

**Amanda Rose Lanham Potts**

*Bachelor of Arts, Master of Applied Linguistics*

***At arm's length:***

**Methods of investigating  
constructions of the 'other' in  
American disaster and disease  
reporting**

Lancaster University, Department of Linguistics and English Language

*Submitted for PhD by Research, December 2013*

ProQuest Number: 11003805

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 11003805

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

**This thesis was submitted in fulfilment of the requirement for the degree of Doctor of Philosophy (Linguistics).**

**This work has not been submitted for any other degree or award at this or any other university or educational establishment.**

**This thesis is the result of my own independent work and investigation. Other sources are acknowledged by explicit references, and a list of these is appended. The views expressed are my own.**

**Amanda Potts**

# **At arm's length: Methods of investigating constructions of the 'other' in American disaster and disease reporting**

*Amanda Rose Lanham Potts, Bachelor of Arts, Master of Applied Linguistics*

Thesis submitted for PhD by Research, December 2013  
Department of Linguistics and English Language, Lancaster University

## ***Abstract***

This thesis is a corpus-based critical discourse analysis of social actor construal in media discourse of major American press publications, 1981-2009. Analysis is based upon two custom-collected corpora: a 36,736,679-word corpus of reporting on Hurricane Katrina, spanning approximately one year of coverage; and a 161,144,924-word corpus of AIDS/HIV reporting, published over the course of nearly three decades. I detail common attribution, argumentation, and predication strategies associated with the most frequent nomination strategies in both corpora, as well as investigating construal via *topoi* and metaphorical representation of actors and actions.

Matched analyses are performed over the course of the thesis, enabling me to: a) uncover common characteristics of discourses of moral panic and risk society in both Katrina and AIDS/HIV reporting, and b) refine a generalizable method of analysing high-frequency items in large corpora varying in word count, diachrony, and topicality. To this end, I propose a reproducible method of downsampling results called proportional semantic collocation, by which a researcher might quantitatively determine salient categories of semantic preference for a given search term, and use these indicators for close, qualitative analysis.

In close analyses of my corpora, shared discursive strategies and representation patterns characterised both people affected by Katrina and people with AIDS. These included: nearly total lack of agency; the construal of threat to the in-group, e.g. through the *topos* of numbers; and highly frequent association with additional othered groups (or 'deviancy doubling'). An additional pattern noted throughout the thesis has been identified as 'social sequestering', or a segregation of certain othered groups into risk or moral panic categories, typified by discourses creating distance or dissuading reader identification with these groups. In this way, major American newspaper publications are found to perpetuate an us/them dichotomy, encouraging social distance to protect the majority group from social panic.

## ***Acknowledgements***

First and foremost, I would like to express my deepest gratitude to my supervisor, Paul Baker, who is my steadiest supporter and greatest role model. He is a wonderful mentor who has showed me how to express myself through research, demonstrating that it is possible to be both passionate and impartial. Many of the opportunities that I have been given are owed directly to Paul's recommendation and reputation, and I feel certain that I have thrived under his supervision in a way that I could not have done working with anyone else.

I also owe a debt of gratitude to two people who made the methods used here possible. Without the excellent corpus analysis tools and technical support provided by Andrew Hardie and Paul Rayson, I might still be wrangling my corpora these three years on. I appreciate their talent, their dedication to their craft, and their patience in lending so much assistance to so many burgeoning corpus linguists, myself included.

The Department of Linguistics and English Language at Lancaster University has been an ideal environment in which to grow as a researcher, and I was fortunate to be surrounded by so many imminent and gracious academics during my degree. I would particularly like to express my appreciation to Jonathan Culpeper, Mick Short, Greg Myers, and Veronika Koller, who have all kindly dedicated time to help me to become a better teacher and, in turn, made me a better learner.

I would also like to thank my examiners, Gerlinde Mautner, Elena Semino, and Rosamund Moon, who inspired me throughout the process of writing this thesis, and whose comments have greatly refined the final product. Receiving feedback and encouragement from these brilliant women is a memory that I will always treasure.

Finally, I express my appreciation to the Faculty of Arts and Social Sciences at Lancaster University for the award of a research scholarship which enabled me to undertake this degree, and for their travel fund, which helped me to present and receive feedback on earlier iterations of this work.

I have never felt so fully supported; this work is the product of an idyllic university, a generous faculty, a vibrant department, an inspiring staff, and an understanding partner, family, and circle of friends. I am so lucky.

# Contents

<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.1 Orientation.....	1
1.2 Research Questions .....	2
1.3 Key terms.....	3
1.3.1 Discourse.....	3
1.3.2 Ideology .....	4
1.3.3 Media/Press.....	4
1.3.4 Identity.....	5
1.3.5 Race .....	6
1.3.6 Social class.....	6
1.3.7 Sexuality .....	7
1.4 Hurricane Katrina .....	7
1.4.1 The 'catastrophic' event .....	7
1.4.2 Social Profile: New Orleans before and after Katrina .....	8
1.4.3 The role of the media.....	10
1.4.4 Towards a critical analysis of Hurricane Katrina.....	11
1.5 HIV/AIDS .....	12
1.5.1 A brief overview of HIV.....	12
1.5.2 Social impact.....	14
1.5.3 Social stigma.....	15
1.6 Rationale for study.....	17
1.7 Structure of thesis.....	19
1.8 Summary.....	19
<b>Chapter 2: Overview of theory and literature .....</b>	<b>20</b>
2.1 Media Discourse .....	20
2.1.1 The Socialized Selection of News .....	21
2.1.2 Writers, Readers, and Cultural Consensus.....	24
2.1.3 Moral Panic and the Risk Society .....	25
2.1.4 The Structuring of News .....	27
2.2 Critical approaches to discourse analysis.....	28
2.2.1 The Discourse-Historical Approach .....	29
2.2.2 The Socio-Cognitive Approach: Critical Discourse Studies.....	32
2.2.3 Criticisms leading to the incorporation of a corpus-based method .....	34
2.3 Corpus Assisted Discourse Studies and corpus-based critical discourse analysis.....	36
2.3.1 A brief overview of Corpus Linguistics.....	36
2.3.2 Schools of Corpus Assisted Discourse Studies and corpus-based critical discourse analysis .....	37

2.3.3	Strengths of the mixed method .....	38
2.3.4	Key corpus-based discourse analyses of identity .....	38
2.3.5	Criticisms and responses .....	46
2.4	<i>Summary</i> .....	48
<b>Chapter 3:</b>	<b>Data and methods</b> .....	<b>50</b>
3.1	<i>Data collection and cleansing</i> .....	50
3.1.1	Factiva search builder .....	50
3.1.2	Sampling frame .....	51
3.1.3	Corpus annotation and markup.....	52
3.2	<i>Description of the corpora</i> .....	53
3.2.1	Hurricane Katrina .....	53
3.2.2	AIDS/HIV.....	56
3.2.3	Corpus of Contemporary American English.....	59
3.3	<i>Methods</i> .....	60
3.3.1	Corpus Linguistics.....	60
3.3.2	Critical Discourse Analysis.....	65
3.3.3	Tools.....	71
3.4	<i>Stages of analysis and study matching</i> .....	73
<b>Chapter 4:</b>	<b>Attributes of ‘people’ affected by Katrina</b> .....	<b>75</b>
4.1	<i>Introduction</i> .....	75
4.2	<i>Methods</i> .....	76
4.2.1	Identification of the naming strategy .....	76
4.2.2	Calculation of collocates.....	76
4.3	<i>Analysis</i> .....	77
4.3.1	USAS category A: general and abstract terms .....	79
4.3.2	USAS category X: psychological actions, states and processes .....	83
4.3.3	USAS B and L categories: disabled, sick, ill and dead .....	84
4.3.4	Collocates from USAS F and H categories: hungry and homeless.....	88
4.3.5	Collocates from USAS I category: ‘low-income’ and ‘moderate/high-income’ sets.....	92
4.3.6	Collocate from the USAS M category: displaced .....	97
4.3.7	Collocates from the USAS N category: many, most, fewer, and assorted figures....	98
4.3.8	Collocates from USAS S category: social actions, states and processes .....	100
4.3.9	Collocates from USAS T category: young and elderly .....	111
4.4	<i>Summary</i> .....	113
4.4.1	Of Findings .....	113
4.4.2	Of methods .....	114
<b>Chapter 5:</b>	<b>Semantic preference of frequent Hurricane Katrina naming strategies</b> .....	<b>115</b>
5.1	<i>Introduction</i> .....	115
5.2	<i>Method: Targeting search terms</i> .....	115
5.2.1	Nominal frequency lists.....	116

5.2.2	Possessive collocations.....	117
5.3	<i>Exploring naming strategies</i> .....	118
5.3.1	Dictionary definitions.....	119
5.3.2	Incidence of naming strategies.....	120
5.4	<i>Method: Downsampling collocates</i> .....	121
5.5	<i>Analysis</i> .....	124
5.5.1	Adjectival collocations .....	124
5.5.2	Nominal collocations .....	130
5.5.3	Verbal collocates.....	134
5.5.4	Water and volume metaphors.....	146
5.6	<i>Summary</i> .....	151
5.6.1	Of findings.....	151
5.6.2	Of method .....	151
<b>Chapter 6:</b>	<b>Katrina's 'refugees'</b> .....	<b>153</b>
6.1	<i>Introduction</i> .....	153
6.1.1	Dictionary definitions of term .....	154
6.2	<i>Method</i> .....	155
6.2.1	Broad categorization of concordance lines.....	156
6.2.2	Description of categories .....	157
6.2.3	Frequency of categories .....	160
6.3	<i>Analysis of categories</i> .....	162
6.3.1	Category: KATRINA, Subcategory: UNCONTESTED.....	162
6.3.2	Category: KATRINA, Subcategory: CONTESTED .....	167
6.3.3	Category: KATRINA, Subcategory: DEFINITIONS.....	172
6.3.4	Subcategory: PARALLELISMS.....	178
6.3.5	Subcategory: CAMPS.....	180
6.4	<i>Summary</i> .....	182
6.4.1	Of findings.....	182
6.4.2	Of method .....	184
<b>Chapter 7:</b>	<b>The attributes and actions of people with AIDS/HIV</b> .....	<b>186</b>
7.1	<i>Introduction</i> .....	186
7.2	<i>Methods</i> .....	186
7.2.1	Identification of the search term.....	187
7.2.2	Categorisation of collocates, and selection of 'salient' domains.....	190
7.2.3	Manual disambiguation of B category collocates.....	191
7.3	<i>Analysis</i> .....	192
7.3.1	A: General and abstract terms .....	194
7.3.2	I: Money and commerce in industry .....	199
7.3.3	N: Numbers and measurement .....	201
7.3.4	S: Social actions, states and processes .....	201



7.4	<i>Summary</i> .....	217
7.4.1	Of findings.....	217
7.4.2	Of method .....	218
<b>Chapter 8:</b>	<b>Semantic preference of frequent naming strategies for people with</b>	
<b>HIV/AIDS</b>	<b>.....</b>	<b>219</b>
8.1	<i>Introduction</i> .....	219
8.2	<i>Method: Targeting search terms</i> .....	220
8.3	<i>Exploring naming strategies</i> .....	222
8.3.1	Incidence of naming strategies.....	222
8.3.2	Collocation and semantic preference of naming strategies.....	225
8.4	<i>Collocational profiles of naming strategies</i> .....	228
8.4.1	Category A: general and abstract terms .....	228
8.4.2	Category B: The body and the individual.....	235
8.4.3	Category G: Government and public .....	245
8.4.4	Category H: Architecture, housing and the home .....	247
8.4.5	Category S: Social actions, states and processes .....	254
8.4.6	Category X: Psychological actions, states and processes.....	260
8.5	<i>Summary</i> .....	264
8.5.1	Of results .....	264
8.5.2	Of method .....	266
<b>Chapter 9:</b>	<b>Diachronic analysis of an identity group in the AIDS/HIV corpus.....</b>	<b>268</b>
9.1	<i>Introduction</i> .....	268
9.2	<i>Methodology</i> .....	269
9.2.1	Targeting search terms.....	269
9.1.1	Identifying key semantic categories of collocates .....	271
9.3	<i>Analysis</i> .....	276
9.3.1	Category A: General and abstract terms .....	277
9.3.2	Category G: Government and public .....	292
9.3.3	Category S: Social actions, states and processes .....	301
9.4	<i>Summary</i> .....	319
9.4.1	Of results .....	320
9.4.2	Of method .....	321
<b>Chapter 10:</b>	<b>Conclusion.....</b>	<b>323</b>
10.1	<i>Introduction</i> .....	323
10.2	<i>Discussion of findings</i> .....	323
10.2.1	Moral Panic versus Risk Society .....	323
10.2.2	Deviancy Doubling and Social Sequestering .....	324
10.2.3	(Lack of) Agency .....	325
10.2.4	Research questions revisited .....	326
10.3	<i>Discussion of methods</i> .....	330

10.3.1	Dual corpora and matched studies.....	330
10.3.2	Corpus-based critical discourse analysis.....	331
10.3.3	Proportional semantic collocation .....	332
10.4	<i>Reflecting on the study</i> .....	333
10.5	<i>Further research</i> .....	335
10.6	<i>Concluding comments</i> .....	336
<b>References</b> .....		<b>338</b>
<b>Appendices</b> .....		<b>354</b>
	<i>Appendix A: Full USAS semantic tagset</i> .....	354
	<i>Appendix B: Full collocation table for people in the Katrina corpus (Chapter 4)</i> .....	359
	<i>Appendix C: Full collocation table for resident in the Katrina corpus (Chapter 5)</i> .....	368
	<i>Appendix D: Full collocation table for victim in the Katrina corpus (Chapter 5)</i> .....	371
	<i>Appendix E: Full collocation table for evacuee in the Katrina corpus (Chapter 5)</i> .....	373
	<i>Appendix F: Full collocation table for survivor in the Katrina corpus (Chapter 5)</i> .....	375
	<i>Appendix G: Full collocation table for UNCONTESTED uses of refugee in the Katrina corpus (Chapter 6)</i> .....	376
	<i>Appendix H: Full collocation table for “people with (AIDS A.I.D.S. HIV H.I.V. GRID G.R.I.D.)” in the AIDS/HIV corpus (Chapter 7)</i> .....	378
	<i>Appendix I: Full collocation table for patient in the AIDS/HIV corpus (Chapter 8)</i> .....	380
	<i>Appendix J: Full collocation table for victim in the AIDS/HIV corpus (Chapter 8)</i> .....	384
	<i>Appendix K: Full collocation table for sufferer in the AIDS/HIV corpus (Chapter 8)</i> .....	386
	<i>Appendix L: Full collocation table for carrier in the AIDS/HIV corpus (Chapter 8)</i> .....	387
	<i>Appendix M: Full collocation table for “(gay gays homosexual homosexuals)” in the AIDS/HIV 1980s subcorpus (Chapter 9)</i> .....	388
	<i>Appendix N: Full collocation table for “(gay gays homosexual homosexuals)” in the AIDS/HIV 1990s subcorpus (Chapter 9)</i> .....	390
	<i>Appendix O: Full collocation table for “(gay gays homosexual homosexuals)” in the AIDS/HIV 2000s subcorpus (Chapter 9)</i> .....	393

## Tables

<b>Table 3-1:</b> Count Hurricane Katrina-related articles by source name .....	54
<b>Table 3-2:</b> Count of Hurricane Katrina-related articles shown as a representative percentage of all articles in the Factiva database, August 2005 – August 2006 .....	55
<b>Table 3-3:</b> Count of AIDS/HIV-related articles by source name.....	57
<b>Table 3-4:</b> Count of AIDS/HIV-related articles shown as a representative percentage of all articles in the Factiva database, 1981 – 2009.....	58
<b>Table 4-1:</b> The significant adjectival collocate lemmas of <i>people</i> grouped by USAS semantic category.....	79
<b>Table 4-2:</b> USAS A category collocates of <i>people</i> , listed in order of descending log likelihood value .....	79
<b>Table 4-3:</b> USAS X category collocates of <i>people</i> , listed in order of descending log likelihood value .....	84
<b>Table 4-4:</b> USAS B category collocates of <i>people</i> , listed in order of descending log likelihood value .....	84
<b>Table 4-5:</b> The sole collocate of <i>people</i> from the USAS L category.....	87
<b>Table 4-6:</b> The sole collocate of <i>people</i> from the USAS F category.....	88
<b>Table 4-7:</b> The sole collocate of <i>people</i> from the USAS H category.....	89
<b>Table 4-8:</b> USAS I category collocates of <i>people</i> .....	92
<b>Table 4-9:</b> Adjectival collocates of <i>people</i> from USAS Broad Category I: Money and commerce in industry .....	92
<b>Table 4-10:</b> Frequency of nodes from the ‘low-income set’.....	93
<b>Table 4-11:</b> Secondary collocates of the ‘low-income set’ .....	93
<b>Table 4-12:</b> Frequency of nodes from the ‘moderate/high-income’ set.....	95
<b>Table 4-13:</b> Secondary collocates of the ‘moderate/high-income set’.....	95
<b>Table 4-14:</b> USAS N category collocates of <i>people</i> , listed in order of descending log likelihood value.....	99
<b>Table 4-15:</b> USAS S category collocates of <i>people</i> , listed in order of descending log likelihood value.....	100
<b>Table 4-16:</b> Frequency and dispersion of <i>care/caring/cares about black people</i> listed by month and year.....	101
<b>Table 4-17:</b> Types of quotation and recontextualisation of Kanye West’s quote.....	104
<b>Table 4-18:</b> Frequency of government actors depicted as <i>caring</i> or <i>not caring</i> about social actors .....	105
<b>Table 4-19:</b> Secondary adjectival collocates of <i>black + people</i> and <i>white + people</i> .....	107
<b>Table 5-1:</b> List of the top 20 frequent common plural nouns with human actors in boldface .....	116
<b>Table 5-2:</b> Top ‘human’ collocates of <i>Katrina’s</i> in the +/-4 span.....	118
<b>Table 5-3:</b> Top ‘human’ collocates of <i>Katrina’s</i> in the +1 span.....	118

