

# **The Linguistics of Laughter**

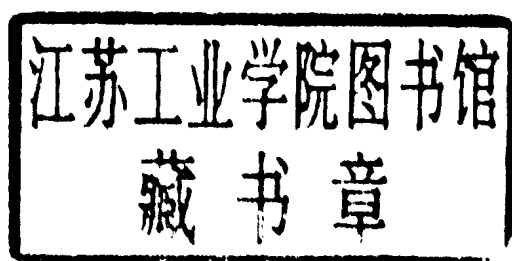
A corpus-assisted study of laughter-talk

**Alan Partington**

# **The Linguistics of Laughter**

A corpus-assisted study of  
laughter-talk

**Alan Partington**



First published 2006  
by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada  
by Routledge  
270 Madison Ave, New York, NY 10016

*Routledge is an imprint of the Taylor & Francis Group,  
an informa business*

© 2006 Alan Partington

Typeset in Times New Roman by  
Newgen Imaging Systems (P) Ltd, Chennai, India  
Printed and bound in Great Britain by  
Biddles Ltd, Kynn's Lyn

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

*British Library Cataloguing in Publication Data*

A catalogue record for this book is available from the British Library

*Library of Congress Cataloging in Publication Data*

A catalog record for this book has been requested

ISBN10: 0-415-38166-5

ISBN13: 978-0-415-38166-6

# The Linguistics of Laughter

This book examines what speakers try to achieve by producing ‘laughter-talk’ (the talk preceding and eliciting an episode of laughter) and, using abundant examples from language corpora, what hearers are signalling when they produce laughter.

In particular, the author focuses on the tactical use of laughter-talk to achieve specific rhetorical and strategic ends: for example, to construct an identity, to make an argumentative point, to threaten someone else’s face or save one’s own. Although laughter and humour are by no means always related, the book also considers the implications these corpus-based observations may have about humour theory in general.

As one of the first works to have recourse to such a sizeable databank of examples of laughter in spontaneous running talk, this impressive volume will be a point of reference and an inspiration for scholars with an interest in corpus linguistics, discourse, humour, wordplay, irony and laughter-talk as a social phenomenon.

**Alan Partington** is Associate Professor of Linguistics in the Faculty of Political Science at the University of Bologna, Italy. He is the author of *Patterns and Meanings: Using Corpora for English Language Research and Teaching* and *The Linguistics of Political Argument: The Spin-doctor and the Wolf-pack at the White House* (also published by Routledge).

## **Routledge Studies in Linguistics**

### **1. Polari – The Lost Language of Gay Men**

*Paul Baker*

### **2. The Linguistic Analysis of Jokes**

*Graeme Ritchie*

### **3. The Irish Language in Ireland**

From Góidél to globalisation

*Diarmait Mac Giolla Chríost*

### **4. Conceptualizing Metaphors**

On Charles Peirce's marginalia

*Ivan Mladenov*

### **5. The Linguistics of Laughter**

A corpus-assisted study of laughter-talk

*Alan Partington*

# Acknowledgements

First of all, I am grateful to the White House Library for making the briefings transcripts publicly available and to C-Span public service television for broadcasting the briefings on their website. I must also thank fellow members of *CorDis*, a computer-assisted research project into modern political English financed by the Italian Ministry for Universities, especially John Morley, Louann Haarman, Alison Duguid and David Brierley, an honorary member, for their invaluable observations on the manuscript. Many thanks too are due to Salvatore Attardo for his comments and encouragement at vital moments. As ever, I am immensely grateful to Peter Levy for the sureness of his linguistic touch, his lateral insights and his cheerful chiding.

