<i>asylum seeker</i>	Someone who leaves their own country because <b>they are in danger</b> , especially for political reasons, and who asks the government of another country to allow them to live there.	Someone who has fled persecution in their homeland, has arrived in another country, <b>made themselves known to the authorities</b> and <b>exercised the legal right to apply for asylum</b> .
<i>immigrant</i>	Someone who enters another country to <b>live there permanently</b> .	-----
<i>migrant</i>	Someone who goes to live in another area or country, especially in order to find <b>work</b> .	<b>[economic migrant]</b> Someone who has moved to another country to <b>work</b> .

## International Association for the Study of Forced Migration

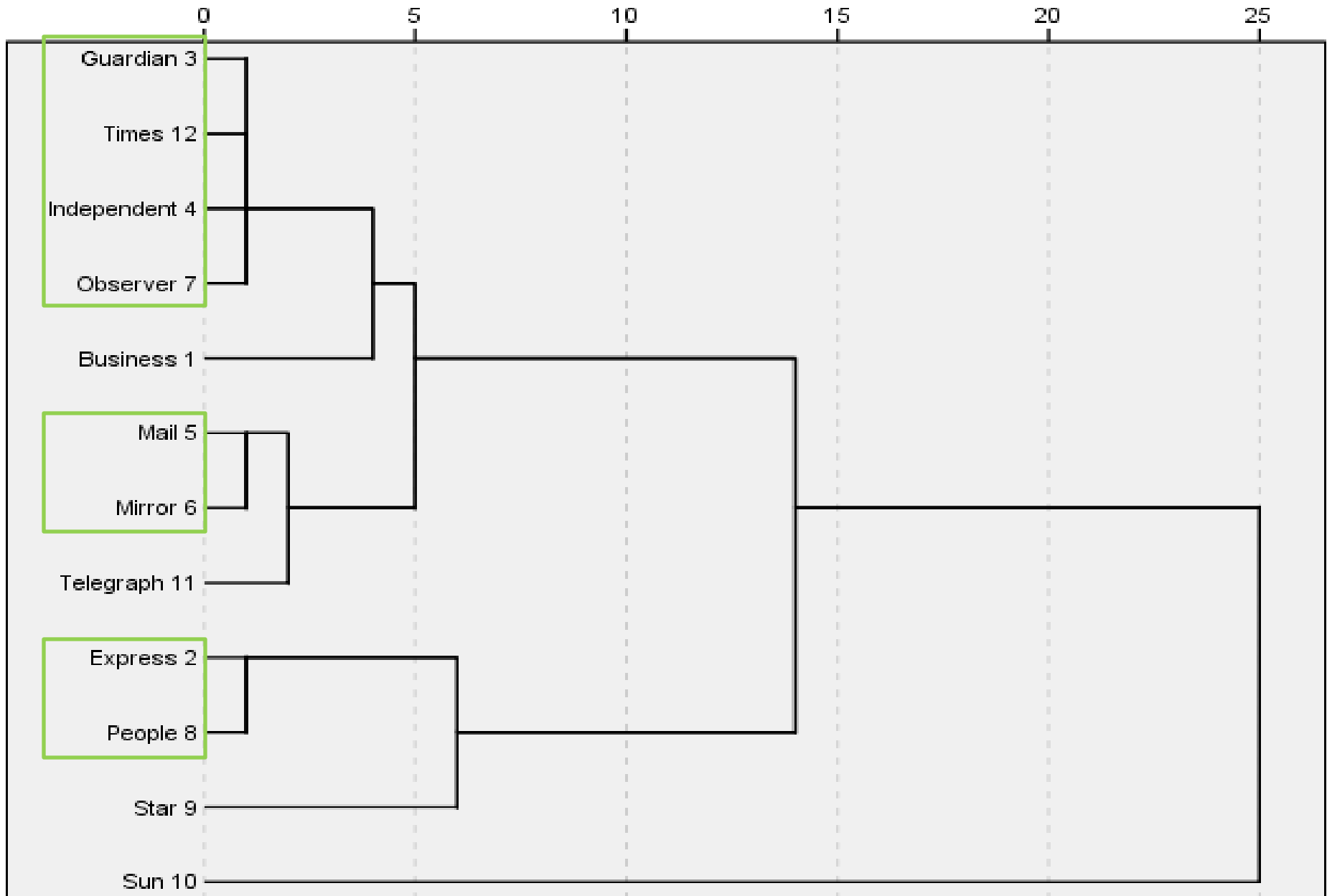
**Forced migration:** refugees and asylum seekers