<b>Table 5-4:</b> Top ‘human’ collocates of <i>of Katrina</i> in the +/-4 span .....	118
<b>Table 5-5:</b> Top ‘human’ collocates of <i>of Katrina</i> in the -1 span.....	118
<b>Table 5-6:</b> Dictionary definitions of <i>resident, evacuee, victim</i> and <i>survivor</i> .....	119
<b>Table 5-7:</b> Collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> , categorized by part of speech and USAS broad semantic categories .....	123
<b>Table 5-8:</b> USAS M category adjectival collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> .....	124
<b>Table 5-9:</b> USAS N category adjectival collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> .....	127
<b>Table 5-10:</b> USAS S category nominal collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> .....	130
<b>Table 5-11:</b> USAS W category nominal collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> .....	133
<b>Table 5-12:</b> USAS H category verbal collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> .....	135
<b>Table 5-13:</b> USAS M category verbal collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> .....	136
<b>Table 5-14:</b> USAS S category verbal collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> .....	142
<b>Table 5-15:</b> USAS X category verbal collocates of <i>resident(s), victim(s), evacuee(s),</i> and <i>survivor(s)</i> .....	144
<b>Table 6-1:</b> A selection of three pre-Katrina dictionary definitions of <i>refugee</i> .....	155
<b>Table 6-2:</b> Frequency of <i>refugee</i> in various categories of semantic sense.....	157
<b>Table 6-3:</b> Collocates of <i>refugee(s)</i> in the UNCONTESTED category .....	163
<b>Table 6-4:</b> Frequency of verbal collocates of UNCONTESTED <i>refugee(s)</i> by USAS semantic category and transitivity .....	165
<b>Table 6-5:</b> Frequency of alternative naming strategies for <i>refugee(s)</i> proposed in the CONTESTED set.....	171
<b>Table 6-6:</b> Frequency of <i>refugee(s)</i> in the DICTIONARY category by reference type and statement on applicability of the term .....	173
<b>Table 6-7:</b> Frequency of secondary components of bi- and tri-grams following <i>refugee</i> in the CAMPS subcategory .....	181
<b>Table 7-1:</b> Frequency of all solutions arising from a search for “people with (AIDS A.I.D.S. HIV H.I.V. GRID G.R.I.D.)” in the AIDS/HIV corpus, expressed by number of occurrences and percentage of all occurrences .....	188
<b>Table 7-2:</b> Frequency of solutions arising from search for “people with (AIDS A.I.D.S. HIV H.I.V. GRID G.R.I.D.)” and “(PWA PWAs)” in the AIDS/HIV corpus, expressed by decade of occurrence.....	189
<b>Table 7-3:</b> Collocates of <i>PWA/H</i> , categorised into USAS semantic categories, and expressed as a percentage of the overall number of collocates .....	193

<b>Table 7-4:</b> USAS A category collocates of <i>PWA/H</i> , listed in order of descending log likelihood value.....	194
<b>Table 7-5:</b> USAS I category collocates of <i>PWA/H</i> , listed in order of descending log likelihood value .....	199
<b>Table 7-6:</b> Comparison between ‘lower-income’ collocates of <i>PWA</i> (AIDS/HIV corpus) and <i>people</i> (Katrina corpus), expressed by Mutual Information value and relative rank within the lower-income collocate set.....	200
<b>Table 7-7:</b> USAS N category collocates of <i>PWA/H</i> , listed in order of descending log likelihood value.....	201
<b>Table 7-8:</b> USAS S category collocates of <i>PWA/H</i> , listed in order of descending log likelihood value.....	202
<b>Table 7-9:</b> Contextual agents of verbal collocates help, treat, <i>serve</i> , <i>assist</i> , <i>protect</i> , <i>enable</i> , and <i>counsel</i> , expressed by number of instances of concordance lines of a given agent and percentage of 777 overall concordance lines of these verb collocates.....	206
<b>Table 7-10:</b> Common agents <i>helping PWA/H</i> , with frequency of <i>by</i> -phrases and secondary clauses.....	207
<b>Table 8-1:</b> Frequency of ‘human’ plural common nouns in the AIDS/A.I.D.S./HIV/H.I.V. <i>x</i> form	221
<b>Table 8-2:</b> Frequency of ‘human’ plural common nouns in the <i>x</i> of AIDS/A.I.D.S./HIV/H.I.V. form .....	221
<b>Table 8-3:</b> Frequency of naming strategies by decade—expressed as raw hits, dispersion over total texts, and frequency per million words in the category .....	223
<b>Table 8-4:</b> Collocates of naming strategies, categorised into USAS semantic categories, and expressed as a percentage of the overall number of collocates for a given naming strategy .....	227
<b>Table 8-5:</b> Collocates for each naming strategy from USAS category A.....	229
<b>Table 8-6:</b> USAS Category A subcategory ‘Exclusion’ collocates by naming strategy.....	229
<b>Table 8-7:</b> USAS Category A subcategory ‘Plight’ collocates by naming strategy.....	231
<b>Table 8-8:</b> USAS Category A subcategory ‘Classification’ collocates by naming strategy.....	234
<b>Table 8-9:</b> USAS Category A subcategory ‘Exclusivizers etc.’ collocates by naming strategy .....	235
<b>Table 8-10:</b> Collocates for each naming strategy from USAS category B.....	236
<b>Table 8-11:</b> USAS Category B subcategory ‘Conditions’ collocates by naming strategy .....	237
<b>Table 8-12:</b> USAS Category B subcategory ‘Drugs and Treatment’ collocates by naming strategy .....	238
<b>Table 8-13:</b> USAS Category B subcategory ‘Medical services and sites’ collocates by naming strategy.....	238
<b>Table 8-14:</b> USAS Category B subcategory ‘Illness’ collocates by naming strategy .....	240
<b>Table 8-15:</b> Frequency of adverbs modifying the collocation of <i>patient(s)</i> with <i>ill</i> .....	240
<b>Table 8-16:</b> USAS Category B subcategory ‘Other groups’ collocates by naming strategy .....	243
<b>Table 8-17:</b> USAS Category G collocates by naming strategy .....	245

<b>Table 8-18:</b> Frequency and distribution of <i>discrimination</i> and <i>discriminate</i> collocating with <i>patient(s)</i> , <i>victim(s)</i> , and <i>sufferer(s)</i> .....	245
<b>Table 8-19:</b> Frequency of various sources/sites of <i>discrimination</i> by appearance in concordance lines of congruently realized PWA naming strategies.....	246
<b>Table 8-20:</b> USAS Category H category collocates by naming strategy .....	248
<b>Table 8-21:</b> USAS Category H subcategory 'Homeless' collocates by naming strategy .....	248
<b>Table 8-22:</b> Frequency of collocation with <i>homeless</i> by various naming strategies, by parallel and compound structuring.....	249
<b>Table 8-23:</b> Frequency of additional social groups appearing in parallel structures.....	249
<b>Table 8-24:</b> USAS Category H subcategory 'Residence' collocates by naming strategy .....	252
<b>Table 8-25:</b> USAS Category S category collocates by naming strategy .....	254
<b>Table 8-26:</b> USAS Category S subcategory 'Care and Advocacy' collocates by naming strategy.....	255
<b>Table 8-27:</b> USAS Category S subcategory 'Commemoration' collocates by naming strategy .....	256
<b>Table 8-28:</b> USAS Category S subcategory 'Sexuality' collocates by naming strategy.....	257
<b>Table 8-29:</b> USAS Category S subcategory 'Groups and Belonging' collocates by naming strategy .....	258
<b>Table 8-30:</b> Collocates from USAS Category X: psychological actions, states and processes.....	260
<b>Table 9-1:</b> Frequency of node words associated with homosexuality, shown in raw hits and as a percentage of all of the search results.....	270
<b>Table 9-2:</b> Frequency of node words associated with homosexuality, shown by distribution and dispersion over three decades.....	271
<b>Table 9-3:</b> Collocates of the gay set by decade, categorised into USAS semantic categories, and expressed as a percentage of the overall number of collocates per decade.....	274
<b>Table 9-4:</b> Social groups appearing in concordance lines where the gay set collocates with high-risk.....	278
<b>Table 9-5:</b> n-grams resulting from the collocation of <i>gay</i> with <i>activity</i> .....	281
<b>Table 9-6:</b> Frequency of actions and roles in concordance lines of <i>exclude</i> .....	285
<b>Table 9-7:</b> Concepts/n-grams co-occurring in context with the collocate <i>oppose</i> .....	289
<b>Table 9-8:</b> Concepts/n-grams co-occurring in context with the collocate <i>oppose</i> .....	297
<b>Table 9-9:</b> Concepts/n-grams co-occurring in context with the collocate <i>legalize</i> .....	297
<b>Table 9-10:</b> Frequency breakdown of the collocation between <i>right</i> and the search nodes, expressed in raw hits and percentage of total collocation of search string per decade.....	300
<b>Table 9-11:</b> Behaviours <i>engaged</i> in, when <i>engage</i> collocates with the gay set, expressed by semantic category and frequency of occurrence in concordance lines .....	303
<b>Table 9-12:</b> Concepts/n-grams co-occurring in context with the collocate <i>acceptance</i> .....	309
<b>Table 9-13:</b> Frequency breakdown of the collocation between <i>community</i> and the search nodes, expressed in raw hits and percentage of total collocation of search string per decade.....	317

## Figures

<b>Figure 3-1:</b> Visual representation of Hurricane Katrina articles in proportion to all Factiva articles by year.....	56
<b>Figure 3-2:</b> Visual representation of AIDS/HIV articles in proportion to all Factiva articles by year.....	59
<b>Figure 3-3:</b> Sample concordance line format.....	64
<b>Figure 3-4:</b> Structure of USAS major discourse field distinctions (from Archer et al., 2002, p. 2)	72
<b>Figure 4-1:</b> Sample concordance line of <i>missing people</i> .....	80
<b>Figure 4-2:</b> Sample concordance lines of the collocation of <i>ordinary</i> and <i>people</i> .....	81
<b>Figure 4-3:</b> Sample concordance lines of the collocation of <i>innocent</i> and <i>people</i> .....	82
<b>Figure 4-4:</b> Sample concordance lines of the collocation of <i>desperate</i> and <i>people</i> .....	83
<b>Figure 4-5:</b> Sample concordance lines of the collocation of <i>wonderful</i> and <i>people</i> .....	83
<b>Figure 4-6:</b> Sample concordance lines of <i>talented</i> and <i>smart</i> collocating with <i>people</i> .....	84
<b>Figure 4-7:</b> Sample concordance lines of the collocation of <i>disabled</i> and <i>people</i> .....	85
<b>Figure 4-8:</b> Sample concordance lines of the collocation of <i>sick</i> and <i>people</i> .....	86
<b>Figure 4-9:</b> Sample concordance lines of the collocation of <i>dead</i> and <i>people</i> .....	87
<b>Figure 4-10:</b> Sample concordance lines of the collocation of <i>dead</i> and <i>people</i> .....	88
<b>Figure 4-11:</b> Sample concordance lines of the collocation of <i>hungry</i> and <i>people</i> .....	89
<b>Figure 4-12:</b> Instances of <i>homeless</i> in each month of the Katrina corpus, standardised to frequency per million.....	90
<b>Figure 4-13:</b> Sample concordance lines of the collocation of <i>homeless</i> and <i>people</i> .....	91
<b>Figure 4-14:</b> Sample concordance lines of low income set .....	94
<b>Figure 4-15:</b> Sample concordance lines of low-income set linked to other social groups .....	95
<b>Figure 4-16:</b> Sample concordance lines of moderate/high-income set.....	96
<b>Figure 4-17:</b> Sample concordance lines of low-income and moderate/high-income set collocating .....	96
<b>Figure 4-18:</b> The sole collocate of <i>people</i> from the USAS M category.....	97
<b>Figure 4-19:</b> Sample concordance lines of <i>displaced people</i> preceded by 'the' .....	97
<b>Figure 4-20:</b> Sample concordance lines of alternate forms of the lemma <i>care</i> .....	102
<b>Figure 4-21:</b> Sample concordance lines of <i>American</i> collocating with <i>people</i> .....	107
<b>Figure 4-22:</b> Sample concordance lines of <i>American</i> collocating with <i>people</i> .....	108
<b>Figure 4-23:</b> Sample concordance lines of <i>generous</i> collocating with <i>people</i> .....	109
<b>Figure 4-24:</b> Sample concordance lines of <i>resilient</i> collocating with <i>people</i> .....	110
<b>Figure 4-25:</b> Sample concordance lines of <i>stranded</i> collocating with <i>people</i> .....	111
<b>Figure 4-26:</b> USAS T category collocates of <i>people</i> , listed in order of descending log likelihood value.....	112
<b>Figure 4-27:</b> Sample concordance lines of <i>young</i> collocating with both the white and <i>black</i> sets. ....	113

<b>Figure 5-1:</b> Frequency of naming strategies over time, expressed in frequency/million .....	121
<b>Figure 5-2:</b> Sample concordance lines of <i>remaining</i> and <i>returning</i> collocating with <i>residents</i> ....	126
<b>Figure 5-3:</b> Sample concordance lines of the collocation between <i>evacuated</i> and <i>resident(s)</i> .....	126
<b>Figure 5-4:</b> Sample concordance line of the collocation between <i>nearby</i> and <i>resident(s)</i> .....	128
<b>Figure 5-5:</b> Sample concordance lines of the collocation between <i>estimated</i> and <i>evacuee(s)</i> .....	129
<b>Figure 5-6:</b> Sample concordance lines of the collocation between figures and <i>evacuee(s)</i> and <i>resident(s)</i> .....	129
<b>Figure 5-7:</b> Sample concordance lines of the collocation between <i>resident(s)</i> and <i>native</i> .....	131
<b>Figure 5-8:</b> Sample concordance lines of the collocation between <i>survivor(s)</i> and <i>son, wife, and daughter</i> .....	132
<b>Figure 5-9:</b> Sample concordance lines of <i>disaster, tsunami, earthquake</i> , collocating with <i>victim(s)</i> .....	134
<b>Figure 5-10:</b> Sample concordance line of <i>shelters</i> (plural common noun) mis-tagged as a verb	135
<b>Figure 5-11:</b> Sample concordance lines of the collocation between naming strategies and <i>scatter</i> .....	137
<b>Figure 5-12:</b> Sample concordance lines of the collocation between naming strategies and <i>flee</i>	138
<b>Figure 5-13:</b> Sample concordance lines of the collocation between naming strategies and <i>stay</i>	139
<b>Figure 5-14:</b> Sample concordance lines of the collocation <i>resident(s)</i> and <i>evacuate</i> .....	140
<b>Figure 5-15:</b> Sample concordance lines of the collocation between <i>resident(s)</i> and <i>return</i> .....	140
<b>Figure 5-16:</b> Sample concordance lines of the collocation between naming strategies and <i>rescue/strand</i> .....	143
<b>Figure 5-17:</b> Sample concordance lines of the collocation between <i>resident(s)</i> and <i>heed</i> .....	145
<b>Figure 5-18:</b> Sample concordance lines of <i>identify</i> and <i>search</i> .....	145
<b>Figure 5-19:</b> Sample concordance lines showing metaphorical use of <i>pour</i> and <i>wave</i> .....	147
<b>Figure 5-20:</b> Sample concordance lines showing metaphorical use of <i>absorb</i> .....	148
<b>Figure 5-21:</b> Sample concordance lines showing metaphorical use of <i>swell</i> .....	148
<b>Figure 5-22:</b> Sample concordance lines showing metaphorical use of <i>swell</i> .....	149
<b>Figure 5-23:</b> Sample concordance lines showing metaphorical use of <i>stream</i> and <i>trickle</i> .....	150
<b>Figure 6-1:</b> Overall frequency of <i>refugee(s)</i> in the Katrina corpus by month, expressed in frequency per million words. ....	156
<b>Figure 6-2:</b> Frequency of compressed semantic categories of sense of <i>refugee(s)</i> over time.....	161
<b>Figure 6-3:</b> Frequency of compressed semantic categories of sense of <i>refugee(s)</i> over time (clipped).....	162
<b>Figure 6-4:</b> Sample concordance lines of <i>refugee crisis</i> .....	164
<b>Figure 6-5:</b> Sample concordance lines illustrating water metaphors associated with <i>refugee(s)</i> .....	166
<b>Figure 6-6:</b> Overall frequency of CONTESTED <i>refugee(s)</i> application by direction (self-reference versus other-reference) and appraisal (appropriateness versus inappropriateness of application).....	168
<b>Figure 6-7:</b> Sample concordance lines illustrating CONTESTED use of <i>refugee(s)</i> .....	169



<b>Figure 6-8:</b> Sample concordance lines illustrating arguments <i>against</i> the use of <i>refugee(s)</i> .....	169
<b>Figure 6-9:</b> Sample concordance lines illustrating ‘un-American’ arguments against <i>refugee(s)</i> .....	170
<b>Figure 6-10:</b> Sample concordance lines illustrating alternative naming strategies provided in place of <i>refugee(s)</i> .....	172
<b>Figure 6-11:</b> Sample concordance lines of DEFINITIONS being incorporated to argue against applicability of <i>refugee(s)</i> in referring to PAKs.....	174
<b>Figure 6-12:</b> Sample concordance of an alternative naming strategy for <i>refugee(s)</i> offered following the political definition subcategory.....	175
<b>Figure 6-13:</b> Frequency of <i>refugee(s)</i> by stance category and definition source.....	178
<b>Figure 6-14:</b> Sample concordance lines of PARALLELISMS between PAKs and Dust Bowl victims .....	179
<b>Figure 6-15:</b> Sample concordance lines belonging to the CAMPS subcategory.....	181
<b>Figure 6-16:</b> A selection of three post-Katrina dictionary definitions of <i>refugee</i> .....	183
<b>Figure 7-1:</b> Frequency of solutions arising from a search for “people with (AIDS A.I.D.S. HIV H.I.V. GRID G.R.I.D.)” in the AIDS/HIV corpus, expressed by frequency/million words in category (year).....	189
<b>Figure 7-2:</b> Instances of <i>quarantining</i> collocating with PWA/H.....	195
<b>Figure 7-3:</b> Instances of <i>quarantine</i> collocating with PWA/H.....	196
<b>Figure 7-4:</b> Instances of <i>isolate</i> collocating with PWA/H.....	197
<b>Figure 7-5:</b> Instances of <i>stigmatize</i> collocating with PWA/H.....	198
<b>Figure 7-6:</b> Instances of the Agencies/Organisations/Centres agent in concordance lines where <i>help</i> and <i>treat</i> collocate with PWA/H.....	209
<b>Figure 7-7:</b> Instances of the Individual agent in concordance lines where <i>help</i> and <i>treat</i> collocate with PWA/H.....	210
<b>Figure 7-8:</b> Instances of the Drugs/Treatments agent in concordance lines where <i>help</i> and <i>treat</i> collocate with PWA/H.....	210
<b>Figure 7-9:</b> Instances of the Financial Assistance agent in concordance lines where <i>help</i> and <i>treat</i> collocate with PWA/H.....	211
<b>Figure 7-10:</b> Instances of the Government/Legislature agent in concordance lines where <i>help</i> and <i>treat</i> collocate with PWA/H.....	212
<b>Figure 7-11:</b> Instances of the Pets agent in concordance lines where <i>help</i> and <i>treat</i> collocate with the PWA search terms.....	212
<b>Figure 7-12:</b> Instances of the ‘Emotional Sense’ of the PWA/H collocate <i>treat</i> .....	213
<b>Figure 7-13:</b> All concordance lines arising from the collocation of <i>deserve</i> with PWA/H.....	215
<b>Figure 8-1:</b> Frequency of naming strategies by year—expressed as frequency/million words in the category.....	224
<b>Figure 8-2:</b> Sample concordance lines showing the collocation between <i>shun</i> and <i>AIDS patients</i> where doctors/medical programs are the agents.....	230