# Contents

<i>Acknowledgements</i>	vi
Introduction: laughter-talk – research questions and methods	1
1 Joke humour theory and language principles	25
2 Laughter in running discourse: shifts of mode, narrative, role and register	57
3 Face-work and the in-group	82
4 Wordplay, phraseplay and relexicalization	110
5 Teasing and verbal duelling	144
6 Irony and sarcasm	182
7 General conclusions	225
<i>Appendix 1: visual puns and verbal-visual puns</i>	237
<i>Appendix 2: levels of sarcasm</i>	239
<i>Appendix 3: irony and popular historiography</i>	240
<i>Notes</i>	242
<i>Bibliography</i>	248
<i>Name index</i>	256
<i>Subject index</i>	259

# Introduction

## Laughter-talk – research questions and methods

### 1.1 Aims

This work investigates the phenomenon of ‘laughter-talk’, that is, the talk preceding and provoking, intentionally or otherwise, a bout of laughter. More specifically, with the assistance of language corpora, I examine what speakers try to achieve by engaging in laughter-talk and what both speakers and hearers may be signalling when they produce laughter. Of particular interest is the tactical use of laughter-talk to achieve specific rhetorical ends, for example, to construct an identity, to make an argumentative point, to threaten someone else’s face or boost one’s own.

Although laughter and humour are by no means coterminous, intuition, experience and past literature tell us that they are closely related and so I go on to consider the implications these corpus-based observations may have for humour theory in general.

Past research into the relationship between language and laughter has fallen into two camps. The first type has been, roughly speaking, cognitive-psychological and has concentrated on laughter as a signal of humour. It has tried to answer a question of the type ‘what does the human mind, or psyche, find funny, and why?’ The data discussed was usually deliberately ‘authored’ material, either comic literature or what are known in layman’s terms as ‘canned jokes’. This was the case for both technical and philosophical reasons. Until comparatively recently the means of recording spontaneous episodes of humour were not available, whilst a feeling dominated the field that authored material was a more proper object of study, more interesting and closer to art and literature. Of late, however, some analysts have felt that such entirely decontextualized studies of humour, devoid of due regard to the producer and receiver of the laughter-talk, may leave important questions unexamined.

The second vein of research is socio-anthropological and looks at laughter in its social contexts. Rather than treating laughter as natural, instinctual and beyond our conscious control, in short, as a response to some stimulus entering the mind, attention is paid to issues such as when, where and in what ways people organize, produce, respond to and interpret laughter as part of the ongoing stream of social interaction. Research of this second variety was made possible by the development



## 2 *Laughter-talk*

of tape recording. One major criticism that has been levelled at such studies, however, is the high degree to which they are subject to the Observer's Paradox, how both the observer and the process of observation interfere considerably with what is observed. Much of the data is collected relatively informally, by recording friends or colleagues who are often both aware of the presence of the tape recorder and of the object of the research. This inevitably begs questions about the spontaneity and authenticity of such discourse. The amount of data collected in such studies is often relatively small.

In more recent times, the advent of corpus technology, allied to the increasing availability of material in electronic form on the Web, has made it possible to compile large corpora of authentic discourse occurring while subjects are going about their everyday business. This enables the analyst to collect and study considerable numbers of episodes of spontaneous laughter in circumstances where participants interact naturally and are unaware that their linguistic or laughter behaviour is an object of study. The main data for the current work consists of around 180 transcriptions containing circa 1,000,000 words of press conferences (or 'briefings') held at the White House over the past six years. The transcriptions contain indications of where laughter occurs – the word 'laughter' in round or square brackets – and it is possible to recover a degree of audio-visual information regarding the contexts in which it occurred, since the briefings are broadcast over the Web by C-Span television.<sup>1</sup> Thus, one of the innovations of the present work is the use of methodologies and software deriving from Corpus Linguistics (especially concordancing) and the integration of quantitative and qualitative approaches in the study of laughter phenomena.

The main concept driving the current work, then, is the need to compare and contrast current developments in both the cognitive and the discourse/conversational fields of linguistics. Indeed, the contention is that these approaches are compatible and that attention to both is necessary if we wish to build up a picture of why human beings use laughter-talk and how they do so to attain strategic goals in everyday interaction. In brief, in order to understand laughter-talk, we need to develop a model which can encompass and render compatible three elements: a theory of language production and reception, a theory of cognition as it relates to humour and a theory of human social interaction. The availability of large quantities of suitable data along with the technological and methodological capacity to analyse it may shed new light on this most ancient of objects of study, which thinkers from Aristotle, through Hobbes, Freud, Bergson to Woody Allen have found endlessly fascinating and utterly vexing in equal measure.