**Voluntary migration:** immigrants and (economic) migrants

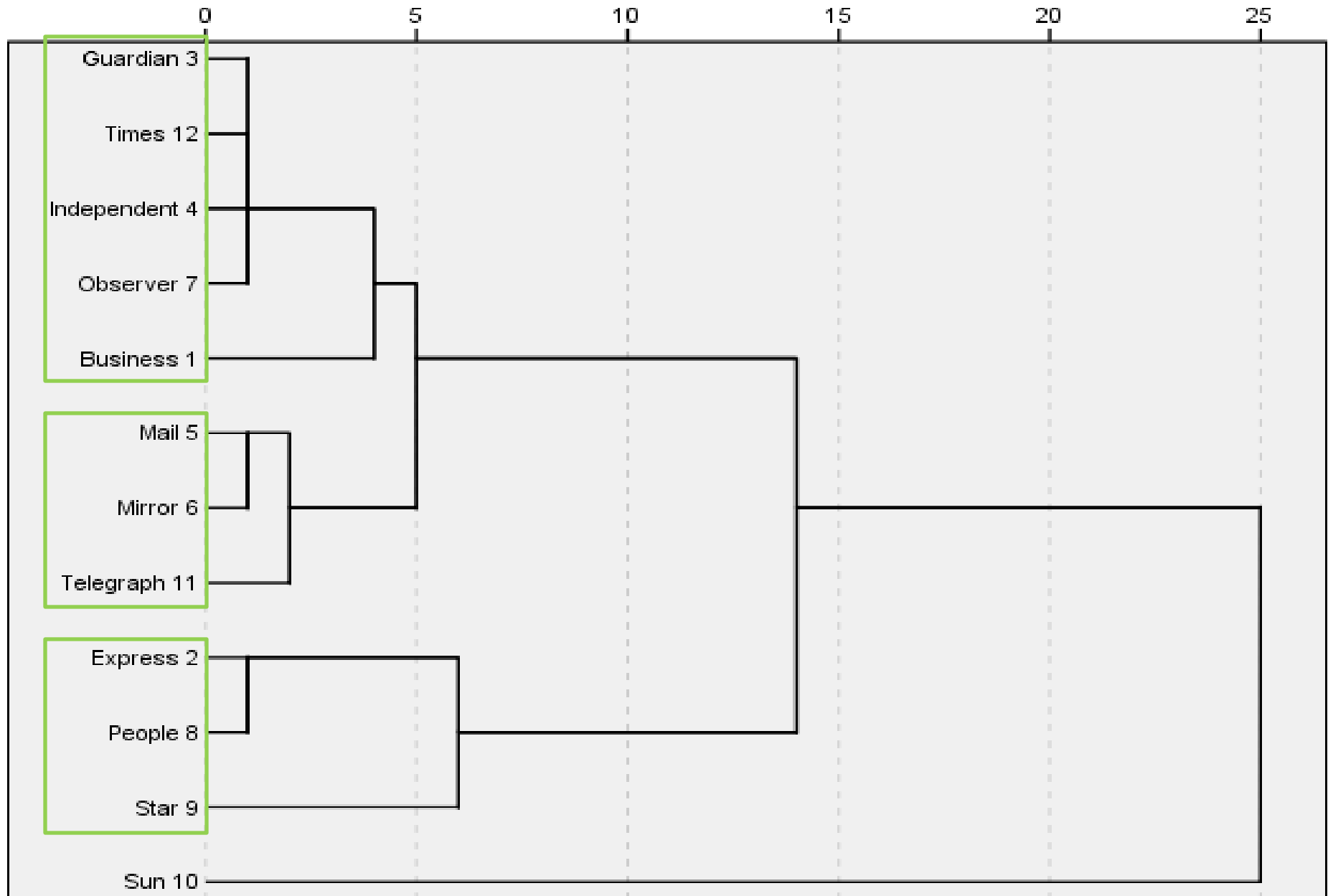
# RASIM: Frequency of nonsensical terms



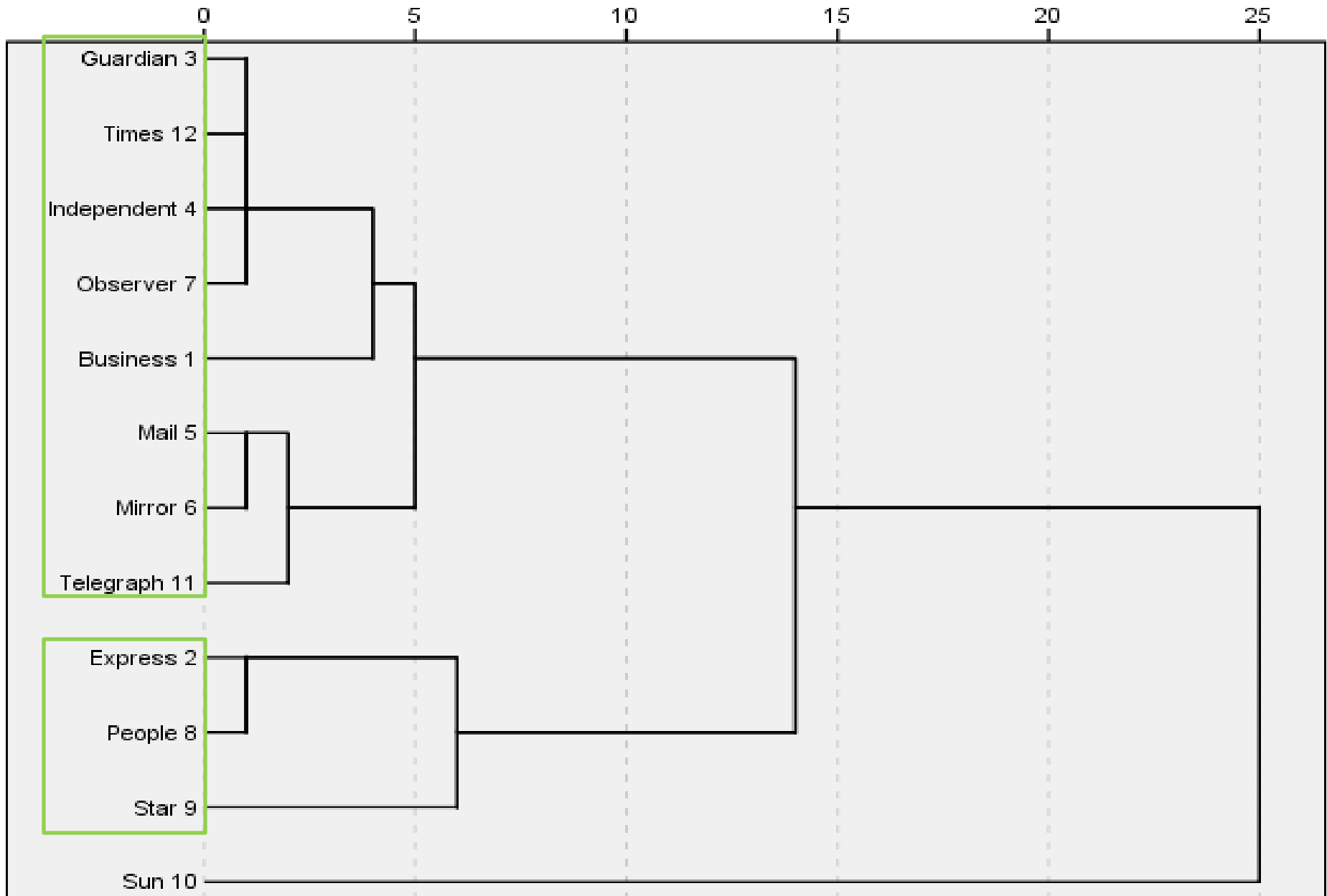
# Nonsensical terms: clusters



# Nonsensical terms: clusters



# Nonsensical terms: clusters





**CL does not examine sufficient amount of (co-)text  
lists of words (frequency, keywords, collocates)  
short concordance lines**

**Corpus research can involve the close analysis of longer  
stretches of text, up to whole texts ...**

**... while explicit annotation of features (e.g. stance)  
enables the quantification of emerging patterns ...  
... and replication of the analysis.**

# CL taking a closer look

Examination of uses of *suffocat*\* and *drown*\* in relation to RASIM.

## Why?

- Some forms corpus-wide collocates of RASIM
  - Not key in broadsheet-tabloid comparison
  - Investigation of (what was expected to be) sympathetic reporting on RASIM
- *illegal* shared collocate of many forms of *suffocat*\* and *drown*\*

Presentation was negative in almost 50% of instances

Negative presentation was **direct** or **indirect**.

# Direct

Through **attribution** by

- modification of victims with adjectives such as *illegal*, *clandestine*, etc.
- co-reference to the victims using *illegal*, etc.
- reference to their attempts to enter as *illegal*, etc.

In June, 58 **illegal immigrants** from China **suffocated** in the back of a lorry in Dover, after a journey across Europe.