<b>Figure 8-3:</b> Sample concordance lines showing the collocation between <i>shun</i> and <i>AIDS patients</i> with agents other than doctors/medical establishments.....	231
<b>Figure 8-4:</b> Sample concordance lines of <i>plight</i> collocating with <i>AIDS victims</i> .....	232
<b>Figure 8-5:</b> Sample concordance lines of <i>confidentiality</i> collocating with <i>AIDS patients</i> .....	233
<b>Figure 8-6:</b> Sample concordance lines for <i>victim(s)</i> collocating with <i>dying</i> and <i>suffering</i> .....	239
<b>Figure 8-7:</b> Sample concordance line of AIDS patient collocating redundantly with <i>ill</i> .....	240
<b>Figure 8-8:</b> Sample concordance lines for <i>victim(s)</i> collocating with <i>dying</i> and <i>suffering</i> .....	241
<b>Figure 8-9:</b> Sample concordance line for <i>carrier(s)</i> collocating with <i>infect</i> .....	242
<b>Figure 8-10:</b> Sample concordance lines for <i>carrier(s)</i> collocating with <i>infect</i> .....	242
<b>Figure 8-11:</b> Sample concordance lines for <i>patient(s)</i> collocating with <i>leper</i> .....	244
<b>Figure 8-12:</b> Sample concordance line of <i>victim(s)</i> collocating with <i>hemophiliac</i> .....	245
<b>Figure 8-13:</b> All concordance lines showing human agents as sources of <i>discrimination</i> .....	247
<b>Figure 8-14:</b> Concordance lines indicating past or future states of homelessness.....	250
<b>Figure 8-15:</b> Concordance lines indicating additional activities of <i>homeless</i> PWAs.....	251
<b>Figure 8-16:</b> Concordance lines illustrating the collocation between USAS H subcategory 'Residence' collocates and various naming strategies.....	253
<b>Figure 8-17:</b> Concordance lines illustrating the collocation between <i>occupy</i> and <i>patient(s)</i> .....	254
<b>Figure 8-18:</b> Sample concordance lines showing 'commemoration' collocates of <i>victim(s)</i> .....	256
<b>Figure 8-19:</b> Sample concordance lines showing <i>embrace</i> collocating with <i>victim(s)</i> .....	259
<b>Figure 8-20:</b> Sample concordance lines showing 'commemoration' collocates of <i>victim(s)</i> .....	259
<b>Figure 8-21:</b> Sample concordance lines showing <i>known</i> collocating with various naming strategies.....	262
<b>Figure 8-22:</b> Sample concordance lines of the collocation between <i>suspected</i> and <i>carrier(s)</i> .....	263
<b>Figure 8-23:</b> Sample concordance lines of the collocation between <i>knowingly</i> and <i>carrier(s)</i> .....	263
<b>Figure 8-24:</b> Sample concordance lines of the collocation between <i>innocent</i> and <i>victim(s)</i> .....	264
<b>Figure 9-1:</b> Count of collocates in each semantic category, stacked by time.....	272
<b>Figure 9-2:</b> Venn diagram of collocates from all USAS broad semantic categories for <i>gay/gays/homosexual/homosexuals</i> in the AIDS/HIV corpus depicted over time.....	276
<b>Figure 9-3:</b> Venn diagram of USAS Category A collocates for <i>gay/gays/homosexual/homosexuals</i> depicted over time.....	277
<b>Figure 9-4:</b> Sample concordance lines where the gay set collocates with <i>high-risk</i> .....	279
<b>Figure 9-5:</b> Sample concordance lines where the gay set collocates with <i>active</i> .....	280
<b>Figure 9-6:</b> Sample concordance lines where <i>homosexual</i> collocates with <i>activity</i> .....	282
<b>Figure 9-7:</b> Sample concordance lines where <i>primarily</i> collocates with the node words.....	283
<b>Figure 9-8:</b> Sample concordance lines where <i>predominantly</i> collocates with the node words...	284
<b>Figure 9-9:</b> Sample concordance lines where <i>predominantly</i> collocates with the node words...	284
<b>Figure 9-10:</b> Sample concordance lines where <i>exclude</i> collocates with the node words.....	286
<b>Figure 9-11:</b> Sample concordance lines where <i>outing</i> collocates with the node words.....	288
<b>Figure 9-12:</b> Sample concordance lines where <i>mainstream</i> collocates with the node words.....	289
<b>Figure 9-13:</b> Sample concordance lines where <i>mainstream</i> collocates with the node words.....	291

**Figure 9-14:** Sample concordance lines where *liberation* collocates with the node words ..... 292

**Figure 9-15:** Venn diagram of USAS Category G collocates for *gay/gays/homosexual/homosexuals* depicted over time..... 293

**Figure 9-16:** Sample concordance lines where *elected* collocates with the node words ..... 295

**Figure 9-17:** Sample concordance lines where the gay set collocates with *agenda*..... 296

**Figure 9-18:** Sample concordance lines where *amendment* and *constitutional* collocate with the node words ..... 298

**Figure 9-19:** Sample concordance lines where *republican* collocates with the node words..... 298

**Figure 9-20:** Sample concordance lines where *legalize* and *equal* collocate with the node words ..... 299

**Figure 9-21:** Venn diagram of USAS Category S collocates for *gay/gays/homosexual/homosexuals* depicted over time..... 302

**Figure 9-22:** Sample concordance line where *promiscuous* collocates with the node words..... 305

**Figure 9-23:** Sample concordance line where *leader* collocates with the node words..... 305

**Figure 9-24:** Sample concordance line where *cater* collocates with the node words ..... 306

**Figure 9-25:** Sample concordance line where *prejudice* and *tolerance* collocate with the node words..... 307

**Figure 9-26:** Sample concordance lines where *latino* collocates with the node words..... 308

**Figure 9-27:** Sample concordance lines where *acceptance* collocates with the node words..... 309

**Figure 9-28:** Sample concordance lines where *parade* collocates with the node words..... 310

**Figure 9-29:** Sample concordance lines where *adoption* collocates with the node words ..... 311

**Figure 9-30:** Sample concordance lines where *adoption* collocates with the node words ..... 313

**Figure 9-31:** Sample concordance lines where *lesbian* collocates with the node words ..... 315

**Figure 9-32:** Sample concordance lines where *bisexual* collocates with the node words..... 315

**Figure 9-33:** Sample concordance lines where *heterosexual* collocates with the node words .... 316

**Figure 9-34:** Sample concordance lines where *straight* collocates with the node words ..... 316

**Figure 9-35:** Sample concordance lines where the gay set collocates with *community*..... 318

**Figure 9-36:** Sample concordance lines where the gay set collocates with lifestyle ..... 319

# **Chapter 1: Introduction**

## **1.1 Orientation**

This study investigates the construal of social actors in American ‘act of God’ news reporting, with an emphasis on detecting markers of identity and in-/out-group qualities as indicators of underlying ideology. Large, custom-collected corpora on two separate catastrophic events (Hurricane Katrina and the AIDS/HIV epidemic) form the basis of analysis. Adopting a corpus-based critical discourse analytical approach, I explore frequent naming strategies and salient identity markers through the lens of semantic collocation (to include semantic preference and semantic prosody). I aim to discover whether established methods of ‘othering’ (based upon concrete events offering a catalyst) remain effective during ‘acts of God’, where blame cannot be so easily negotiated or even assigned completely to human agency. This is achieved through a series of three matched analyses across the two corpora, allowing me to consider whether representations of the ‘other’ in differing circumstances share common linguistic constructions. The development and refinement of a generalizable method of analysing and comparing corpora of varying (but somewhat related) topics, and vastly differing sizes and periods of diachrony was an ongoing ancillary aim of this thesis, in support of the analytical goals.

For each corpus, I aim to first describe the most frequent naming strategy as a gauge of the discourses most commonly constructed around the people most affected by the storm and the epidemic. The scope shifts in subsequent chapters to contrastive analyses, wherein multiple naming strategies are compared to illustrate the functional choices made when selecting one nomination strategy over another. Finally I consider the corpora diachronically, studying the changing use of a single word or construal of a single social group over time. In these sections, I propose novel methods of introducing time series analysis incorporating reference to lexicography and semantic collocation.

Corpora on two separate topics have been consciously selected for a variety of reasons, including the fact that both events hold deep personal meaning for me. I was born in a rural town just outside of New Orleans, which was (and continues to be) hugely affected by Hurricane Katrina. Furthermore, I have had a family member pass away from AIDS-related complications. It is my experiences as a (critical) consumer of the media that compel me to undertake this study unpicking popular press construal of the people of

the American Gulf Coast and people with AIDS. As a reasonably well-educated white, straight, female, raised for the most part in New York, I can be considered within the 'in' group of newspaper readerships, or certainly not at high risk of being personally affected by natural disasters in the Deep South, or by the spread of AIDS/HIV. However: I have been. Here, I problematize the media practice of institutional iteration of out-groups that, as a reader, I feel endangers society by creating false risk dichotomies and artificially distances citizens from one another in time of crisis. This can be done through selection of naming strategy, use of attribution, and construal of transitivity, as I will soon show in two corpora of varying sizes, extending through differing lengths of time, and based upon divergent event types.

With these strengths acknowledged, I must also concede that working with two corpora of very large size (nearing 200 million words in combination) presents its own unique challenge. Therefore, an additional contribution incorporated into each chapter of this thesis is the development of a method to systematically downsample results in corpus-based discourse analytical studies in a meaningful way: I call this *proportional semantic collocation*.

More detailed description of the research questions supporting these analytical choices can be found below, in section 1.2. This is followed by a brief dictionary of key terms (1.3). In sections 1.4 and 1.5, I describe Hurricane Katrina and AIDS/HIV respectively, both in terms of their financial-political impacts, and their social effects in the American context. My rationale for undertaking this study is detailed in section 1.6, and the structure of the entire thesis is detailed in section 1.7, just before summary section 1.8.

## **1.2 Research Questions**

Building upon the research tradition of 'othering' and the critical theory of subversive perpetuation of prejudice, this research will turn from generalisation of culpability following crimes perpetrated by groups of people, and begin to explore media discourse surrounding events lacking obvious, blame-worthy 'others', specifically Hurricane Katrina and the AIDS epidemic. To these ends, **the overarching research question is as follows:**

1. How does American media discourse negotiate the concept of risk in 'Acts of God' reporting on Hurricane Katrina and the AIDS/HIV epidemic, particularly in relation to attributes of identity such as ethnicity, class, and sexuality?

A series of **actionable research questions** in support of this overarching research question are addressed in each analysis chapter.

Chapters 4 and 7 respond to the sub-questions:

2. Which identities are associated with the highest frequency of in- and out-group qualities in the Hurricane Katrina and AIDS/HIV corpora?
  - a. To what extent do these identities correlate to existing 'folk devil' categories? What does negative/positive construal of identities indicate about the underlying ideologies of the publications represented in the corpus?

Chapters 5 and 8 respond to the sub-questions:

3. What are the most frequent naming strategies employed in Hurricane Katrina and AIDS/HIV reporting?
  - a. What can corpus-based discourse analysis of associated attributes, transitivity, and cognitive metaphors tell us about the in- and/or out-group qualities attached to the people that the naming strategies represent?

Finally, Chapters 6 and 9 respond to the sub-questions:

4. What evidence does corpus-based discourse analysis of the two corpora give of changing or resistant discourses surrounding othered social groups over time?
  - a. What kind of contextual or cultural changes result in shifting preference for terms or interest in discourses surrounding certain identities?

### **1.3 Key terms**

Before proceeding to an overview of the events and a justification for the study, I will briefly outline terms used throughout this thesis.

#### **1.3.1 Discourse**

Discourse is extremely polysemous, being traditionally used to refer to something as general as a length of written or spoken language or, more recently, to refer to language of a certain genre enacted in a specific register. This thesis adopts Reisigl and Wodak's (2001, p. 36) definition, where discourse "can be understood as a complex bundle of simultaneous and sequential interrelated linguistic acts which manifest themselves

within and across the social fields of action as thematically interrelated semiotic (spoken or written) tokens that belong to specific semiotic types (genres)". This explanation defines language as an ideological *social practice*, or a tool by which social domination and power hierarchies can be constructed and enforced (Habermas, 1971). The social aspect of discourse means that "surrounding any one object, event, person etc., there may be a variety of different discourses" each conveying different representations of reality (Burr, 1995, p. 48). In this way, we can refer to discourse on the basis of genre (i.e. *media discourse*), topic (i.e. *political discourse*), or ideology (i.e. *racist discourses*). All of these subcategories will be addressed in the following chapters.

### **1.3.2 Ideology**

Discourse is one method of encoding and communicating ideologies. This thesis adopts van Dijk's definition of ideology as "the basis of the social representations shared by members of a group. This means that ideologies allow people, as group members, to organize the multitude of social beliefs about what is the case, good or bad, right or wrong, *for them*, and to act accordingly" (1998, p. 8). Ideologies are not always (or even usually) based in scientific reality, and can be used as instruments of domination, competition, or even opposition (van Dijk, 1998, p. 11). Conceptualisations of power relationships and social roles are intrinsic to ideologies, and "values, attitudes, opinions, knowledge, and mental models of events" are also cognitively related (van Dijk, 1998, p. 12). It is for this reason that critical analysis of language in context is grounded here in analysis of ideologies presented in discourse.

### **1.3.3 Media/Press**

One powerful distributor of ideologically encoded discourse is the media, a collective noun referring to means of mass communication. This includes the component (historically) appearing in physical print: the Press. I adopt here a working definition of the Press as encompassing both the physical output (e.g. newspapers) and the discourse creators (e.g. reporters and editors). The purpose of using these as the basis of the study is that:

Media texts are perceived as dialogic, and the readings depend on the receivers and on the settings. Researchers presume, therefore, that readers/listeners or viewers interact with media (not only by writing letters to the editor but also by interpreting and understanding them in specific subjective ways). Media texts also depend on intertextual relations with many other genres, diachronically or synchronically. Texts relate to other texts, represented by the media, through quotes or indirect references, thus already adding particular meanings or

decontextualizing and recontextualizing meanings. Media thus produce and reproduce social meanings. (Wodak & Busch, 2004, p. 106)

It is my intent to explore the Press representation of identity in reporting on AIDS/HIV and Hurricane Katrina.

### **1.3.4 Identity**

Identity is understood here to be a form of social constructionism (Berger & Luckman, 1967; J. K. Hall, 1996; Kroskrity, 2000). I use identity as defined by De Fina, Schiffrin and Bamberg (2006):

a process that (1) takes place in concrete and specific interactional occasions, (2) yields constellations of identities instead of individual, monolithic constructs, (3) does not emanate from the individual, but results from processes of negotiation and entextualization (Bauman & Briggs, 1990) that are eminently social, and (4) entails “discursive work” (Zimmerman & Wieder, 1970)

A critical discourse analytical approach to the investigation of construal of identity is highly compatible as “both social and discourse practices frame, and in many ways define, the way individuals and groups present themselves to others, negotiate roles, and conceptualize themselves” (De Fina et al., 2006). The construction of an identity is often “related to the definition of categories for inclusion or exclusion of self and others, and to their identification with typical activities and routines” (De Fina et al., 2006). Groups sharing common ideologies often may also come to have shared, stable evaluations about what it means to perform an identity, including agreement on goals and parameters of accepted behaviour (Wodak & Busch, 2004, p. 120). While individuals oftentimes identify *themselves* as belonging to a group, identities may also be assigned by others. Outsiders, or those not belonging to an in-group, are sometimes referred to as ‘others’. This process is referred to henceforth as ‘othering’ (discussed more fully in Chapter 3). Identities can be constructed around anything from political beliefs to nationalities, sexual practices to biological characteristics.

Some parameters of inclusion in these groups can be found below. These are incorporated for ease of reference throughout the thesis, and to reflect the contents of the corpus, not due to my own agreement with these identifiers as appropriate delimiters for the categorization of human beings.



### **1.3.5 Race**

I use the word 'race' in this thesis to describe a group of persons sharing biological, genetically transmitted traits, though I agree with Ward Connerly of the American Civil Rights Institute (U.S. Commission on Civil Rights, 2009, p. 32) in that:

Were I empowered to do so, I would purge immediately from the public arena all classification of Americans based on skin color, texture of hair, nose width, lip size and slant of eyes—all attributes that serve to define "race." I recognize, however, that my view is an ideal that is not yet ready, if ever, for implementation in a society in which "race" seems to seep out of every pore of the public domain...

I acknowledge that this term has been criticised as being a socially constructed concept that has been used to reify both differences and similarities between groups of people and even to further stereotypes (see Palmié, 2007; Perry, 2011), and some modern researchers have preferred to use 'ethnicity' in its place. However, in the American context 'ethnicity' can refer to both biological traits and also cultural, linguistic, religious, or nationalistic similarities, so I believe that it is too broad for the purpose of this research. In the United Kingdom, this same distinction might be discussed by using the term 'ethnicity', though in the data itself, the term 'race' is used far more frequently, and it is on this specific dimension that arguments are based. Citizens are still asked to identify their 'race' in the U.S. census. The search terms used here (e.g. *black, white*) will only be able to provide results where, e.g., skin colour is cited as the variable, and (as an American, with an American understanding of the terms 'race' and 'ethnicity') I hesitate to say that black and white Americans are wholly different ethnicities, given that so many other 'ethnic' qualities are shared<sup>1</sup>. To do so, I believe, would indicate my endorsement of a binary distinction between citizens of a single country based only on their skin colour. Therefore, the use of this word occurs primarily as a reflection of the data, as problematic as that might be.

### **1.3.6 Social class**

When I discuss social class, I am very broadly referring to a person or group's relative access to the (unequal) distribution of resources in society, normally measured on financial capital, but also in reference to social capital (e.g. resources based on group membership and networks) and symbolic capital (e.g. legitimation based on prestige,

---

<sup>1</sup> The term 'African-American' can denote cultural context extending beyond skin colour, but this term is not always (or predominantly) used with sensitivity to the actual origin of the groups referenced; instead, it is a marker of political correctness or cultural sensitivity (Neal 2001).

honour, etc.) (Bourdieu, 1989). Again taking a cue from the data, most of the emphasis in this thesis will be on description of social class as economically determined (Skeggs, 2004, p. 31) and very coarse-grained, normally lower-class/poor, middle-class/working-class, and upper-class/wealthy.