### 1.2 **Press briefings**

The current research into laughter utilizes data from three corpora of White House press briefings. Briefings are press conferences held on a regular basis – in the case of the White House, almost daily. They are a particular type of *institutional talk* (Drew and Heritage eds 1992), which is basically defined as talk between professionals and lay people, but the definition can be stretched, as here,

to include talk between two groups of professionals with an audience of lay persons (the TV and Internet audience). In fact, briefings are a particularly fascinating genre of institutional talk in that they combine features of informal talk, given that the participants meet so often and know each other well, and confrontational or 'strategic' talk. The two parties involved – the spokesperson or podium (officially known as the 'White House Press Secretary') and the press – have very different interests and aims in life, which are in conflict on several levels. The podium wishes to project his political ideas and particular view of the world, the press to test that view, often suggesting more critical alternatives. The press hopes to uncover ever more information, including any evidence of weakness, malpractice, internal dissension and so on, the podium ideally wants to give as little away as possible outside the official line. They adopt and exploit different participant roles or *footings* (Goffman 1981; Levinson 1988), command non-symmetrical sets of discursive resources and employ different discourse strategies; they use different metaphors to describe the world and probably even see the whole nature of the business being conducted in different ways (Partington 2003).

What transpires in these briefings can also be extremely important and highly delicate from a political perspective:

Anything McCurry [press secretary during the Clinton administration] uttered from the podium magically attained the status of official White House policy, and if he deviated later on the administration would be accused of the dreaded sin of flip-flopping.

(Reaves White)<sup>2</sup>

Not only are the podium's words often treated by the press as White House policy, but they risk interpretation by non-American bodies as official US policy. Since they are broadcast both on television and on the Internet, 'any misstep can be beamed instantaneously around the world' (*CNN-allpolitics*). All this exposure, of course, means fame: 'the chief White House spokesman's face is probably as well known as any cabinet member' (*CNN-allpolitics*). In Galtung and Ruge's (1981) terms he is 'newsworthy', has become an 'elite person' in his own right. Many of the journalists, too, are well-known television faces or newspaper by-lines. Clayman and Heritage (2002) suggest that press conferences, including briefings, occupy the same vital space in the US political-media arena as the news interview in the British. An outline of the typical structure of White House press briefings is given in section 2.2.1.

### **I.3 Corpus-Assisted Discourse Studies**

#### ***I.3.1 Quantitative and qualitative approaches combined***

This research into laughter-talk is an instance of a project in the nascent interdisciplinary field of Corpus-Assisted Discourse Studies (CADS). This arose from the realisation that some of the methodology and instruments commonly used in

Corpus Linguistics might be adapted for the study of features of discourse (see especially Louw 1993; Stubbs 1996, 2001; Partington *et al.* 2004). In other words, that it was possible to combine the *quantitative* types of analysis used in Corpus Linguistics, which generally take into consideration large quantities of texts and subject them to statistical analysis, with the *qualitative* methods more typical of discourse studies which examine in detail much smaller amounts of discourse, frequently single texts. In its purest form:

the quantitative paradigm hinges on a hypothetical-deductive mode of inquiry and a fairly rigid sequence of interventions which foresee the performance of experiments in controlled situations and the statistical measurement of data in order to reach reliable and replicable results which allow for generalisations and the prediction of a cause and effect relationship.

(Haarman *et al.* 2002: 56–57)

whereas:

qualitative methodology instead proceeds in a non-experimental or exploratory fashion, draws considerably on insight and intuition and derives results from the systematic observation of phenomena in such a way that theories or hypotheses emerge inductively and are said to be ‘grounded’ in data.