[*The Express*, Nov. 2000]

A Dutch lorry driver was jailed for 14 years for killing 58 Chinese **immigrants** who **suffocated** in his trailer as he tried to **smuggle** them into Britain. Perry Wacker, 33, closed an air vent during the Channel crossing so that the ferry crew could not hear his **illegal cargo**

[*The Daily Mail*, June 2001]

# Indirect

Through **framing** the report within ...

- general references to illegal immigration
- indirect references to the ‘illegality’ of RASIM  
(e.g. *suspected asylum seeker, sneak across the perilous straits*)
- references to *smuggling, trafficking, illegal entry/transport* etc.
- references to problems with immigration/asylum or RASIM
- references to problems with, or laxity of, the existing immigration / asylum system

A SUSPECTED asylum seeker drowned and another was seriously ill with hypothermia last night after they tried to cross the Channel in a 12ft kayak.  
[*The Daily Mail*, June 2002]

China is among the top four countries whose citizens are sneaking in. It does not want a repeat of such tragedies as the drowning last year in Morecambe Bay or in 2000 when 58 Chinese suffocated in the back of a lorry heading for Dover.  
[*The Times*, Sept. 2005]

The risks of trafficking were highlighted last summer by the deaths of 58 Chinese immigrants found suffocated in a Dutch-owned truck which arrived in Dover from Belgium. Illegal immigration is also expected to be high on the agenda of an Anglo-French summit in Cahors, southern France ...  
[*The Guardian*, Feb. 2001]

# Comparison of tabloids and broadsheets

<i>suffocat*</i> + <i>drown*</i>	<b>T</b> ( <i>n</i> =250)	<b>B</b> ( <i>n</i> =409)	<b>T</b> %	<b>B</b> %	<b>LL</b>
Negative attribution [Direct]	75	93	30.0	22.7	3.15
Negative framing [Indirect]	46	80	18.4	19.6	0.11
Negative presentation [TOTAL]	121	173	48.4	42.3	1.28

LL = 6.63 → probability of results being due to chance 1%

LL = 3.84 → probability of results being due to chance 5%

## *suffocat\* / drown\**

- ❑ Negative presentation is unexpectedly high in both groups.
- ❑ There is no statistically significant difference between B and T in the proportion of negative presentation.
- ❑ Indirect negative presentation is equally favoured.
- ❑ Tabloids seem to prefer direct negative presentation ...
- ❑ ... but difference is not statistically significant.
- ❑ Both B and T make sure to project sympathy when reporting tragedies involving RASIM during their journey to the destination country ...
- ❑ ... but they very frequently communicate the notion that the victims were party to an illegal act, and, consequently, were somehow responsible for their fate.



*suffocat\* / drown\**

Following the CL leads and digging deeper

The coherence of racism

or

Racism sandwich

['Interview' with 'the man in the street', *The Sun*, June 2000]

Fifty-eight Chinese immigrants suffocating to death in a lorry after paying £18,000 each to be smuggled into Dover?

WVM: Absolutely tragic and I hope in some sort of way it's a lesson to the rest of them not to take such huge risks with their lives just to get into Britain. But it's the ones behind the smuggling that should be made to pay the price not the poor souls desperate to flee.

[Letter, *The Sun*, June 2001]

WHAT a horrific, callous man Perry Wacker is to let those 58 Chinese migrants suffocate in the rear of his truck. If our Government had stood firm and made it difficult to enter Britain - turning migrants back instead of looking after them - they would not try to smuggle themselves here. Then this tragic waste of life and the anguish of the people who found them might not have happened. The manslaughter charge should have been shared by the Government for not sorting out the problem.

# Sting in the tail

(see also Morley, 2004)

HEADLINE: Sun, sea sand and corpses: It's an idyllic scene: a young couple enjoy a picnic on the beach. But a few yards along the sand lies a body - one of many lives lost in the struggle to reach Europe. In our second **special report on immigration**, Peter Lennon visits Zahara de los Atunes in southern Spain: Real Lives

Zahara de los Atunes, situated on part of the Spanish coastline that has become one of the most popular windsurfing areas in Europe, recently achieved unwelcome notoriety when a photographer took a picture of a young couple sunbathing on the beach within yards of the body of a [drowned immigrant](#).

It was a little unlucky for Zahara which, at 41km west of Tarifa, is not the preferred destination of the immigrants. They make for the closer beaches of Punta Paloma and Bolonia, seven and 15km out from the town. But crossing in increasing numbers in light craft, at the mercy of winds and tides, Zahara is now getting its share.

[*The Guardian*, 13 December 2000]

# Slow-release racism

ANYONE returning to these shores after a few years abroad will look in awe at how Ireland has changed beyond recognition.