### **1.3.7 Sexuality**

I use the term sexuality to describe “who and what one desires sexually” (Evans, in Eadie, 2004, p. 207); this is one term that has “achieved widespread acceptance...for instance in anti-discrimination legislation” (Evans, in Eadie, 2004, p. 207) and appears to have been adopted as the term of choice in the most recent portion of my AIDS/HIV corpus. Other terms appearing in the corpus (often in the 1980s and 1990s) to refer to the same concept include ‘sexual orientation’ and ‘sexual preference’, though the latter has “implied favoured choice” and has “more transitory connotations” (Evans, in Eadie, 2004, p. 208). This term is also preferable as it can encompass sexual desires held but never acted upon, in contrast to ‘sexual practices’, a term which occurs in my AIDS/HIV corpus. Sexual practices refer specifically “to the kind of sexual acts people practice and as such they can provide an important basis for identities”, e.g. the category of ‘men who have sex with men’ “was developed in the context of HIV prevention in order to capture men who hold on to their heterosexual identities but have sexual encounters with other men” (Evans, in Eadie, 2004, p. 2009).

## **1.4 Hurricane Katrina**

The analyses in Chapters 4, 5, and 6 are based upon data from a corpus of newspaper reporting on the topic of (or containing references to) Hurricane Katrina. In the sections below, I summarise the natural disaster (1.4.1), then provide a social profile of New Orleans (the most dramatically affected and highly reported upon metropolitan area) before and after the storm (1.4.2). I then discuss the role of the media in shaping interpretation of the event (1.4.3) before justifying a critical analysis of the storm (1.4.4).

### **1.4.1 The ‘catastrophic’ event**

On August 24, 2005, a tropical cyclone over the central Bahamas developed a deep inner core and, in moving toward the south-eastern coast of Florida, it strengthened until “the cyclone became Katrina, the 11th tropical storm of the 2005 Atlantic hurricane season” (Knabb, Rhome, & Brown, 2005, p. 2). Katrina made its first landfall on August 25 in Miami-Dade County as a Category 1 hurricane, with maximum sustained winds of 70 knots (Knabb et al., 2005). The storm then moved across the Gulf of Mexico,

strengthening over the water to a Category 3 hurricane with 100 knot winds by August 27, and doubling in intensity to a Category 5 storm with 150 knot winds as it neared the mouth of the Mississippi River on August 28 (Knabb et al., 2005, p. 3). On August 28, tropical storm-force winds extended 200 miles from the storm's centre; 90 miles of these were hurricane-force winds (Knabb et al., 2005). Due to internal deterioration of the eyewall, Katrina was a Category 3 hurricane during its final two landfalls in Louisiana on August 29. As it moved north in the days that followed, the storm rapidly weakened, and was downgraded to tropical depression on August 30 before being absorbed into a frontal zone in the Great Lakes on August 31 (Knabb et al., 2005).

Much of the Mississippi River settlements are now below sea level, exposing residents to floods and surges. New Orleans itself is below sea level in a 'bowl' surrounded by a protective ring of dams and levees meant to guard against surges from the Mississippi River to the south and the west and Lake Pontchartrain to the north. During Hurricane Katrina, storm surges of nearly 30 feet were observed in areas of Mississippi. Surges of 5-19 feet around New Orleans strained the water protection systems and led to multiple cases of overtopping/breaching of floodwalls and levees, many of which failed catastrophically during or immediately following landfall (Knabb et al., 2005, p. 9). "Overall, about 80% of the city of New Orleans flooded, to varying depths up to about 20 ft. within a day or so after landfall of the eye" and storm waters remained in the city for 43 days after touchdown (Knabb et al., 2005). The storm destroyed homes and businesses across the entire northern Gulf coast, and disrupted a significant percentage of American oil refinement capacity while also causing millions of gallons of oil to be spilled into the Louisiana coastline (Knabb et al., 2005, p. 12). Overall, \$81 billion of damages are attributed to Hurricane Katrina, only half of which were insured. This makes Hurricane Katrina the most costly natural disaster in the history of the United States (Knabb et al., 2005, p. 1) to date. It was also one of the most deadly. Direct and indirect fatalities attributed to Katrina are numbered at 1,833 (Knabb et al., 2005, p. 11) with 1,464 of these from Louisiana (Louisiana Department of Health and Hospitals, 2006). As a large percentage of the storm damages (and therefore the reporting) centred on the city of New Orleans, this will serve as the locus of my research.

#### ***1.4.2 Social Profile: New Orleans before and after Katrina***

Before going on, it is important to note that New Orleans is not representative of demographic averages across America. The city has greater populations both of non-white inhabitants and of residents living beneath the poverty level than the national

statistical averages (Bishaw & Iceland, 2003; Grieco & Cassidy, 2001; U.S. Census Bureau, 2004; U.S. Census Bureau, 2004b). According to the 2000 United States census, 75.1% of Americans were white, 12.5% were Hispanic, and 12.3% were black or African-American, with all other races accounting for the remainder of the 281,421,906-person population (Grieco & Cassidy, 2001). Nationwide, approximately 12.4% of individuals in 2000 were considered "below poverty level" (Bishaw & Iceland, 2003). By contrast, a 2004 United States census of the city of New Orleans found 68.3% of residents to be black or African-American, with 23.2% of all residents living below the poverty line (U.S. Census Bureau, 2004; U.S. Census Bureau, 2004b). New Orleans also has significantly higher levels of crime compared to national averages; even in the years before Hurricane Katrina, it was often referred to as one of the nation's most dangerous cities. "The murder rate for other US cities of similar size was 13.2 per 100000 person-years during 2004; the New Orleans rate of 57.1 was more than 4 times as high as the national average for comparable cities" (VanLandingham, 2007, pp. 1614–1615). These demographic details are important as we begin to deconstruct the rhetoric of a 'poor and black New Orleans' in juxtaposition to the assumed 'white and middle class' readership of major national newspapers.

Though natural disasters can affect almost any geographical area or social group, even these 'acts of God' tend to have the greatest impact on those living in poverty or other socially disadvantaged positions (Prilleltensky, 2003). After Hurricane Katrina, "although the coincidence of race and class was starkly evident, it was poverty which primarily determined who lived in the most vulnerable, low-lying neighbourhoods (that flooded first and emptied last), who was uninsured, who was unable to escape the storm and flood (and thus who lived and who died), who had fewer choices in relocating and who did not have the resources to return and rebuild" (Voorhees, Vick, & Perkins, 2007, p. 417). Affluent white neighbourhoods in New Orleans and the surrounding parishes tended to be on higher land, whereas poorer black neighbourhoods were in lower-lying areas closer to the levees. Flood risk continues to be unevenly distributed; "[t]he poor are exposed to the greatest danger...they are relegated to the most flood-prone land, and can afford little of the flood-control protection which the rich may enjoy. Relief and rehabilitation programmes may actually encourage poor people to remain in flood endangered areas" (Wijkman & Timberlake, 1984, p. 7). Federal Emergency Management Agency (FEMA) storm damage data shows that "damaged areas were 45.8% black, compared to 26.4% in undamaged areas", that "45.7% of homes in damaged areas were occupied by renters, compared to 30.9% in undamaged

communities”, and that “20.9% of households had incomes below the poverty line in damaged areas, compared to 15.3% in undamaged areas” (Logan, 2006, p. 5). Of those Hurricane Katrina fatalities identified at St. Gabriel and Carville morgues in Louisiana, 53% were African-American, 39% Caucasian, 2% Hispanic, <1% Asian/Pacific, <1% Native American, <1% Other, with 5% Unknown (Louisiana Department of Health and Hospitals, 2006).

The government’s (delayed) call for evacuation further emphasized an existing socio-geographical divide: whereas middle- or upper-class residents were able to take swift measures to board up their homes and relocate to distant hotels or houses of family members and friends, those in the lowest social classes and elderly residents had fewer relocation options and extremely limited funds with which to realize them. Nearly all of New Orleans’s 455,046 pre-Katrina residents (U.S. Census Bureau, 2009) were evacuated (either voluntarily or forcibly) two weeks after the storm, leaving a skeleton population of no more than two thousand within city limits until flood water had been pumped out through September and October 2005 and residents were slowly allowed to return (Fussell, Sastry, & VanLandingham, 2010). By mid-2007, the population stood at only 63% of its pre-storm number (U.S. Census Bureau, 2009), with little way of knowing how many of these were original residents. At the most conservative estimate, one-third of the population of New Orleans has been permanently displaced—most of these poor, uneducated, and/or black residents whose rented properties were most affected by flood damages—understandably reluctant to return to “a post-disaster environment with poorly functioning schools, hospitals, clinics, public services, and infrastructure, as well as scarce and costly rental housing and elevated crime rates” (Fussell et al., 2010, p. 4). Other social effects of Hurricane Katrina are also pervasive and enduring. The 2006 New Orleans murder rate of 96.6 per 100,000 person-years was over 11 times the national average, and represents a 69% increase over 2004 city-specific figures and a 48% increase over 2005 city-specific figures (VanLandingham, 2007).

### **1.4.3 The role of the media**

In the days and weeks following the storm, reports of widespread destruction, human loss, and residential flooding dominated the news media. Citizens depend most deeply on the media when other avenues of communication fail, such as during catastrophic crises (Livingstone, 1998). Without electricity, cellular service, or even a clear plan of action from government emergency services, the citizens of the affected Gulf Coast

region—and those who cared for them—turned to the news for information on casualties, evacuation advice, and the location of aid organisations and resources. What was presented instead was a bleak image of New Orleans as a drowned ghetto, “a state of anarchy; anomie; chaos; disorganization; regression to animal-like behaviour; and a total collapse of social control, agencies, and personnel” (Rodriguez, 2006, p. 83).

This style of disaster reporting echoed a generalized panic and anxiety which will be discussed in the next chapter more deeply. But it stands stating here that:

“[t]he research shows that, with political stakes so high, major disasters and impending threats can fuel elite panic on both local and national levels. Such panic takes a variety of forms, including: pathological fear of social disorder and of segments of the population that are not part of the elite; practices designed to protect private property and other elite prerogatives; and postevent efforts to identify and punish scapegoats and hastily usher in new ‘reforms’”. (Kathleen Tierney, 2008, p. 131)

Media discourse reproducing elite panic and reiterating this discourse had the side effect of widely distorting the circumstances of the post-storm south. For instance, “[w]hile some antisocial behaviour did occur, the overwhelming majority of the emergent activity [following Hurricane Katrina] was prosocial in nature” (Rodriguez, 2006, p. 84), contrary to widespread reports of looting, violence, and so on. In fact, “elite panic was shockingly evident during Katrina, as evidenced by media and public officials’ obsessions with looting and lawlessness, the issuing of shoot-to-kill orders arising primarily out of a concern with property crime, and the rush to act upon rumors that circulated regarding the ‘savage’ behavior of lower-class community residents, immigrants, and people of color” (K. Tierney, Bevc, & Kuligowski, 2006, cited in Tierney, 2008, p. 131).

In this study, I cast a critical eye toward the media and also to its disparagers. Though traces of ‘elite panic’ seem to present themselves at the folk level, it remains to be proven whether these emerge as a result of an empirical, quantitative study.

#### **1.4.4 Towards a critical analysis of Hurricane Katrina**

While a natural disaster might appear better fodder for a meteorological study than a linguistic one, “Katrina can be best understood as a collision between a natural force (itself of human construction to the extent that global warming or something of the sort can be said to have been involved) and what turned out to be a strangely vulnerable social order” (Brunsma, Overfelt, & Picou, 2007, p. xx). New Orleans, the city most

affected by this storm, was vulnerable due to geographical positioning, a governmental neglect of decaying infrastructure, and the socioeconomic makeup of its citizenship. The event has tremendously changed the entire racial, economic, and social makeup of the lower Mississippi River region. The influence of the media in shaping the discourse that followed the storm, in combination with the academic discretization of many of the reported 'facts' positions this event perfectly for a critical analysis of the construal of social actors.

Though this event came later (in 2005) than the start of the AIDS/HIV outbreak (beginning in the 1980s), Hurricane Katrina has been deemed a more appropriate starting point on the basis of two factors: 1) the corpus is smaller, allowing for more methodological experimentation on a smaller data set; and 2) reporting on the event follows a very clear (reverse-) distribution reflecting a single spike in interest, compared to various distributional anomalies in the AIDS/HIV corpus. In some ways, the Katrina corpus can be considered a 'control' against which discursive and methodological findings from the AIDS/HIV corpus will be compared.

## **1.5 HIV/AIDS**

In this section, I will provide a very brief overview of the epidemic (section 1.5.1) before discussing the attendant social impact (1.5.2) and social stigma (1.5.3) of AIDS/HIV in America.

### **1.5.1 A brief overview of HIV**

In 1979 and 1980, American physicians began to note an increased incidence of "a mononucleosis-like syndrome, marked by hectic fever, weight loss, and swollen lymph nodes" (Grmek, Maulitz, & Duffin, 1990, p. 3). The first patients shared a common demographic: they were all young men from the areas surrounding New York, Los Angeles, and San Francisco, and they all stated a sexual preference for other men. This mystery virus quickly progressed to include the skin cancer Kaposi's sarcoma, as well as oral and anal thrush and respiratory infections such as *pneumocystosis carinii* pneumonia—opportunistic cancers and diseases indicative of a compromised immune system. On July 3, 1981, the medical correspondent for the *New York Times* published an exposé on the rise of a new "gay cancer" (Grmek et al., 1990, p. 10). The following day, the Centers for Disease Control and Prevention (CDC) released an official report on clinical findings of the outbreak, and later that month, a major team of medical investigators published clinical findings concluding that the "unusual

occurrence of Kaposi's sarcoma in a population much exposed to sexually transmissible disease suggests that such exposure may play a role in its pathogenesis" (Hymes et al., 1981, p. 574).

Within months, several dozen similar cases were reported both domestically and internationally. While early reports (drawing upon a small population of gay men) indicated that 'popper' (alkyl nitrates) use or sexual promiscuity could lead one to 'catch' AIDS, it soon became clear that AIDS is caused by the lentivirus HIV, and HIV is transmitted through the exchange of bodily fluids (e.g. semen and blood) (Conrad, 1986). While all humans are at risk of contracting the virus, those who participate regularly in activities involving the exchange of bodily fluids (e.g., intravenous drug users, haemophiliacs, prostitutes) are considered to constitute a 'heightened risk group'.

Uncertainty regarding both the source and the pathogenesis of the disease beset researchers through the 1980s, and paucity of information continues to this day. It is now commonly agreed that HIV descended from an African strain of Simian Immunodeficiency Virus (SIV), though the process by which SIV (which affects monkeys) 'jumped' species to HIV (which affects humans) is still contested. Further, an 'AIDS Test' still does not exist; the syndrome is characterized by the positive presence of HIV, and "by more than two dozen different illnesses and symptoms as well as by specific indications on blood test findings" (Smith, 1998, p. 64). HIV is said to have progressed to AIDS in the presence of a so-called "AIDS-defining illness", or when a blood test shows T-cells expressing the immunoglobulin glycoprotein CD4 below a certain level, indicating severe immunodeficiency (Smith, 1998, p. 64).

Though American transmission rates have dropped steadily since the mid-1980s (Center for Disease Control, 2011), the epidemic is still spreading. The CDC estimates that over 1.1 million people are currently living with HIV in the US, with an additional 50,000 Americans becoming infected each year (Center for Disease Control, 2011). Since the start of the American epidemic, a set number of minority social groups have borne the brunt of both infection and mortality rates. In 2008, over half of an estimated 56,300 Americans newly infected with HIV were gay and bisexual men; in the same year, it was estimated that Americans of black/African ethnicity were seven times more likely to contract the virus as white Americans (H. I. Hall et al., 2008). Overall, "HIV continues to take a severe toll on multiple communities in the US, with gay and bisexual men of all races, African Americans, and Latinos bearing the heaviest burden" (Center for Disease



Control, 2008, p. 3). It is for this reason that a diachronic critical analysis of the social impact of the naming strategies of the epidemic is proposed.

### **1.5.2 Social impact**

This initial link between the spread of the epidemic and homosexuality as causative practice has had a deep impact that continues to resound today. While the 1960s and 1970s were generally seen as a progressive era for gay rights—many ‘sodomy’ laws were repealed and communities were formed around ideals of acceptance and experimentation—the initial isolated appearance of HIV within the community sharply curtailed burgeoning social support and even reignited moral fear of gay people.

The earliest official designation of this new and terrifying disease was Gay-Related Immune Deficiency Syndrome (GRID), though it was more popularly known as “gay plague” during the early 1980s (Conrad, 1986). As with many diseases, framing has historically relied upon linguistic techniques such as nomination and metaphor (Sontag, 1989). It was not until September 1982—after People with AIDS (PWAs) were demonstrated to include haemophiliacs and Haitians—that indication of sexuality was dropped from the nomenclature, and the term *Acquired Immunodeficiency Syndrome* (AIDS) was officially adopted by the CDC (Center for Disease Control, 1982). But even this—a generalized name that did not include definitions of carrier demographics or details about methods of ‘acquiring’ the syndrome—was soon officially narrowed to include only minority groups.

On March 3, 1983, the CDC pronounced that it had identified four groups at heightened risk of contracting AIDS: 1) homosexual men who engage in unprotected sexual acts with multiple partners; 2) intravenous heroin users; 3) Haitian immigrants; and 4) haemophiliacs (Siplon, 2002, p. 6). While this was intended as the epidemiologists’ empirical description of ‘at-risk’ individuals, it was not long before “all members of the identified risk groups were seen as potentially contagious; from there it was but a short step to perceive those populations as ‘responsible’ for AIDS” (Fee & Krieger, 1993, p. 324). The public policy debates that followed publication of this (rather explosive) information were particularly heated, in part due to the extraordinary friction between policy makers attempting to halt the spread of an epidemic and affected social groups struggling to maintain their values and lifestyles. For instance, “[i]t was hard for many public health officials, who had always dealt with infectious diseases by finding infected individuals and trying to limit their opportunities for infecting others, to understand elements of the gay community, who sought to preserve their sexual freedom and thus

fought hard to keep open bath houses and other places that facilitated sexual encounters” (Fee & Krieger, 1993, pp. 327–328).

Traditional methods of infectious disease control have largely employed “a combination of coercive measures and moralistic messages” (Fee & Krieger, 1993, p. 328). This governmental response often works in tandem with popular social reaction; traditionally, if a “disease occurred mainly within marginalized groups or was believed to spread through morally suspect behaviour, society blamed the victim for becoming infected or transmitting the disease to others” (Fee & Krieger, 1993, p. 328). Initially, both governmental and societal responses were similar to outbreak patterns in previous epidemics. Stories of HIV+ children (or even their HIV- siblings) being withheld from schools, medical workers refusing treatment of PWAs, and religious leaders denouncing infected clergy members inundated the media (Conrad, 1986, p. 76). Thus, “the AIDS epidemic—with its genuine potential for global devastation—is simultaneously an epidemic of a transmissible lethal disease and an epidemic of meanings or signification” (Treichler, 1987, p. 295). The spread of negative sentiment was multiplied by the doubly stigmatizing nature of HIV: PWAs likely belong to some marginalized social group (gay men, drug users, sex workers), the perceived deviance of which is only amplified by the contraction of a chronic and fatal illness. This had a large social impact: “[w]hile AIDS is contagious, so is the fear and stigma” (Conrad, 1986, p. 79).