(Haarman *et al.* 2002: 57)

Many experimenters, especially in the social sciences, have questioned this rigid dichotomy, maintaining that elements of both paradigms can usefully be employed in the research process. Just as the experimental researcher subjectively intervenes in the research design when formulating hypotheses and in deciding cut-off points for statistical analysis, so the qualitative researcher cannot be insensitive to quantity in the interpretation of data.<sup>3</sup> In this school of thought, research is ‘a dynamic process which links together problems, theories and methods’ (Bryman and Burgess 1994: 4) and the researcher is free to shunt back and forth among hypotheses, data-collection, analysis, evaluation and even speculation, as long as these phases are kept separate and the movements among them are closely chartered. Data *creation* (as in, say, arranging circumstances for laughter to occur and then recording it for analysis) is another matter and should only be employed if there is no alternative. In terms of a debate which has recently opened up in the field (Provine 2000; Attardo 2003), a CADS approach to laughter studies is neither entirely *performance* based nor wholly *competence* based, but combines features of both.

What follows, then, is an outline of the CADS methodology employed in the course of the present research. This description is included in the Introduction since the intricacies of the methodology will not always be explained in detail in the following chapters, first, because many readers will be more interested in discourse and humour studies than in Corpus Linguistics and, second, because many of the techniques are fairly repetitive.

### 1.3.2 *Corpus-Assisted Discourse Studies methodology*

The initial phases of CADS methodology are as follows:

- 1 data collection and corpus compilation;
- 2 data / corpus organization;
- 3 corpus interrogation.

#### (1) *Data collection and corpus compilation*

##### THE MAIN CORPORA

A number of the corpora used in the current research were compiled for Partington (2003) where briefings were analysed from a number of political and discourse angles including stance and footing, journalistic attribution, politeness phenomena and metaphor.

During this research it became clear that the briefings corpus also constituted an ample database of laughter occurring in semi-spontaneous speech. These bouts of laughter were in transcribed texts but many briefings were also available to audio-visual scrutiny because they are frequently screened as webcasts. This presented the opportunity to examine what makes people laugh in real-life spontaneous speech in a working environment. Moreover it was possible to do so with a degree of 'blindness' missing from some research into authentic discourse, in that the subjects are unaware of the objectives of the research and there is no danger of the subjects being influenced by interaction with the researcher. In the present case the raw data is produced by participants who most probably never in their wildest dreams imagined that their interaction might be studied from the point of view of laughter-talk. Moreover, the data was transcribed by parties (professional transcribers at the White House library) different from the final analyst (myself). The separation of data formulation and data analysis is, of course, fundamental in the elimination of contamination between researcher and subject.

The technology of data collection developed rapidly as this research evolved, that is, from 1998 to the present. The briefings transcripts for the earliest versions of the corpora were collected 'manually' by downloading them one by one from the White House Library website. The very first corpus (*Dems*) was of briefings held during the last years of the Clinton administration, 48 in all, composing a total of 250,000 words of spoken discourse. By watching the webcasts I was able to make notes on interesting features of paralanguage. At this early stage I was not exclusively interested in laughter phenomena but in all aspects of this discourse type.

The second corpus (*Reps*) was compiled in a different fashion. I began to collect a batch of briefings, still downloading one by one, in September–October, 2001, that is, during and immediately after the September 11th attack, in order to study the podium's and the press's reaction to such dramatic events. I subsequently decided to collect batches of briefings at six monthly intervals in order to

construct what is known as a 'monitor' corpus (Sinclair 1982). Since the transcriptions remain on the White House Library website until the end of the administration (they disappear when the administration changes – a 'new broom' policy), I also collected a batch from six months *before* the attack, for purposes of comparison. Each batch contains circa 125,000 words of running discourse. The *Reps* corpus, then, currently consists of a series of subcorpora:

- *Reps0* (Mar/Apr 2001)
- *Reps1* (Sep/Oct 2001)
- *Reps2* (Mar/Apr 2002)
- *Reps3* (Sep/Oct 2002)
- *Reps4* (Mar/Apr 2003)
- *Reps5* (Sep/Apr 2003)

for a total of 750,000 words which, when combined with *Dems*, makes a grand total of 1 million words of briefings. It has subsequently become possible, thanks to expanding software capability and disk storage, to download files from the Web in automatic fashion, using programmes such as *Nettransport* or *Winhttrack*. These programmes are capable of downloading onto a hard-disk or other memory support all the files in a given website. By specifying the briefings section of the White House Library site, *WHBig* was compiled, containing all press secretary briefings held from the beginning of the Republican reign until 17th June 2004 (the last available transcript at the time of compilation), a total of approximately 6 million words.