We have on this tiny island an incredible mix of Africans, East Europeans, Scandanavians and people from the Middle-East who have chosen to make Ireland their home.

Most are here to work, to make money and raise a family.

Others are here to escape the horror of their blood-soaked native lands, where they have seen their mothers, fathers, sisters and brothers butchered before their eyes.

They came here wearing nothing but the clothes on their back, probably having handed over obscene money to ruthless people traffickers - the price of their freedom.

Dozens of them, some pregnant, never get this far, drowning on flimsy rafts on the Strait of Gibraltar, crossing from North Africa to the promised land of Europe.

Others, like the 58 Chinese immigrants found in Dover in 2001, suffocate in the back of articulated trucks.

Under European law, we owe these people shelter, food and protection once they arrive here.



But there is one word which gives politicians a huge headache - and poses the greatest challenge in how we deal with immigrants: Deportation. Two Irish cases in particular spring to mind - Nigerian student Kunle Elukanlo, the teenager who begged to be allowed to sit his Leaving Cert here - and Nimota Bamidele, who claimed she would be stoned to death if returned to the same country.

These people clearly pose no threat to our national safety. But soon we will have to deal with people who do.

And that's why Bertie Ahern, Mary Harney and Michael McDowell must look across the water at British Home Secretary Charles Clarke's rules on what should constitute deportation.

In the wake of the London bombings last month, he has promised to kick people out of Britain if they:

SUPPORT terrorists and justify or celebrate violence  
FOSTER hatred which could lead to inter-community violence, or

WRITE, produce, publish or distribute material inciting violence.

Sounds sensible, doesn't it?

And, yet, we have to listen to the warped bile of civil rights groups questioning the legality of giving the prophets of hate their marching orders.

We have to put up with militant Islamic groups crying foul - saying objections to the new rules were ignored.

Maybe they should try selling that argument to Eileen and Paddy Tallon, whose fireman son Sean died in the 9/11 attacks on New York in 2001 doing what he loved - rescuing people.

Or John Falding, who was speaking to his girlfriend Anat Rosenberg on a mobile phone when she was killed in the No30 bus bomb in London's Tavistock Square.

The Government should photocopy Clarke's **deportation** rules and rewrite it for the statute book here.

**Genuine foreign nationals** who want to make Ireland their home, who want to embrace our way of life and enrich it with theirs, have nothing to fear from such legislation.

But those who preach or drum up support for violence - round them up and kick them out. Now.

[*The Mirror*, 26 August 2005]

# Conclusions (1)

- The quality of reporting can be quantified by examining (among other aspects) ...
  - the link between the *reaction+sustain* in entity-specific articles and ‘trigger’ events.
  - the frequency of use of explicitly loaded terms in relation to entities in focus (semantic prosodies).
  - negative attitudes hidden in / sandwiched within superficially / partially ‘sympathetic’ reporting (discourse prosodies).
- Such targeted analysis ...
  - provides a means of principled and transparent text selection.
  - expands the current family of CL techniques by ‘raiding’ compatible methodologies.

# Conclusions (2)

- The broadsheet-tabloid distinction is better seen as a cline.
- Some tabloids may show some broadsheet features – and vice versa (see also Duguid, forthcoming).
- Some newspapers show B/T features more consistently (e.g. *The Business*, *The Sun*).
- The same newspaper may vary its approach to reporting - and its stance towards the topic/entities involved - according to the reported event.
- Studies of groups of newspapers may miss important individual differences (see also Marchi & Taylor, forthcoming).
- Studies of particular newspapers can safely generalise only about the particular newspaper.

# Questions

- How helpful / generalisable are conclusions drawn from the analysis of a small number of articles ...
- ... particularly when they have been selected because they are 'interesting' or 'telling'?
- Can depth of analysis compensate for lack of representativeness?

# Suggestions (1)

- Objective vs. subjective distinction not useful – misleading.
- Objective processes involve subjective decisions at some point:
  - Setting thresholds for statistical test results (why are we happy with 1% probability of chance/error- but not 1.1%?)
  - Lexis-based research is not a-theoretical (McEnery & Gabrielatos, 2006).

# Suggestions (2)

- Useful distinctions:
    - explicit annotation
    - discrete categories
    - precise counting
    - stat. testing
    - total accountability
  - replicability
  - implicit annotation
  - non-discrete categories
  - approximate/fuzzy counting
  - general impression
  - selective treatment
  - non-replicability
- Replicability applies to analytical categories, procedures and results - **not** interpretation.

What corpus linguists tend to do



What corpus linguistics can do