### **1.5.3 Social stigma**

It was not until the late 1980s and especially the 1990s that a rights-based counter-narrative arose from newly formed international AIDS/HIV organizations and the enormous mobilization efforts of PWA support groups. But, despite the best efforts of organized human rights organizations and special interest groups, AIDS-related social panic continued well through the 1980s, 1990s, and even 2000s. In 1996-1997 and 1998-1999, Herek and a team of researchers (2002) undertook a series of random telephone surveys from a random-digit-dialling sample of all American households with telephones in the contiguous 48 states, and compared results with a previous 1990-1991 dataset. The study aimed to measure HIV-related stigma and knowledge in the United States over the course of the decade.

Though some findings were somewhat heartening—e.g. support for coercive policies such as forced quarantine or public naming of PWAs declined steeply over the decade—many of the results ran counter to my initial belief that HIV-related stigma would weaken steadily from the ‘dark ages’ of AIDS-as-plague into the current era of cocktail

therapies and AIDS-as-chronic-disease. In 1991, 27.7% of respondents admitted to being angry at PWAs; 34.6% were afraid; 26.6% were disgusted. All indicators declined in 1997, and in 1999, 14.8% claimed to be angry; 20.2% were afraid, and 16.0% were disgusted (Herek, Capitanio, & Widaman, 2002, p. 372). Though all figures represent a precipitous drop, it is noteworthy that even 20 years after the arrival of AIDS in the public eye, such a large percentage of those polled still felt anger, disgust, and especially fear. Other indicators are bleaker still. "The proportion of respondents believing that 'people who got AIDS through sex or drug use have gotten what they deserve' peaked in 1997 at 28%", a significant increase over the 1991 figure (Herek et al., 2002, p. 374). Agreement declined in 1999, but approximately one fourth of survey respondents still expressed endorsement (Herek et al., 2002, p. 374). "This pattern is worrisome because individuals with an undesirable condition are generally subjected to greater stigma when they are perceived to be personally responsible for their situation" (Herek et al., 2002, p. 375). For many survey participants, PWAs were also held directly accountable for the spread of the epidemic, with 25.5% in 1997 and 21.8% in 1999 believing that "most people with AIDS don't care if they infect others with the AIDS virus" (Herek et al., 2002, p. 375). From 1997-1999, at least 20% of respondents (and up to 30%) additionally stated that they would avoid shopping at a grocery store owned by a PWA, avoid wearing a sweater or drinking from a glass once used by a PWA, and feel discomfort if a PWA were attending their child's school, working in their office, or working in their supermarket. The researchers conclude that "[a]lthough support for extremely punitive policies toward PWAs has declined, AIDS remains a stigmatized condition in the United States" (Herek et al., 2002, p. 371). "The persistence of discomfort with PWAs, blame directed at PWAs for their condition, and misapprehensions about casual social contact are cause for continuing concern" and represent a major hurdle in HIV education programs and in the ongoing battle against the spread of AIDS (Herek et al., 2002, p. 371).

A slight drop in the stigma associated with HIV might correlate to the shifting 'face' of those who were infected. "[A]lthough many have perceived the epidemic as affecting mostly White gay men" (Sutton et al., 2009, p. 351), there is little question that so-called 'risk groups' expanded quickly and rapidly outward from this demographic through the 1980s. But nearly as soon as the CDC declared 'risk groups', the targets had shifted. Technological advances in medical testing in the 1990s all but eliminated the risk of contraction through blood transfusion in America. Social advances like 'risk reduction' policies on clean syringe and needle exchange programmes have had broad, positive

effects on limiting the spread of HIV in drug-using communities (Hurley, Jolley, & Kaldor, 1997). These descriptions do downplay the major struggles undertaken by proponents of the distribution of clean needles (and free condoms), who still battle opposition from those who believe that such measures encourage illicit activity. However, the single greatest risk of contracting HIV is still one that most adults on the planet engage in: sexual transmission (of both homosexual and heterosexual natures) accounted for 90% of new HIV cases reported in 2006 (Sutton et al., 2009, p. 351).

Black people and African-Americans have also been disproportionately affected by the disease: between 1981 and 1988 they represented just 12% of the American population, but accounted for 26% of AIDS cases (Sutton et al., 2009). This trend has only become more pronounced in recent times. In 2006, black people still roughly represented 12% of the population, though they now “account for almost half (46%) of people living with HIV in the US, as well as nearly half (45%) of new infections each year” (Center for Disease Control, 2008, p. 2). This disproportionate distribution of ‘bads’ (or social risks) echoes the social context of Hurricane Katrina reporting, one aspect leading me to link the two events together in a single mirrored analytical series.

## **1.6 Rationale for study**

Previous studies have focused on the creation of US/THEM discursive patterns following events wherein a degree of social disruption can be attributed (broadly and fallaciously) to human individuals or groups. This study undertakes linguistic examination of news reporting on socially disruptive events that strike somewhat indeterminately, and which cannot be so readily attributed to specific human agents.

To explore the discourses of risk and blame in ‘act of God’ reporting, I have selected two relatively modern, socially disruptive circumstances: the Hurricane Katrina natural disaster (2005) and the AIDS epidemic (1980-2009). I refer to these two loosely as ‘acts of God’, a legal term meaning “an extraordinary interruption by a natural cause (as a flood or earthquake) of the usual course of events that experience, prescience, or care cannot reasonably foresee or prevent” (“Merriam Webster Dictionary Online,” 2013). While this term is used most often in the context of natural disasters (and is therefore more suitable to Hurricane Katrina), I believe that it also applies to AIDS/HIV in this context for the following reasons: 1) the epidemic caused extraordinary interruption; 2) the cause (the lentivirus itself) is natural; 3) at the time of its outbreak, experience, prescience, or care could not have foreseen or prevented its occurrence or spread through unaware ‘risk groups’; 4) as is the case with victims of natural disasters, victims

of AIDS/HIV are often precluded as insurance holders from receiving full [medical, housing, employment] benefits, meeting the legal criteria. Finally, 'act of God' conveys an additional meaning not embedded in 'epidemic' or 'natural disaster'; while both events were superficially outside of human control, the effects were magnified due to human action (e.g. climate change, intravenous drug abuse), and detriments are sometimes reported as being holy retribution for immoralities. Though the term 'act of God' might be imperfect (and could be falsely misinterpreted as my endorsing the concept of mortal reprisal), it is the most apt that I have been able to find in these three years of research.

As is the case with most natural disasters, Hurricane Katrina unfolded in a short time frame, but over the course of a week, killed nearly 1,833 people and caused an estimated US\$81 billion of damage (see page 8). On the other hand, HIV/AIDS was first reported in the media over 30 years ago, and has been responsible for approximately 500,000 deaths in the United States alone and 25 million deaths worldwide. Therefore, Hurricane Katrina lends itself to a 'snapshot' study of reporting practices immediately following the event, whereas HIV/AIDS allows for a diachronic study of media discourse over three decades. To normalise for a shared cultural context (with which I, as the researcher, am familiar), and to narrow the scope of research, these circumstances will be explored only through American media discourse.

To expand existing research in discriminatory 'othering' in media discourse, a new series of phenomena will be defined and presented. In an extension of documented US vs THEM schema, a new theory called Social Sequestering will be developed. By tracing, analysing, and cross-comparing large samples of discourse, I aim to illustrate that in the face of destabilising events, the media collectively sequesters minority populations into 'risk groups', perhaps with the aim of protecting the majority group from social panic—a combination of risk society (Beck, 1992) and moral panic (Cohen, 1972). Over time, the parameters for inclusion in risk groups may change, though pilot studies indicate that all risk groups belong to pre-existing categories of 'deviants' or 'folk devils' (Cohen, 1972), and that assignment of one sort of deviancy to social actors is often followed by compounded assignment of other out-group characteristics.

This study is designed to discover whether 'othering' can be shown to exist in American 'act of God' reporting, even when the 'other' must be fabricated. This is the first large-scale, diachronic, corpus-assisted, study of the representation of social actors in AIDS/HIV reporting available. Comparison to a large corpus of Hurricane Katrina

reporting allows me to develop a robust method of downsampling that I recommend to other researchers interested in the representation of identity in large data sets.

## **1.7 Structure of thesis**

Following this introductory chapter, I provide an overview of key theories guiding my work in Chapter 2, highlighting previous scholarly contributions of particular relevance. Here I will explain and justify the incorporation of a mixed methods approach of corpus-based critical discourse analysis, citing successes in similar work. In Chapter 3, I outline the method of collecting my data and describe the corpora. It is in this chapter that I detail and justify broad methodological choices, describe techniques incorporated from the fields of both Corpus Linguistics and Critical Discourse Analysis, and justify methods undertaken in the chapters to follow. In Chapters 4 and 7, semantic collocation of the most frequent naming strategy in the Katrina corpus (Chapter 4) and AIDS/HIV corpus (Chapter 7) is undertaken to establish a discursive baseline of identity construal in the corpora. In Chapters 5 and 8, four additional frequent naming strategies from each corpus (Katrina in Chapter 5, AIDS/HIV in Chapter 8) are compared on the basis of proportional semantic collocation. It is by this method that the covert meanings of such naming strategies can be somewhat demystified. The final analysis chapters for each corpus are both diachronic studies. In Chapter 6, I study the changing use of naming strategy over one year in the Katrina corpus while in Chapter 9, I examine the evolving construal of an identity group over nearly three decades. The final Chapter, 10, concludes the study and offers a critical summary of findings as well as some reflections on the development of a method of semantic collocation. Here, I will offer some final remarks on the perceived efficacy of this method to downsample large (and somewhat dissimilar) corpora.

## **1.8 Summary**

In this chapter I have provided an introduction to my research topic, an orientation to my position as a researcher, and detailed research questions to be explored in the following sections. Key terms have been outlined and defined, and the two events forming the basis of my analyses have been discussed. Following the rationale for my study and an overview of the structure of the thesis, I now turn to an overview of key theoretical and literary contributions influencing this work.

## **Chapter 2: Overview of theory and literature**

In this chapter, I provide an overview of the theory and literature underpinning my research. In the following sections I will: discuss the motivating factors in selection of media discourse as primary data source (section 2.1); provide brief overviews of various forms of critical discourse analysis with an emphasis on research in the media (section 2.2); discuss the incorporation of methods in corpus analysis as supporting critical studies (section 2.3); and finally, identify some of the research gaps that I attempt to address in this thesis (section 2.4).

The methods needed to undertake research of this description must support: 1) theorization of socio-cultural hierarchies, 2) problematization of accepted values, and 3) empirical representation of substantiating evidence. To meet all of these aims, media discourse will be considered within a (largely) theoretical framework—Critical Discourse Analysis—in combination with a (largely) methodological framework—Corpus Linguistics. Wherever possible, reviews of relevant literature will be incorporated to illustrate points and to position this research within the context of previous studies.

### **2.1 Media Discourse**

A 2009 report from analytical firm Scarborough Research “finds nearly three in four adults, nearly 171 million, in the U.S. read a newspaper – in print or online – on a weekly basis” (2009, p. 1). Given the implications of such pervasive ideological impact upon the public, it is natural that researchers seek to gain a deeper understanding of the mediated engagement between the Press and the people (Talbot, 2007, p. 3).

Emerging technologies have contributed to a slow decline in print media consumption; however, this trend has also had an intensifying and diversifying effect on makeup of the ‘average’ news audience: more Americans than ever before now access the news on a regular basis (Scarborough Research, 2009). While still older, more educated, and more affluent than the American median measures, readers represent an ever-diversifying picture of the public. The modern media—an institutionalized purveyor of discourse—is tasked with homogenizing the expectations of their enormously diverse audience and providing (at the very least, an illusion of-) objective, immediate information, in a principally linguistic form. “Language analysis, then, can help anchor social and cultural research and analysis in a detailed understanding of the nature of media output” (Fairclough, 1995, p. 16).

This thesis takes Halliday's (1985) position: language is a tool by which people process, organize, and communicate experience. Therefore, "a useful working assumption is that any part of any text (from the media or from elsewhere) will be simultaneously representing, setting up identities, and setting up relations." (Fairclough, 1995, p. 5). The news, then, is not a list of 'facts' updated and distributed at regular intervals; it is the communication of stories which have been selected for their supposed social interest, processed by powerful peoples and institutions, and generated with a pre-imbued set of existing moral, social, and ideological values (S. Hall, 1997).

Perhaps owing to this complex interaction between socio-political context and linguistic form and function, analysis of media discourse informs a multitude of fields: "[i]n addition to extensive interest in media and cultural studies, it is the subject of scrutiny in linguistics—particularly conversation analysis, critical discourse analysis, ethnography of communication, linguistic anthropology, pragmatics and sociolinguistics—and also in cultural geography, psychology, sociology and tourism studies." (Talbot, 2007, p. 3). Contemporary linguistic approaches to media discourse have included: linguistic and sociolinguistic analysis of the 'form' of texts (Bell, 1991; Mårdh, 1980; Ungerer, 2000); conversation analysis of institutionalized and/or media discourse (Atkinson & Heritage, 1984; Greatbatch, 1986; Heritage, 1985; Hutchby, 1996); critical linguistics (Fowler, Hodge, Kress, & Trew, 1979; Fowler, 1991); semiotic analysis of linguistic codes and conventions (Fiske & Hartley, 1978); social semiotic investigations of representation and choice (Hodge & Kress, 1988); and socio-cognitive perspectives placing equal emphases on discursive production and reception (van Dijk, 1988); among many others<sup>2</sup>. My own research follows primarily upon established work in two of these areas: critical analysis and socio-cognitive perspectives of discourse. In the following sections, I will discuss the sociological implications of news selection, presentation, and reception, drawing upon examples from previous studies and alluding to the research ultimately contained within this thesis, wherever possible.

### **2.1.1 The Socialized Selection of News**

Reporting every event transpiring in every corner of the globe would require an infinite number of hours in any news cycle. In order to maintain high circulation numbers and maximize income, media organizations must work quickly and selectively—constantly generating content on matters they believe to be most relevant to their customers

---

<sup>2</sup> See Fairclough's *Media Discourse* (1995), van Dijk's *Discourse and Power* (2008), or *Approaches to Media Discourse* (1998) edited by Bell and Garrett for extended reviews of research in media discourse.



(Galtung & Ruge, 1973) , and taking pains to present it in a manner more enticing than that of their competitors (White, 1997). Therefore, the generation and presentation of news can be described as a culturally determined social practice negotiated between readers and journalists (Galtung & Ruge, 1973). The selection of news by publishing companies is governed by two broad principles: relevancy/news values (Galtung & Ruge, 1965, 1973) and ease of access (Fowler, 1991).

Galtung and Ruge (1965, pp. 53–55) claim that the media tend to select reporting subjects on the basis of eight criteria of newsworthiness (news values), or relevancy:

1. **Frequency:** Stories occurring over a relatively short life span (or in a long lifespan punctuated by a short-lived climax) have a great probability of being covered in any given news cycle.
2. **Amplitude:** The more dramatic the event, the more newsworthy.
3. **Clarity:** Newsworthy events contain clear delineations between right and wrong, and are free from ambiguity of meaning.
4. **Meaning:** Events occurring in cultural similar context (or to people in a shared culture or like culture) are more familiar and therefore more interesting to readers.
5. **Consonance:** Events corresponding with pre-established predictions or expectations are better received and registered by readers.
6. **Unexpectedness:** Events that are both meaningful and consonant can also be unexpected (or rare).
7. **Continuity:** Once an event is deemed newsworthy, it continues within the news cycle as the previously unexpected becomes the commonplace.
8. **Composition:** When lacking stories of a certain composition (i.e. international, domestic), the threshold of newsworthiness for this composition goes down, and vice versa.

In summary, newspapers tailor meanings created in their reporting to remain compatible with the ideologies of their readers and individual political standpoint, but a variety of cultural factors come into play alongside these economical-political features (Bell, 1991). This concept of media output as a socialized, selective expression of culture is what leads van Dijk (1991, p. 2) to assert that “news is a *practice*: a discourse which, far from neutrally reflecting social reality and empirical facts, intervenes in what Berger and Luckmann call ‘the social construction of reality.’” In other words: all stories are not created equal. Events garnering the most coverage tend to be those with higher news

values, which concern elite nations or people, can be accessed from an interpersonal level, or are particularly negative (Galtung & Ruge, 1965).

American reports on domestic events are, therefore, both rich in *meaning* (sharing a common cultural context with readers) and also *elitism* (concerning powerful nations). Reports of natural disaster are additionally marked with extreme *frequency, amplitude, unexpectedness, and continuity*; journalistic accounts of the spread and impact of an epidemic, while lacking dramatic peaks in *frequency* or *amplitude*, do rate highly in *consonance* and *continuity* (in addition to being interpersonally loaded and deeply negative). Ostensibly, both of these topics would additionally appear to have a firm aspect of moral *clarity*; the materialization of a sudden, catastrophic threat to life and livelihood is undeniably 'bad' in any culture. However, reports from the disaster/epidemic genres lack the ability to assign straightforward fault; the natural world is not a suitable scapegoat. The continued (and covert) demarcation of good versus bad nonetheless existent in these reports can perhaps provide an insight into the underlying ideology of media sources.

News isn't 'delivered' but rather 'gathered' (Fowler, 1991). Journalists have developed a strategy to conserve time and effort: the close monitoring of institutions offering consistently 'newsy' output—including political/judicial departments, public service organizations, corporations and unions, emergency/military services, and prominent individuals (Fowler, 1991, pp. 20–21). These sources are:

“...highly privileged: they are *established* by official authority, by social status or by commercial success; they are *organized*, with a bureaucratic structure which embodies spokespersons, and a regular scheduling of statements; and they have the *resources* to pay for publicity and public relations. Thus, the most convenient sources for journalists to monitor are also, necessarily, institutions and persons with official authority and/or financial power.” (Fowler, 1991, p. 22)

Powerful persons or institutions have primary access to information about events in the world, and the press has the means to interpret and propagate this information to billions of news consumers. This reciprocal relationship between the powerful and the press has heavy influence over the selection and presentation of news, resulting in both the reflection and perpetration of existing ideologies. Linguistic forms are likewise constructions of social and cultural values; “[l]anguage analysis, then, can help anchor social and cultural research and analysis in a detailed understanding of the nature of media output” (Fairclough, 1995, p. 16).