#### OTHER CORPORA

One of the axioms of CADS is that discourse study is necessarily comparative in two separate but related ways. First, within an individual discourse type, only by comparing the choices being made by speakers or writers at any point in a discourse with those which are normal, that is, usual within the genre, can we discover how *meaningful* those choices are. Observations from a single source (even an authentic text) are of limited value and are essentially anecdotal: 'by and large, we are not methodologically justified in interpreting the significance of a particular linguistic event unless we can compare it with other similar events' (Partington 1998: 146). Testing observations and findings against corpus data can provide 'background information' against which particular events can be judged.

Second, if we are also interested in the characteristics of the discourse type itself, it is vital to be able to compare its particular features and patterns with those of other discourse types. In this way we discover *how* it is special, and can go on to consider *why*. All genre/register/discourse type analysis is thus properly comparative. In the wider field of discourse studies, this requirement has unfortunately not always been observed in practice.

In sum, CADS analysts hold that, if texts are not compared to other bodies or corpora of texts it is not possible to know or to prove what is normal and only

against a known background of what is normal and expected can we detect the unusual and meaningful.

A number of other corpora were utilized in the course of this research as a basis for comparison with briefings discourse. These include several corpora of journalistic texts: a collection of British news interviews (*INTS*) of similar size to the first briefings corpus (250,000 words), a 100-million-word corpus of written British broadsheet newspaper texts (*Papers*) and a circa one-million-word corpus of editorials and reports from British broadsheets and tabloids (*EdsReps*). The *Frown* (one million words of general US English) and its sister *Flob* (one million words of UK English) corpora, the *Colt* corpus of teenager talk and the *Wellington Spoken Corpus* (*WSC*) of general conversation were all used when appropriate. The *British National Corpus* (*BNC*) on the Web was also occasionally consulted (<http://www.natcorp.ox.ac.uk/>).

## (2) Data/corpus organization

### EDITING

Corpora can be edited in various ways, a process normally referred to as *mark-up*. Most of the original corpora, *Dems* and *Reps*, have been edited so that, when required, it is possible to treat the journalist's contributions and the podium's contributions as separate subcorpora. Thus, for instance, we can contrast the way journalists typically use a particular expression with the way the podium employs it.

With these corpora subdivisions and this mark-up, it became possible to compare (i) the podium's speech with that of the journalists, (ii) the speech of different podiums and (iii) briefings from different periods (including Democrat and the Republican periods of office).

### ISOLATING LAUGHTER EPISODES

This can be carried out in two ways. First, one can go through each file and use the search and cut and paste facilities offered by a Word Processor to extract all episodes where the item *laughter* appears in the transcript. Alternatively, one can run the concordancer (see section 1.3.4) to collect all instances of the use of the word *laughter* and then save the resulting concordance list in a separate file. It is possible to ask the program to include considerable co-text around the word or phrase to be sought (the *searchword*), which generally allows the analyst to study the episode in context. The second method clearly saves a great deal of time, but subsequent editing of the file may be needed to put the episodes in the correct chronological order or to remove episodes which appear more than once because one instance of *laughter* has occurred in close proximity to another – a very frequent phenomenon. This second method was the one I used and separate concordances of *laughter* were made for *Dems* and for each of the *Reps* collections. These will be referred to collectively as the 'laughter files'.

Subsequent analysis, classification and evaluation of the laughter bouts is, of course, largely a 'manual', that is, a human task. However, if the concordances of *laughter* are transferred into a word processor file, the analyst can add notes to the raw data. This then allows him or her, at a later date, to concordance the annotated version and call up, for instance, episodes where *wordplay* has occurred, or where *sarcasm* or *facework* are in play (if, of course, these were categories employed). By using the concordancer's *Context* facility, it becomes possible to perform cross-referenced enquiries: for example, to call up all episodes where *wordplay* and *facework* were deemed to be involved in the laughter-talk.