### **2.1.2 Writers, Readers, and Cultural Consensus**

Above, it was posited that the media is in the business of gathering information from powerful sources and circulating it to the masses, for profit. The business relationships between both suppliers (contacts) and clients (readers) are most ably attended when the interests of the political powerbrokers align with those of the consumers; both the Press and the government stand to benefit from the brokering of agreement on a range of consumerist or civic interests. It is by supporting an ideology of consensus that the Press can further the institutional aims of its sources, and develop a stronger relationship with its audience (Fowler, 1991, p. 49).

Cultural consensus does not relate to common agreement on a set of facts. Instead, the ideology of consensus assumes that the interests of any given population “are undivided, held in common; and that the whole population acknowledges this ‘fact’ by subscribing to a certain set of beliefs” (Fowler, 1991, p. 49). This idea relates to the assumption that cultures have common ideological systems of values (Hartley, 1982) that are enacted through institutionalized communication and reinforced during times of crisis (Fowler 1991: 49), for instance, in the event of epidemic or natural disaster. Entire cultures may thus develop common positive values for concepts such as *order*, *co-operation*, and *responsibility*, while negatively evaluating *extremism*, *chaos*, and *violence* (Chibnall, 1977). Such differentiation between ‘right’ and ‘wrong’ allows for categorization of actions or ideas (and responsible parties) into those inside the consensus—the general population, ‘us’—and those outside of consensus—deviants, dissidents, ‘them’ (Fowler, 1991, p. 53).

This concept is echoed in van Leeuwen (2008), who defines ‘social practices’ as socially acceptable models of how activities should be accomplished in order to achieve harmony, coordination, and cooperation within society. In this framework, ‘social actors’ are the participants highlighted within the discourse in the scope of their social practices, thereby signifying to the intended reader appraisals of these social actors, their actions, and their overall place in the societal context (van Leeuwen, 2008). In this thesis, I will be attempting to reconstruct the social model that social actors were framed in by following linguistic traces of appraisal, as “[a]ll texts code the ideological position[s] of their producers” (C. Caldas-Coulthard, 1996, p. 228). One linguistic aspect of this ideological position is traces of the construction, reiteration, and problematisation of moral panics and risk society discourse.

### 2.1.3 Moral Panic and the Risk Society

Examining crossovers in discourses relating aspects of moral panic versus risk society is a particularly interesting challenge given the diachronic scope of the project. As world conditions and concerns change, so, too, do discourses employed to describe and process the human experience. Throughout the late 1960s and early 1970s, Stanley Cohen transfigured the sociological study of deviance by defining the idea of 'moral panic', or a period where a "condition, episode, person, or groups of persons emerges to become defined as a threat to societal values and interests", is vilified within the Press, moralized by powerful peoples, 'diagnosed' by experts, and subjected to coping mechanisms until which point the hysteria subsides (Cohen, 1972, p. 9). Goode and Ben Yahuda characterize moral panics by: heightened *concern* over a group's behaviour; an increase in *hostility* aimed toward this group; emergence of a *consensus* that behaviour poses a social threat; a *disproportional* response taken by other society members; and *volatile*, sudden peaks and troughs in concern or interest (2010, pp. 37–41). They note that "moral panics are likely to 'clarify [the] normative contours' and 'moral boundaries' of the society in which they take place, demonstrat[ing] that there are limits to how much diversity can be tolerated in a society" (Goode & Ben-Yehuda, 2010, p. 29).

Though Cohen's theory is now almost fifty years old, modern researchers continue to reference the phenomenon in their work. Moral panics still surround the discourses concerning subcultures observed by Cohen (i.e. ethnic/racial minorities, violent youths, sexual transgressors). Baker observed moral panics relating to sexual deviancy in UK tabloid newspapers, where homosexuality was consistently linked to "sexual aggression, deviancy and impermanence", particularly in the Daily Mail, whose "discourses of gay men as powerful, ubiquitous, secretive and proselytising suggests an underlying fear and anger – that a change is occurring which they are powerless to prevent" (2005, p. 91). This is extremely relevant to the AIDS/HIV chapters, in particular Chapter 9, where I undertake a close analysis of this identity group. Goode and Ben Yahuda (2010) note the continued presence of drug abuse panics (particularly in relation to youth panics), which has obvious links to the panics related to the spread of HIV through recreational intravenous drug use. In a large-scale corpus-based discourse analysis of Muslim people in the media, Baker, Gabrielatos and McEneaney identified this group as "the targets of a 'scrounger' moral panic conducted by the tabloids during the 2000s" (2013, p. 187). As I will demonstrate later, a similar discourse of unfair or undeserved doling out of governmental resources to people with AIDS/HIV and people affected by Katrina (even 'phony victims') arises in both corpora.

However, analysis of two happenings that extend beyond the unadulterated power of human influence (the presence of a natural disaster and the mutation of an epidemic) necessitate the inclusion of a further framework. Beck (1992) proposes that the evolving conditions of a post-modern, post-industrial world have inspired a new variety of moral panic: discourses of the risk society. Cohen (1972) conceived of moral panics relating to chiefly localized incidents—rising breast cancer rates among certain demographics, increasing instances of violence attributed to a certain group, etc., whereas Beck (1992) concludes that contemporary society is increasingly concerned with the possibility of catastrophic global events. While moral panics (and the ‘folk devils’ assigned responsibility for them) could be ‘managed’ by powerful institutions in Cohen’s decades-old description, the sheer scale of modern risk—multi-national nuclear/environmental disaster, cataclysmic collapse of the world financial system, the rapid spread of epidemics on a ‘shrinking’ planet—have resulted in a new breed of social anxiety that even the world’s most powerful must grapple to manage (Beck, 1992). “These risks have novel impacts that are: 1) very complex in terms of causation; 2) unpredictable and latent; 3) not limited by time, space, or social class (i.e., globalized); 4) not detectable by our physical senses; and 5) are the result of human decisions” (Ungar, 2001, p. 3). In other words, while ‘folk devils’ could be ‘hunted’ down and subjected to measures of social control, the risk society—lacking its former system of identifying blameworthy ‘others’—is more apt to attempt to place blame on institutions tasked with ensuring and communicating safety: for instance, the government and/or the Press (Ungar, 2001).

Controversially, Beck rejects class as an important concept in the new social reality (1992), leading to a series of extremely persuasive scholarly responses demonstrating that class continues to correspond to life-changes (Goldthorpe, 2002), and that risk invariably runs in parallel to disadvantage (Mythen, 2005a, 2005b). In addition to the simple reality that mortality rates after natural disasters are concentrated in lower socio-economic strata (Ibarrán & Ruth, 2009, p. 54), which also face much greater difficulties in recovering from said disasters (Blaikie, Cannon, Davis, & Wisner, 1994, p. 48), this thesis, particularly in Chapters 4-6, provides additional evidence that the concept of risk is construed in tandem with the concept of poverty. This corroborates findings by Cutter, Boruff and Shirley (2003), Adger (2006), and Redefining Progress (2004), who find that relative wealth determines who occupies the physical and ideological spaces most vulnerable to the continued exposure to ever-more diverse and catastrophic risks (Curran, 2013).

It is my hypothesis that this results in a reflexive move on the part of powerful bodies: namely, the *manufacture* of a deviant 'other' (who resides in the familiar subjective territory of former 'panics') to quash the prospect of widespread anxiety in the 'us' or 'have' group and dramatically decrease the possibility of mutiny in the 'them' or 'have-not' group in this emerging global society. This relationship must be managed powerfully:

As the risk society develops, so does the antagonism between those afflicted by risks and those who profit from them. The social and economic importance of knowledge grows similarly, and with it the power over the media to structure knowledge (science and research) and disseminate it (mass media). The risk society in this sense is also the science, media and information society. Thus new antagonisms grow up between those who produce risk definitions and those who consume them. (Beck, 1992, p. 46)

#### **2.1.4 The Structuring of News**

Peter White (1997) describes the generic organization of 'hard news' media texts as well as the socio-ideological objectives realized by these structures. Lead-dominated hard news stories can portray either events or issues, and are ideationally controlled by events or situations considered disruptive to social or moral order, namely: aberrant damage (natural forces, accidents, disease), power relation (politics, wars, takeovers), or normative breach (crime, corruption, negligence) (White, 1997, pp. 104–106). While journalists are inclined to avoid inscription of specific or overt judgment, linguistic choices encode meaning and flag reasons for significance or avenues for engagement—lexical intensification commonly inscribes interpersonal engagement, and comparisons imbue events with a sense of severity or significance (White, 1997, p. 109). Because hard news is concerned with the relation of disruptive events, selection and ideological structuring of news stories relies on a shared socio-cultural context defining normative society.

Anthropological linguists Edward Sapir and Benjamin Whorf and systemic functional linguist Michael Halliday were among the first to hypothesize that languages are structured to communicate ideas and experiences that are often specific to the speech communities in which they are used, and differences in these structures/grammars facilitate and perpetuate socialized ways of viewing and describing encounters with the world (Fowler, 1991, pp. 29–38). Since the 1960s, research in the field of linguistics has increasingly intersected with sociology, gender, and culture studies. It is in the modern social scientist's interest to investigate the ideological implications of language in use, particularly in the case of media discourse, where the ideologies of the powerful are

replicated and disseminated to millions of society members (Fowler, 1991). The theoretical toolkit utilized in such a study must, then, allow for equal emphasis on analysis of both discourse and society.

## **2.2 Critical approaches to discourse analysis**

In 1991, Norman Fairclough, Ruth Wodak, and Teun van Dijk, along with two other Critical Discourse Analysis (CDA) pioneers, Gunther Kress and Theo van Leeuwen (who have since moved toward other approaches), attended a small symposium in Amsterdam (Wodak & Meyer, 2009, p. 3). This significant ‘meeting of the minds’ allowed participants to compare their agendas and to discuss their perspectives. It was only after this meeting that CDA emerged as a multi-disciplinary blend, drawing from the five scholars’ composite interest in the ideological theories of Louis Althusser, philosophical viewpoints of Antonio Gramsci and Jürgen Habermas, sociological studies by Michel Foucault, and functional linguistic systems of Michael Halliday and Mikhail Bakhtin (Titscher, Meyer, Wodak, & Vetter, 2000).

These assorted areas of social science still bear upon the various threads of CDA to varying degrees; as a result, the programme remains highly socially sensitive. In fact, most critical discourse analysts understand the unifying theme of the CDA program to be a shared commitment to describing cultural context, exposing power relationships enacted through communicative practices, and acting as change agents on behalf of the powerless (Meyer, 2001, p. 15). Regardless of its particular concentrations of social scientific influence, each branch of CDA is additionally “characterized by the common interests in de-mystifying ideologies and power through the systematic and reproducible investigation of semiotic data (written, spoken or visual)” (Wodak & Meyer, 2009, p. 3) and disciplined critique of the effect of sociological ‘norms’ or values on resultant linguistic forms (Fairclough, 1995; Fowler et al., 1979). The Hallidayan view of language as a functional result of a series of choices informed by societal, cultural, and personal context is foundational in CDA (Fowler et al., 1979; Hodge & Kress, 1979), and serves as a common backbone.

Though their research converged for a period, the interests that initially led each researcher to CDA differ; thus the sorts of academic study carried on in their research legacies have continued to expand and diverge<sup>3</sup>. As a result, Critical Discourse Analysis is a label that has been applied to many schools of thought, each with their own specific

---

<sup>3</sup> See Blommaert (2005) for a complete historical overview of a wider variety of CDA programmes.

epistemology; in the intervening decades, many offshoot theories have developed. Space and time do not allow for in-depth reconsideration of the foundational classics of CDA (Foucault, 1980; Habermas, 1971; Halliday, 1978; Moscovici, 1982; etc.) herein. These seminal works will instead be considered in relation to the CDA theories that have followed and developed in their tradition.

### **2.2.1 The Discourse-Historical Approach**

Like all other forms of CDA, the Discourse-Historical Approach (DHA) is interdisciplinary as it follows the principle of triangulation. This implies taking a whole range of empirical observations, theories and methods as well as background information into account (Reisigl & Wodak, 2009, p. 89) when analysing discourse. DHA triangulation is based on four levels of analysis: co-textual/co-discursive, intertextual/interdiscursive, extra-linguistic register, and historical/socio-political context (Reisigl & Wodak, 2009, p. 93). The final level of these four is what truly sets DHS apart from the other schools of CDA.

The discourse-historical approach, predictably, places supreme emphasis on historical context. DH analysts take a *diachronic* approach to analysing intertextual and interdiscursive relationships between utterances, texts, genres and discourses, as well as extra-linguistic social/sociological variables, histories of organizations or institutions, and situational variables (Reisigl & Wodak, 2009, p. 90). It is within this historical frame that the researcher may trace patterns of recontextualization (transference of allusions, evocations, arguments, etc. from one text to others) or decontextualization (the removal of one of these elements) and gain insights into the progression of social meanings over time (Reisigl & Wodak, 2009, p. 90).

One of the highest principles of the discourse-historical approach is the pursuit of practical application or change in discursive practices. As such, DH analysts are primarily concerned with the language of the powerful; influencing positive change in these discourses has the highest potential for wide-ranging benefit (Reisigl & Wodak, 2009, p. 88). DH analysts often focus on discourses of politics (Wodak, 2009, 2011) or the press (Krzyzanowski, Triandafyllidou, & Wodak, 2009), especially in regards to the ideologies of identity (Delanty, Wodak, & Jones, 2011; Wodak, Cillia, Reisigl, & Liebhart, 2009) and 'belonging' (Krzyzanowski & Wodak, 2009; L. Pearce & Wodak, 2010). The researchers cite the work of Marx and the Frankfurt School, believing that "texts are often sites of social struggle in that they manifest traces of differing ideological fights for dominance and hegemony, and find communication of power primarily in lexico-grammatical analysis, but also in the production of preferred genres, and in differences



in access to discourse by different social groups" (Reisigl & Wodak, 2009, p. 89). Discursively, traces of power patterns often appear in the form of us/them discourse (or 'othering' in other research traditions), legitimation strategies, employment of argumentation strategies such as *topoi*, "attribution of responsibility or guilt", "victim-agent reversal" and the "scapegoat strategy", and linguistic generalizations, analogies, and/or cross-mapped representation through metaphors (Matouschek, Januschek, & Wodak, 1995, p. 60). The current work of Ruth Wodak and other members of the so-called 'Vienna School' of DHA broadly concern investigations of 'othering', i.e. anti-Semitic, xenophobic, or racist discourses of political bodies. The analysis in this thesis is primarily focused on othering, of the racist, homophobic, and classist sorts.

In one of the earliest DHA papers, Menz, Wodak, Gruber, and Lutz (1988) examine news discourse from seven days on either side of the Hainburgh crisis<sup>4</sup>, exploring the underlying ideology of news publications through analysis of linguistic choice. 'Ideologies' are said here to meet four main criteria: 1) They serve the interests of dominant groups and reinforce the depreciation of lesser groups; 2) They exist without history, with naturalized origins and views; 3) They represent complete structures of oversimplified and falsified "truth"; 4) They exclude the possibility of challenge from sidelined or minority ideologies (Menz et al., 1988). These criteria echo the media discourse concept of hierarchical power relationships and the establishment of consensus, augmented by a further criteria (#4 above) expanding upon the concept of linguistic access as commodity. The researchers also introduce four categories of ideological language: "the adjuration of mythical groups" and "the strategy of seduction" relate to vilification of absent others who cannot be called upon to defend themselves (Menz et al., 1988). The remaining two ("the scapegoat strategy" and "black-and-white" depiction) can be tied to the more modern elaborated concept of 'othering' discussed in the previous chapter (van Dijk, 1988, 1991, 1998, 2006, 2008). By analysing a set of media texts and reports, the researchers are able to arrive at some interesting conclusions regarding positive/negative labelling strategies employed by the Press (van Dijk, 1991, p. 451), and explore the construal of interpersonal involvement by journalists (van Dijk, 1991, pp. 449–450). While the hypothetical framing of this study somewhat

---

<sup>4</sup> In December 1984, the construction of a hydroelectric power plant was planned to begin in an area 40km east of Vienna, in the Au Forest. In response to the scheduled clearing of the forest, thousands of members of the public camped out in protest, resisting multiple attempts at policed evacuation. Media outlets aligned themselves either alongside the protestors, the contractors, or the government, resulting in illustration of the event as equal parts environmental concern and civil disobedience (Menz et al., 1988).

overpowers the results, it is an interesting **foundational example of the applications of DHA in news analysis.**

More recently, Wodak (2003) has applied the **discourse-historical approach to join interpretation of the historical dimension of discourse with socio-linguistic analysis.** Examples are drawn from **media genres propagated by coalition government populist actions in the year 2000.** By analysing references, **predication, argumentation, perspectivation, and intensification/mitigation (Wodak, 2003, p. 139),** Wodak defines and identifies methods for systematic media exclusion and **states that, through careful rhetorical engineering, sociocultural ideology can be manipulated and history rewritten.** Referential and predicational strategies are the focus of all **analysis chapters in this work.**

Though not explicitly invoking a DHA framework, Greenslade (2005) has effectively undertaken a linear investigation of these same discursive **patterns in an Institute for Public Policy Research Asylum and Migration Working Paper exploring British and American media discourse on the topic of asylum.** Three historical case studies are used in support of his research: 1940s post-war Jewish portrayals, 1950s **West Indian immigration commentaries,** and 1960s reactions to the 'rivers of blood' Enoch Powell speech. Greenslade's work is performed in a loose CDA framework whereby the media is perceived as both reflecting and enhancing public opinion (Greenslade, 2005, p. 6). The researcher offers further commentary on the media's declining interest in impartial and factual reporting, linking this trend to increased stress on profit growth (Greenslade, 2005, pp. 9-12) and misinformed mimicry of readership ideology (Greenslade, 2005, pp. 13-14), allegations also levelled by other media discourse analysts (notably: Bell, 1991; Fowler et al., 1979; Fowler, 1991). In comparing these three historical cases with trends in modern asylum journalism, Greenslade reports an increase in hysterical terminology, lexical repetition, and ideological potency of the media (2005, p. 21). In addition to **compiling a so-called 'layman's' timeline of linguistically realised discrimination in British journalism,** Greenslade issues a call to action for creators and consumers of media. Though this article's research method is somewhat lacking in relations to existing work and ultimately in linguistic rigor, the sociological contexts provided enrich the findings with critical meaning, exemplify the cross-disciplinary allure of DHA-style studies, and underline the opportunity for social change through research.