### (3) *Corpus interrogation and data analysis*

#### SOFTWARE

A corpus by itself is simply an inert archive. However, it can be 'interrogated' using dedicated software. The packages used to interrogate the various corpora were *MicroConcord* and *WordSmith Tools*, both very widely available.<sup>4</sup> Such software can help supply us with various kinds of information on the frequency of occurrence of lexis, whilst the queen of corpus tools, the *concordancer*, is essentially a collector and collator of examples. Though in themselves quite simple (and easy to use), these tools can provide a great deal of information about the texts contained in a corpus not always easily available to the naked eye.

#### 1.3.3 *Some frequency data*

Figure I.1 shows the incidence of laughter episodes in the various subcorpora. In *Dems* there are 220 occurrences of [laughter], one every 1,171 words. In *Reps* overall there are 323, an incidence of one laughter episode every 2,340 words. However, in *Reps1*, the briefings which take place during and immediately after 9/11, the incidence drops to one every 3,474 words; understandably it was a

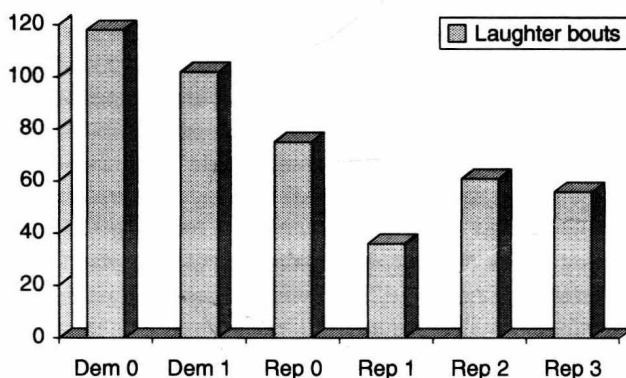


Figure I.1 The relative number of laughter bouts in each of the subcorpora of briefings.

sombre moment. There were other clear indicators too of a particular tension between the press and the podium. There is still more to the question, however. In *Reps0*, the briefings held before the attack, the laughter incidence was analogous to that of *Dems*: one episode every 1,664 words, whilst in *Reps5* the incidence was one every 2,508. The briefings had not recovered their gelastic quality thanks, most probably, to the further serious political and military developments.

*WordSmith* provides a distribution *Plot* tool which displays in a visual form where any particular word or phrase appears in a file. The distribution of *laughter* in first half of the *Dems* files is shown in Figure I.2.