Following on the initial concept of language as constructing and perpetuating social norms, comes the idea that a society's ethical statutes and moral values, then, must

somehow be encoded within its discourse. Lillie Chouliaraki has published a series of CDA-supported books and papers on morality in the media as it pertains to the construal of distant suffering. Chouliaraki believes that the manner in which the media exposes its audience to “stories of human tragedy and suffering” has the capacity to “[cultivate] audiences’ dispositions towards human affairs” and effectively “[propose] to them how to think and feel about the world” (Chouliaraki, 2008, p. 838). In a collection of small case studies, Chouliaraki undertakes multimodal analysis of a sample of stories from prime-time Greek, Danish, and British television that deal with the suffering of (mostly marginalized, minority) distant ‘others’. She notes that moral positions may be encoded through varying aesthetic presentation styles: suffering is construed as *cinematic entertainment*, an opportunity for *philanthropy*, or, most rarely, a *pamphleteering* exercise demanding action (Chouliaraki, 2008). The viewer is also encouraged to experience different degrees of moral agency: a spectator’s awe, an emotional but largely fruitless empathy, and only rarely, an impetus to act on the behalf of the distant ‘other’ (Chouliaraki, 2008). These stories exist within a hierarchy, and modern technology has transformed “the nature of media representation in ways that thoroughly fictionalize or manipulate suffering, thus cancelling or minimizing the demand for public action on the condition of suffering” (Chouliaraki, 2008). As a Critical Discourse Analyst, Chouliaraki uses her findings to advocate further investigation into the presentation of morality (or lack thereof) in the news. Though my own research concerns textual analysis of stories depicting suffering closer to ‘home’, it is interesting to see the overlap in the categorization of these events with epidemics and natural disasters in America, and indeed, to see how my data echoes the constriction of moral calls to action observed by this researcher.

### **2.2.2 The Socio-Cognitive Approach: Critical Discourse Studies**

The Socio-Cognitive Approach shares Foucauldian and Habermasian lineage with Norman Fairclough’s Dialectal-Relational Approach to CDA, and its political interests, but Wodak and Reisigl put forward that the “discourse-historical approach should be seen as an extension of van Dijk’s socio-cognitive model” as “[i]t attempts to incorporate historical-political and affective levels” (Wodak & Reisigl, 1999, p. 186). Rather than naming the Critical Linguistic tradition as the main influence of his work, van Dijk asserts that the roots of his ‘brand’ of CDA grew in Discourse Analysis, itself based on classical rhetoric (van Dijk, 1993). To a certain extent, these factors explain an enhanced emphasis on reception and/or cognition—the feature that sets van Dijk’s approach apart from other threads of CDA.

Though he is considered one of the forefathers of the program, as of late, Teun van Dijk has preferred the term Critical Discourse Studies (CDS) to Critical Discourse Analysis. This nominative distance away from 'classical' CDA foregrounds his conceptualization of CDS as encompassing more than just analysis; he envisages the approach as mutually inclusive of critical theory and critical applications (van Dijk, 2009, p. 62). CDS retains the socio-political motivations driving other schools of CDA, but places more emphasis on investigating problematic discursive patterns in comparison to fundamental 'norms', i.e. human rights (van Dijk, 2009). Analysis of 'normative' versus 'divergent' language is of dual use in this thesis; in the coming chapters, we will establish both socio-cultural and linguistic 'norms' before analysing popular conception of socio-cultural deviants and linguistic mobilization of atypical markers.

In order to measure 'norms' and identify deviations, van Dijk proposes research within "the discourse-cognition-society triangle" (van Dijk, 2009). Following in this research tradition, Jonathan Charteris-Black has lately pioneered work in what he calls Critical Metaphor Analysis (2004), a blend of CDS/CDA and Lakovian cognitive metaphor analysis. Charteris-Black's particular areas of interest include the impact of class, age, and gender on discussions of taboo topics (Charteris-Black, 2012; Seale & Charteris-Black, 2008), metaphor in political discourse (Charteris-Black, 2005, 2009), and metaphorical representations of immigration (Charteris-Black, 2006)—a paper that will be reviewed shortly. The researcher's continued influence from the discipline of gender studies and the field of cognitive linguistics emphasizes both the interdisciplinarity *and* intradisciplinarity of CDA.

Van Dijk stresses that CDS is not a method of CDA, but in fact a critical perspective from which one might conduct research on the interaction of cognition, discourse, and society (van Dijk, 2009). This denial of CDS as a method will become extremely relevant in the following section, where Corpus Linguistics is proposed as a viable method complementing CDA/CDS.

Using the CDS framework, Lupton found that the western media has a tradition of depicting both "innocent" and "guilty" victims of disease: those who have become ill through wilful engagement in deviant behaviour are construed as 'deserving' of their disease, whereas 'innocents' are considered, rather, to be victims of "accidental deviance" (1993, p. 309). Whereas the majority of research into AIDS/HIV risk discourse of has largely focused on "the representation of stigmatized minorities" (1993, p. 308) as 'guilty' parties, Lupton investigates Australian AIDS/HIV risk discourse as it relates to

'innocent' or non-deviant groups: heterosexual racial majority groups. The researcher allowed the texts to dictate salience, and chose areas of analysis by observing those topics considered most newsworthy by the press. The 'Grim Reaper' AIDS awareness campaign represented the threat metaphorically as a "vengeful killer" and an "enemy" who could be defeated with the "weapon" of health education (Lupton, 1993, p. 314). This tactic distanced the government from any blame for the spread of disease - having 'armed' the public with knowledge and fear, the people were left to monitor deviance within themselves and others (Lupton, 1993, p. 315). AIDS is also metaphorically represented as "an apocalyptic disaster" or "latter-day plague" (Lupton, 1993, p. 315) which strikes indeterminately. Quantification rhetoric hyperbolized the proportion of infected or at-risk people, and reduced populations to numerical values (Lupton, 1993, pp. 318–319), a theme which also emerges in my data. Beyond the previously common implication that AIDS is a celestial response to deviant behaviour, the Australian media texts warn that the disease is damnation for the sexual revolution and exists to encourage and enforce ubiquitous moral reformation. The researcher concludes that the media during this period preyed upon "anxieties, the need to define boundaries between Self and Other, to construct a *cordon sanitaire* between the contaminated and those at risk of contamination, to protect against invasion" (Lupton, 1993, p. 353). The categorization of at-risk groups expanded during this time to include taboo behaviour previously ignored in preference for already side-lined groups (i.e. drug users, homosexuals). In light of government 'education', those who still contract AIDS might be re-classified as wilfully deviant, sinful, and deserving of blame and Godly retribution. Similar discourses arise in analyses of my corpora, though these are identified through corpus-based methods, which are not employed in Lupton's study, indicating that prosodies 'not visible to the naked eye' might have been overlooked, and frequency statistics were not included.

### **2.2.3 Criticisms leading to the incorporation of a corpus-based method**

The Critical Discourse Analytical approach and others within this tradition are not without its criticisms. Critics claim that CDA can be Euro-centric (Chilton, 2005), biased in data selection (Tyrwhitt-Drake, 1999), and overly subjective in analysis (Titscher et al., 2000). These points are addressed in turn below. Perhaps owing to the fact that this approach was developed by European researchers (in European universities), influential applications of CDA are almost entirely Euro-centric. However, van Dijk's current focus on Latin American discourse (e.g. 2005) might indicate expansion in this area.

Furthermore, this thesis, written by an American researcher investigating American discourse, is itself a testimony to the growing international adoption of CDA.<sup>5</sup>

The accepted—and in fact prescribed—overt political stance taken by researchers is also a point of contention (Titscher et al., 2000). Critics are wary that critical discourse analysts might formulate conclusions from a position of personal intuition or distaste, and work backward. However, holding researchers to a standard of absolutely impartiality is a utopian fallacy; Burr argues that “objectivity is an impossibility, since each of us, of necessity, must encounter the world from some perspective or other...and the questions we come to ask about that world...must also of necessity arise from the assumptions that are embedded in our own perspective” (1995, p. 160). The socio-culturally motivated CDA has two tenets that work together to disable subjectivity running amok. On one hand, researchers must offer an explicit statement of political positioning. But on the other, they are obligated to constantly question this position in relation to the emergence of discursive proof. I believe that this is a strength of CDA; because critical discourse analysts make their motivating opinions clear, they are required to remain vigilant against undue influence of emotion and/or cognitive bias exerted over science. If cultural ideology is in constant flux, and language is one expression of the existing ideological structures of a given society, then it stands to reason that a study of the evolution of language should be grounded in a study of the historical development of the language’s society. With its explicitly political agenda and stress on the analysis of language as a product of socio-cultural context, CDA offers a unique platform from which one might survey the interface of society and discourse.

As much of the existing research in the CDA tradition is qualitatively oriented, leading some critics (e.g. Widdowson, 2004) to reproach critical discourse analysts for ‘cherry-picking’, or manually selecting instances which fit with a given hypothesis, I have decided upon the incorporation of a quantitative method to: assist in diachronic analysis; allow for statistical representation of results; and provide a more objective view of the data. This method is Corpus Linguistics.

---

<sup>5</sup> In May 2011, the University of Utah hosted its first conference on North American Critical Discourse Analysis, citing this shortage of cohesive geographic scholarship in the field. Speakers included: Jennifer Clary-Lemon of the University of Winnipeg, a discourse-historical analyst investigating construction of immigrant identities in Canada; Eric M. Haas of WestEd and Gustavo E. Fischman of Arizona State University, who apply Fairclough’s dialectal-relational approach in combination with Lakoff’s prototypes to analyze a 1,000-text sample of news reporting on higher education policies; and, notably, Kirk Johnson of the University of Mississippi and Randi Reppen of Northern Arizona University, who contributed to the Johnson, Dolan, Johnson, Sonnett, & Reppen (2010) paper concerning racial identity construction in broadcast news following Hurricane Katrina.

## **2.3 Corpus Assisted Discourse Studies and corpus-based critical discourse analysis**

Critical Discourse Analysis, interdisciplinary and politically motivated, is an ideal programme with which to carry out socio-cultural research. However, “CDA does not necessarily include a broad range of linguistic categories in each single analysis; one might get the impression that only a very few linguistic devices are central to CDA studies” (Wodak & Meyer, 2009, p. 21). The corpus linguistics approach to CDA is arguably the newest of the CDA ‘schools’, and the only one that was developed on the basis of existent framework. This approach draws (obviously) on corpus linguistic methods, in addition to the associated fields of computational sciences, mathematics and statistics. Though my conception of and approach to corpus linguistics is entirely methodological, a brief overview of the concept appears below. More detailed methodological information will be found in the next chapter, 3.

### **2.3.1 A brief overview of Corpus Linguistics**

Since the 1980s, utilization of a ‘new’ mode of computer-mediated analysis of language has been on the steady rise. Corpus Linguistics (CL) emerged in the past three decades in response to various shortcomings attributed to earlier methods, for instance: small sample sizes, reliance on simulated or intuitive examples of discourse, toilsome manual analyses, high rates of subjectivity and/or lack of accountability, and finally, lack of empirical, verifiable, statistical results (Svartvik, 1992, pp. 8–10). A corpus, or “a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language” (Sinclair, 1996) can now be compiled with relative ease using (predominantly) digitised collection methods<sup>6</sup>. Contemporary corpus linguists have exploited access to a generation of ever newer, faster, cheaper means of computational support by applying statistical and linguistic analyses to more comprehensive, diverse collections of texts.

At the time of writing, the method has been applied to nearly all areas of ‘traditional’ linguistics, facilitating a previously unheard of level of quantitative measure of language patterns and uses. Recently, CL has been adopted and applied across a variety of other

---

<sup>6</sup> While a corpus today is generally understood to be a principled collection of machine-readable texts (Sinclair, 1996) much of the method’s groundwork was actually laid in an era before machines were capable of such work. For overviews of historical development, see: Hunston, 2002; McEnery & Hardie, 2012; Svartvik, 1992.

linguistic fields. Corpus-based analyses now offer insights into register (Biber, Johansson, Leech, Conrad, & Finegan, 1999), lexicogrammar (Hunston & Francis, 2000; Sinclair, 1991), English language pedagogy (Coxhead & Byrd, 2007; Coxhead, 2000, 2008), and description of language, and phraseology, e.g. collocation (Kennedy, 1991), lexical bundles (Biber & Conrad, 1999), and semantic sequencing (Hunston, 2002). As researchers from branches of applied and theoretical linguistics find new ways to exploit the powerful statistical toolset bundled in CL, the popularity of the method continues to grow exponentially. A good deal of knowledge about language 'in general' (as a product) has been derived from a multitude of early studies on large, multi-generic corpora, but in the 1990s, a number of researchers, often linked to Birmingham University, began publishing research that used corpus linguistics approaches to take a critical approach to language analysis.

### ***2.3.2 Schools of Corpus Assisted Discourse Studies and corpus-based critical discourse analysis***

In this recently developing field of research, several 'schools' of discourse analysis have recently applied corpus-based methodologies in the aim of developing a deeper understanding of language as a product of context, including processes of production and reactions in reception in extension of more traditional studies of lexicogrammatical output (Stubbs, 1996, p. 8). Some researchers (notably: Wodak, Fairclough, and van Dijk) remain aligned with the Marxist tradition and seek critical insight into ideologies and power imbalances encoded in language; their work is politically motivated, and as such, researchers in their traditions are expected to provide extensive contextual analysis, acknowledge their own ideological biases and undertake research in the pursuit of societal change. Other researchers (i.e. Partington, Stubbs, and Duguid) are equally interested in allowing corpus analysis to expose antithetical or unexpected language features while sometimes backgrounding the role of the researcher (and his or her identity or political agenda). My research is a critical analysis of discourses of discrimination and the context that they occur within; however, the primary agenda of this thesis is not to affect cultural change, but rather to explore the application of computing technology to discourse analysis, and to draw conclusions about the strengths and limitations of such when investigating ideologies encoded in discourse. Common motivations for similar studies (both explicitly critical and otherwise) are illustrated below.



### **2.3.3 Strengths of the mixed method**

A major strength of the corpus linguistics approach to discourse analysis is allowance for 'triangulation', or examination of data from perspectives in a variety of theoretical or methodological frameworks (Wodak & Meyer, 2009, p. 31). In referring to the smaller discourse samples, there is a danger that discourse analysts might "proclaim features as typical rather than build up the notion of 'typicality' on the basis of frequency" and therefore select atypical texts or over-generalise localized results (Mautner, 1995, p. 3). Using statistical measures, Mautner suggests that linguistic 'norms' can be accurately represented, and deviations from these norms can be more objectively identified (Mautner, 1995). Drawing upon a large enough corpus, this degree of accrued certainty also has implications for the study of "semantic prosody" (or an attitude associated with a lexical item over time and across texts)—an area of interest to discourse analysts previously immeasurable with any statistical significance (Louw, 1993).

The combination of corpus linguistic methods with discourse analytical approaches offers benefits for both fields. Using a corpus, analysts can choose a 'top-down' approach, giving statistical support to hypotheses, or a 'bottom-up' approach, allowing the corpus to guide and inform analysis of items of statistical significance (Mautner, 1995). Finally, automated concordance, collocation, and generation of key words (or words significantly over- or underrepresented in one corpus compared to another) are time-saving devices in a field that is conventionally thought to require a good deal of timely manual analysis (Mautner, 1995).

In summary, combining CL with a discourse analytical theory "helps researchers cope with large amounts of textual data, thus bolstering...empirical foundations, reducing researchers' bias and enhancing the credibility of analyses" (Mautner, 2009, p. 138) and the synergy is at the centre of a rapidly developing field of research.

### **2.3.4 Key corpus-based discourse analyses of identity**

Gerlinde Mautner arguably spearheaded the adoption CDA-CL combined approach with the presentation of a Lancaster University Centre for Computer Corpus Research on Language (UCREL) working paper in 1995. In that early piece on the cooperative application of CDA and corpus linguistics, she details the process of extracting concordance data from media corpora. The researcher provides some examples of CDA in combination with corpus analysis, most notably in the differing treatment of actors and pronouns across a variety of news sources; for example, comparing concordance profiling for markers of self ("this newspaper") and readership ("people"). The practice

of funnelling top-level concordance lines through to more granular analysis is interesting (and relevant), as this seems to continue as a standard starting point in more modern work. Since this time, Mautner has continued to make a variety of important contributions to this field, both individually and collaboratively.

Koller and Mautner (2004) offer an overview of linguistic analysis performed in the developing collaborative space between CL and CDA. The researchers concede to popular criticisms of this combined method, namely, the time-consuming nature of keying data and converting it to a computer-readable file, and the attendant loss of multimodal accompaniment. However, they maintain that the potential for quantitative insights into what has historically been a largely qualitative field outweighs these shortcomings. Koller and Mautner survey existing work in this field, reiterating the popularity of focus on "features such as action type, pronouns, proper names, and vocabulary" (Mautner & Koller, 2004). Because researcher interest dictates the starting point in a vast universe of available corpora, the example cases within this article explore Mautner's intuitions about pronoun usage and semantic prosody in European news. The negative semantic prosody of "federal\*7" uncovered within the corpus reinforces the researchers' claim that collocation analysis can add a layer of ideological appraisal analysis to collections of text too broad for traditional CDA tools to manage. Future researchers are warned against the fallacy of favouring only highest-frequency data, to the detriment of infrequent and highly significant marked items. In keeping with the CDA tradition, researchers are further reminded to consider text within context by referring to extended samples of data in cases of semantic ambiguity. To avoid making erroneous, large generalizations based on overly small discursive samples, it is recommended to compare texts within a common genre, or to make use of large-scale corpora available online for comparison. In this study, the researchers conclude that the exhaustive nature of CL in combination with the initial analytical bases provided by collocation and concordance complement the qualitative nature of CDA neatly.

In response to Koller and Mautner's negative appraisal of menial aspects of data entry, it must be mentioned that the number of available digitized texts—especially in the media genres—has grown exponentially in years following publication of their 2004 study. While this entails a dramatic reduction in manual entry, the technology has also caused an even greater disconnect between text, typography, imagery, and layout. Incompatibility with multimodal study remains a major shortcoming of corpus-based

---

<sup>7</sup> \* is a wildcard indicating a string of any number of characters, allowing for primitive lemmatisation.

analysis of media discourse, though some researchers, notably Dawn Knight (e.g. Adolphs, Knight, & Carter, 2011) and Monika Bednarek (e.g. Bednarek & Caple, 2012) are making some important strides in the development of frameworks for precisely this sort of analysis.

Over the past decade, an increasing number of researchers and scholars have exploited the complementary features of discourse analysis and corpus linguistics to analyse portrayal of relatively powerless individuals and social groups; for instance, in frequent lexico-grammatical constructions in large general corpora, Hunston (2002) found the deaf people are construed in a variety of ways (as disabled, to be pitied, as a group with rights) in British English, and Mautner (2007) exposes negative stereotyping of ageing and the elderly. Because the field of research in corpus-based analyses of discourse and identity is no longer as small as it was 15 years ago, this literature review will concentrate most deeply on construction of identities based around qualities researched in this thesis, namely: gender, sexuality, race, poverty, and politics. As discursive choices are heavily influenced by both context and genre, where possible I will focus on analyses of identity in media discourse, highlighting methodological/theoretical contributions that will be drawn from each study.