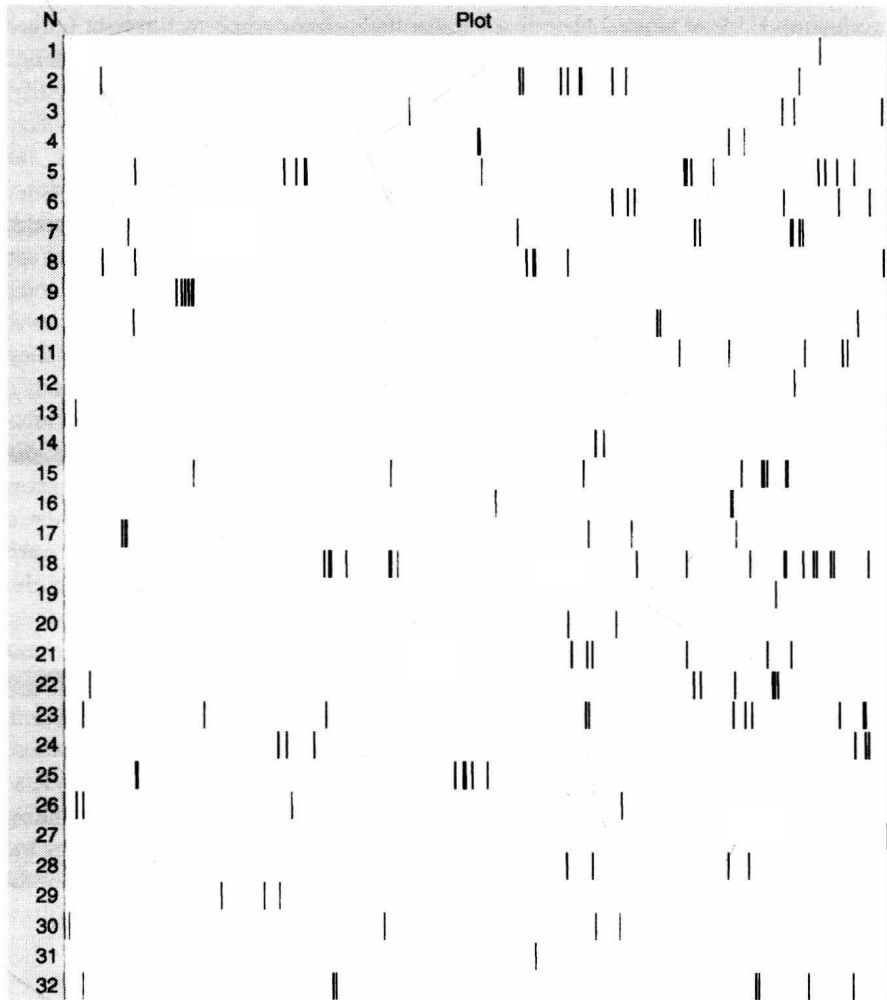


Figure I.2 A plot over time of the incidence of bouts of laughter in 32 briefings in the *Dems* corpus (*WordSmith Tools*, version 4.0).



This is a graphic demonstration of how bouts of laughter tend to cluster together, in other words, that laughter seems to spawn more laughter. As we shall see in Chapter 2, these constellations occur within phases of *interaction*, whilst the long stretches where laughter is absent tend to indicate phases of *transaction*, when, as it were, the 'real' business of briefings is under way. We can also see that there is also a tendency for laughter – and therefore interaction – to occur towards the end of a briefing.

The *Collocation* facility provided by *WordSmith* gives a list of the most frequent words which appear in the immediate co-text of the searchword. Collocation is one of the most important technical concepts in Corpus Linguistics (Stubbs 2001: 29–30; Hoey 2005: 2–15), but it is defined in slightly different ways (Sinclair 1991: 170; Partington 1998: 15–17) and the terminology can be confusing. Here, it is used to indicate the actual, observed co-occurrence (either noted by a human analyst or picked out of the ocean of a corpus by software) of one lexical item with others within a short span of text. When using the facility, the researcher can decide the span but it is usual to set it at between three to five words to the left and right of the searchword. Those lexical items which are observed to occur within this span are called the *collocates* of the searchword. Thus, if we were to take the word *corpus* in the third sentence in this paragraph as our searchword and set the span at three items either side, the collocates are *ocean*, *of* (twice), *a*, *by*, *software*, *like*. If, using the *WordSmith Collocation* tool, we ask for all the collocates of the word *corpus* every time it appears in this book and ask for them to be ordered by frequency, we find, not surprisingly, that they include *linguistic* (co-occurring with *corpus* 13 times), *data* (10), *evidence* (7), *discourse* (6) and *briefings* (4). However, and just as predictably, if we spend a moment to reflect, the most common collocates are the grammar words *the* (41) and *in* (14). When issues of frequency are being discussed, the term *collocation(s)* is often employed, but in two ways. In the first, *x* is said to be a frequent or common collocation of *y*, that is, the items *linguistic* and *evidence* are said to be a frequent or common collocations of *corpus*. In the second, on the other hand, it is the *combination* of items *xy*, for example, *linguistic corpus* or *corpus evidence*, which is called a common collocation. By the same token, rare or nonce combinations such as *stone blind* (Edward Thomas) or *a grief ago* (Dylan Thomas) are often referred to as unusual collocations.

*WordSmith's* so-called *Clusters* facility, meanwhile, gives a list of the most frequent word-strings the searchword appears in. Neither of these tools at first blush furnishes any remarkable information at all on the co-text surrounding *laughter*, which is evidence that there is no such thing as a special vocabulary for conversational laughter. Examples of the use of these two tools are discussed in section I.7.

### I.3.4 *Concordancing*

The concordancer is a collector and collator of examples. It extracts as many instances as the analyst wishes of the searchword or expression under analysis