The field of gender representation is arguably one of the richest areas of corpus-based discourse analysis. While gender does not play as central a role in research as—for instance—sexuality, some methods employed by researchers in this area have been influential in this thesis and will be quickly surveyed. Biber and Burges (2001) examine the ARCHER corpus of dramatic dialogue to draw conclusions about changes in gender representation over time, and metaphor in gendered discourses has been the subject of a series of studies by Koller (2004). Sigley and Holmes (2002), Pearce (2008), and Baker (2010) all investigate frequencies, collocations, and grammatical behaviours of search words such as *men*, *women*, and *girls* in reference corpora. Pearce (M. Pearce, 2008) finds that, overall, males collocated with verbs of aggression while women collocated more strongly with verbs of a sexual nature, while Baker (2006) finds that *bachelors* are imbued with a greater positive semantic prosody contrasting to a negative semantic prosody associated with their female counterparts, *spinsters*. Caldas-Coulthard and Moon (2010) contrast the adjectival pre-modification of *man*, *woman*, *girl* and *boy* in corpora of British broadsheet and tabloid news. They find that while males are evaluated on the basis of their behaviour, societal function and status, females are evaluated on their appearance and sexuality. Semantic preference of node words was not the only site of difference; in frequency analysis, researchers were also able to

describe and quantify divergent preferences in *type* of adjectives employed by broadsheets and tabloids, showing heightened use of 'academic' adjectives in broadsheets and use of playful misspellings, blends, and diminutives in the tabloids (C. R. Caldas-Coulthard & Moon, 2010). Similar analyses of adjectival collocates serve as the first stage of study in Chapters 4 and 7, but as American publications are not so neatly delineated as broadsheets or discourse, highest- and lowest-frequency patterns will be discussed in the frame of specific, single publications, or targeted publication centres.

While sometimes related to gender identities in research, dedicated corpus-based discourse analysis of identities tied to sexuality are markedly more rare. Baker (2004) compares one corpus of gay erotic texts to a similarly-sized corpus of lesbian erotic texts<sup>8</sup>, aiming to uncover deviations or similarities in their constructions of gender identity. Qualitative analysis at the lexical level is extended to include determination of 'sense'; it is through this layer of analysis that Baker distinguishes semantic differences in the appearance of keywords shared between the two corpora. Once salient keywords have been identified, researchers might extend their focus beyond lexemes and consider the key combinations of 'bundles' to map specific semantic preference of one corpus in comparison to another. The researcher further advises categorization of keywords for semantic analysis, though this practice may introduce an element of subjectivity that corpus analysts wish to minimize. In response to both the methodological suggestion and possibility of limitation, semantic preference of node words will be explored in this thesis using an automated semantic tagger in combination with manual analysis. Baker has also undertaken two studies of construction of gay men from both the 'inside' and the 'outside': 1) In 2005, he analysed a corpus of personal ads in *Gay News/Times* to illustrate how gay men attempted to construct their own identities in agreement with 'in-group' values of masculinity, while 2) another study used a corpus of speeches from the British House of Lords to investigate representation of this social group by another party, and found that out-group qualities (such as danger and deviancy) were frequently correlated. This has implications about the *direction* of representation for this thesis; where possible, published discourses from both the 'inside' (e.g. HIV+ individuals, African-Americans) must be contrasted to 'outside' discourses (e.g. politicians) for a more accurate understanding of the accepted parameters for social group inclusion. The literature does not provide large-scale diachronic studies of the representation of social

---

<sup>8</sup> Liz Morrish and Veronika Koller are two other researchers developing work in the field of sexual identity construction in discourse, sometimes with corpus-based methods. However, their work focuses mainly on representation of lesbian identity, an area that will not be explored in this thesis, and has therefore been excluded from the literature review.

groups defined by their sexuality, and this will be approached in Chapter 9 informed by the findings above and shaped by a method grounded in semantic categorization.

This caveat also holds true for identification and representation of racial/ethnic groups. The modern usage of the words 'ethnic', 'racial', and 'tribal' (with their morphological forms) is analysed by Krishnamurthy (1996) through a comparison of discursive application in the media, COBUILD definitions and glosses, and Bank of English (BEC) statistical frequencies. Krishnamurthy finds that journalists make selections between the near-synonyms on a largely geographical basis (i.e. 'here in Britain' versus 'there in Yugoslavia'). Findings from both the COBUILD dictionary and the BEC show that word selection is largely related to field and genre. Unfortunately, the basis of media text selection in Krishnamurthy's (Krishnamurthy, 1996) work is not described, and it could be reasoned that specific texts were provided in support of conclusions already reached through corpus and COBUILD analysis. This study is an interesting example of 'nested' research, in which semantic choice is described prescriptively (through dictionary definition), critically (in actual media occurrence) and statistically (with corpus analysis). This thesis is also 'nested', with frequent naming strategies and identity attributes being targeted through a corpus of media texts, analysed critically, and contrasted to dictionary definitions.

In a 2007 study of a corpus of New Zealand reportage, Barnett, Hodgetts, Bikoa, Chamberlain and Karapufind examples of 'victim blaming' and emphasis on individual responsibility in media representations of the poor. The study uses a 'text-and-context' approach (2007, p. 301), following the narrative development of one news story, and found that the media "coverage relied on construction of...distinctions between the deserving and undeserving poor" (ibid.: 305). This is achieved through emphasis on either individual responsibility (and blame) or community responsibility (and pity). Attributes of people determining whether they are represented as deserving or undeserving of support (social, medical, financial, emotional) will be explored in Chapters 5, 6, 7, and 8 of this thesis.

It is important to note that representation is not always tied to well-defined identities; corpus-based analyses of more broad-reaching forms of 'othering' are also salient here. In a 2006 appraisal analysis of the Xhosa English corpus, Adendorff and de Klerk posit that people tend to 'other' AIDS/HIV "in the belief that it only happens to other people, especially homosexuals, poor people, and people of 'other' ethnic groups" (2009, p. 71). Corpus evidence supports this hypothesis; frequent emotive and appraising lexis

indicate attempts to “identify *others* as being responsible for the spread of AIDS or for being responsible for curbing it – but absolve themselves”. Here representation is not so pointedly fixed on one or more specific identities, but to construction of an ‘other’ belonging to any social group outside of one’s own.

The linguistic construction of political identities (spanning from those of individual politicians to those of entire parties) has been another interesting field of research focus in recent years. In 2000, Norman Fairclough published an influential volume on the way that the British ‘New Labour’ party or the ‘Third Way’ was “effectively talked into being” through public construction of ‘open’ and ‘responsive’ but simultaneously ‘tough’ and ‘fair’ (Fairclough, 2000, pp. 5–10). L’Hôte (2010) analyses a similar corpus of new Labour texts using Wmatrix to calculate key semantic categories, a tool and method applied later in this thesis. Partington (2003) employs a corpus of White House press briefings to investigate the expression of ideologies of (and relationships between) speakers, in part through their persuasive strategies, employment of metaphors, and conveyance of opinion and stance. Duguid (2007) contrasts outsider perceptions of the identities of those working at Number 10 Downing Street with a corpus of courtroom transcriptions wherein (among others) Prime Minister Blair and special advisors construct their *own* identities. Contrary to press and popular construction of these leaders as masterful or in control, keyness analysis exposes high preference for lexis in the semantic field of difficulty and high instantiation of locative/existential verb forms (i.e. *there were*) constructing the speaker as a static, helpless narrator. Exposed through analysis of high-frequency and key words, the build-up of these and other patterns in the corpus belie outsider identity perception of politicians’ identities with insider constructions of immobility, impotency, and uncertainty (Duguid, 2007). Negotiation of political identity in media discourse (and reproduced political discourse) will be analysed in this thesis, often employing similar tools, e.g. investigation of highly frequent and key words or semantic fields, collocation restricted to part-of-speech, and cognitive metaphor analysis.

In a more recent study, Gabrielatos and Baker (2008) combine corpus linguistics and CDA in an analysis of political identity construction around refugees and asylum seekers in the media. Media texts mined from the LexisNexis database are analysed for keyness and collocation of asylum-related nodes to draw conclusions about their semantic/discursive profiles and socio-ideological contexts. Counting the appearances of asylum-related articles in the news, Gabrielatos and Baker find an upward overall trend punctuated with spikes corresponding to political or social events. This indicates a rising

media interest in the topic, as well as a growing discursive connection between asylum-seekers and matters of socio-political upheaval. A new technique—the calculation of “consistent collocates”—is developed therein as method of highlighting collocates appearing across a corpus while disregarding statistically skewed “seasonal” collocate appearances. This is an improvement over previous CDA/CL studies, sometimes criticized for relying too heavily on pure frequency statistics and therefore concluding in with somewhat poorly distributed results. Focusing on c-collocates allows for more accurate description of unfolding prosody in discourse. This in itself opens possibilities for a deeper study of the relationship between discursive partners, the analysis of which has been a contested issue with critical discourse analysts. Consistent collocates are grouped in semantic clusters and measured on rate, source, and time of appearance—all aspects of analysis that will be echoed in this thesis. A methodological combination of CDA and CL results here in an interesting diachronic view of semantic prosody. The researchers suggest further work in measuring the appearance of keywords over time, a method that will be explored in Chapter 9.

Five studies additionally worth mentioning take a different angle, focussing on construal of social actors through the lens of journalists’ choices, thereby uncovering ideologies supported by specific publications and the cultures they are ‘representative’ of. Caldas-Coulthard (C. R. Caldas-Coulthard, 1993, 1995) critically analyses corpora of news language, concluding that women are often mis- or underrepresented in attribution and quotation patterns, which indicates restriction from power in media discourse. A large corpus of news texts from the *Guardian* is analysed by Scott (2002) to depict publication ideology by ‘clumping’ key words together by sense and semantic field, a concept which is refined here. Bondi (2007) combines aspects of CL, CDA, and SFL in the analysis of institutional Bloody Sunday Inquiry discourse and its associated media reportage. Purpose-built corpora of court proceeding transcripts and news stories are compared both to one another and to the full ‘baseline’ BNC, a method of cross-reference which promises interesting intertextual results. By running concordance lines of the major ideational ‘actors’, performing lexical analysis of associated modifiers, and grouping these by category, the researcher is able to show that the semantic construal of participants is largely emotive (and negative), and highly dependent on journalist positioning. Negative and positive judgement of participants is a large factor in the present study, though this analysis is undertaken with a far more quantitatively determined method. Another project titled ‘Corpus and Discourse: a quantitative and qualitative linguistic analysis of political and media discourse on the conflict in Iraq in

2003' also considers the evaluative stance of politicians and journalists, with the aim of exposing "cultural presuppositions and expectations...which provide a sort of cultural frame or backdrop for the coverage" (Haarman & Lombardo, 2009, p. 2). In a volume resulting from the project, multi-modal analysis of a variety of features (including patterns of attribution) leads all researchers to find "strong links between the *cultures* of news providers and the *news* they produce" (Haarman & Lombardo, 2009), a facet which I also seek to explore here. In one chapter, Marchi and Venutti (2009) remark on the usefulness of including XML-valid corpus mark-up, particularly in compiling subcorpora on external linguistic features such as date. The corpora used in this thesis have employed a similar method to meet similar ends.

Nomination strategies have also been studied from the perspective of Cognitive Metaphor Theory. Carrying on the tradition of cognitive CDA developed by van Dijk, Jonathan Charteris-Black (2004, 2005, 2006) has extended Lakoff's cognitive metaphor theory to a blended method he calls Critical Metaphor Analysis. Most influential in my own research is Charteris-Black's (2006) paper investigating rhetorical strategies of British immigration metaphors employed during the 2005 elections. Metaphorical analysis is twofold. The researcher first compares appearances of 'immigrant' to its nominalized counterpart 'immigration' in centre and far-right communication, then describes portrayal of immigration in variations of disaster/container metaphors. When contrasting the language of opposing political standpoints, Charteris-Black claims that "[t]he lexical analysis of the corpus shows that the centre-right discusses 'immigration' while the far right tends to discuss 'immigrants'" (2006, p. 568). This result characterizes the centre-right's general preference for discussing immigration as a process, while reserving congruent (often racist) attributive language for the far-right. The construction of disaster metaphors (i.e. "floods of immigrants") is presented as a journalistic method of invoking emotive responses to immigration akin to that of social destruction. Similar uses of disaster metaphors will be investigated in the Hurricane Katrina corpus in the following chapters. Charteris-Black also finds that container metaphors (wherein Britain is described as a bounded box, often 'full') are employed to both define the metaphorical boundary between 'us' and 'them', and communicate a fear of 'bursting' supposedly shared with the readership. The research presented in this article is enriched with inclusion of broad socio-historical background information, representative sample texts, and psychological theory of metaphor perception, highlighting the harmonious blend of CDA and CL techniques.



In a brief audit of early-1900s debate texts, Gerald V. O'Brien (2003) calls attention to the variety of metaphors employed in the depiction of anti-immigration sentiment in the United States. The researcher notes the appearance of metaphors portraying immigrants as a type of: organism, object, natural catastrophe, or animal. The organism metaphors are mobilized describing immigrants as infectious invaders in the pristine social body of the existing nation, reproducing at the rapid rate of dire disease, and defiling the environment with ghettos and other pockets of filth. Object metaphors, wherein the immigrants are described as interchangeable non-human items, often tie into organism metaphors; the portrayal of immigrants as a unit (i.e. of commerce) contributes to an ideology of segregation whereby the population can be set apart from native culture and considered in terms of mass numbers or over-generalized characteristics. As in the 2008 study by Gabrielatos and Baker, O'Brien notes a tendency toward natural disaster (especially flood) metaphor in delineation of an 'other'. This us-against-them metaphorical process is also realized in schemata of war, where the immigrant is an invader capable of dominating or destroying a native people. Finally, animal and other sub-human metaphors set immigrants apart from civilized (humanized) Americans under threat of alien attack. The researcher does make concessions when explaining the popularity of certain metaphorical expressions—for instance, immigrants often *were* carriers of disease. Though O'Brien ultimately utilizes a very small discursive sample to support his arguments, this article's strength is in the firm definition of metaphorical development and parameters. The relationships drawn to the 1900s sociological climate offer a critical insight into the reasons for metaphorical emergence and adoption; though the researcher does not explicitly name CDA as a contributing framework, the program's same historical, socio-cultural concern underpins his work. O'Brien's conclusion corroborates previous findings by Calavita (1984) and Fairchild (1926), namely that "early immigration restriction policies in the United States focused not on race or nationality, but rather keeping out those immigrants who possessed 'undesirable' characteristics, such as various disabling conditions, infectious diseases, pauperism, and anarchistic tendencies" (O'Brien, 2003).

### **2.3.5 Criticisms and responses**

Owing to the recent technological revolution, nearly all of the traditional criticisms of Corpus Linguistics (i.e. difficulty of collecting texts, time-consuming and subjective nature of manual searching) are now out-dated. Most of the current criticisms of CL stem from reservations about one essential trait of the discipline: there is no defined framework within which the method is best applied. In the absence of a universally

accepted 'manual' giving step-by-step instructions for carrying out corpus-based analysis, researchers have been at liberty to treat CL as a veritable methodological grab-bag, choosing complementary theories, selecting analytical tools, and adjusting statistical thresholds on an as-needed basis. This level of adaptability is simultaneously considered (by critics) to be the method's greatest weakness and celebrated (by corpus linguists) as one of the field's greatest strengths. Let us explore this dichotomy further.

Like all fields of linguistics, the data selected for analysis must be well suited to the hypothesis being tested. Biber (1993) stresses the importance of balance and representation in the process of corpus-building. Though neo-Firthian linguists might argue that a corpus itself could serve as a hypothesis, we are dealing here with corpora that have been specifically built either to represent a reasonable cross-section of language, or to offer a large sample of discourse from a selective genre. Therefore it is important to address questions of balance and representativeness during the process of data collection, not to assure that generalizations can be drawn about language as a whole, but to assume a fair stance in analysis of media texts across a variety of genres. Because the entire population of possible media texts cannot be truly measured in a manner suggested by Biber (1993), a cross-sample of all media texts concerning specific events within a bounded time period will populate each corpus.

While technological advancement has allowed for ever-larger corpora to be assembled and analysed at ever-faster rates, the use of computers in the study of discourse is far from problematic. Corpora are read using machines, and machines are unable to 'read' texts as humans would. To a machine, discourse is a collection of strings whose frequency or co-occurrence might be calculated, but whose intuitive meaning cannot be measured (Widdowson, 2000) and whose context remains largely isolated from the text itself (Partington, 1998, p. 145). Further, frequency disambiguation between words occupying multiple parts of speech or multiple senses of polysemous words (e.g. *treat* or *conflict*) cannot be fully automated. Due to these methodological issues, Stubbs and Gerbig suggest that corpus linguists are guilty of "counting what is easy to count" (1993, p. 78). This limitation necessitates the combination of CL (as a method) with another branch of linguistics (as informing theory). This thesis regards intuition or discourse 'reception' in the frame of Critical Discourse Analysis.

Gerlinde Mautner (a critical discourse analyst in the Foucauldian tradition) became interested in computer-mediated critical quantitative methods after noting that corpus methods had, through the early 1990s, been applied mainly to lexicographical studies or

other descriptive fields. In her 1995 technical paper, Mautner noted that “[i]t is not (yet) common practice to harness the computer in the service of some form of ‘critical inquiry’” (Mautner, 1995), attributing this trend in part to a popular sentiment voiced by Roger Fowler: humans—not machines—possess the sort of sociological sensitivity required for critical analysis (1991, p. 68). However, this researcher takes issue with the common CDA sacrifice of breadth for depth. She refers to the likelihood that CD analysts might “proclaim features as typical rather than build up the notion of ‘typicality’ on the basis of frequency” as a risk factor in the selection of atypical texts or the overgeneralization of localized results (Mautner, 1995). In a perfect world, researchers from all disciplines of the ‘interdisciplinary’ CDA approach could be employed on a single project to assure that such skews are minimized, if not eliminated. But in the absence of the necessary funding or time to carry out such an interdisciplinary exercise on every text, Mautner believes that the adoption of computerized methods could grant a similar level of objectivity and speed (Mautner, 1995). Simply, she denies that suggestion that qualitative critical analysis and quantitative research methods must be mutually exclusive.

## **2.4 Summary**

Critical Discourse Analysis, a theoretical framework with few standardized tools, seems perfectly suited to the incorporation of Corpus Linguistics, a methodological tool with few standardized theoretical goals. CDA can be guided through machine-generated statistical data, and carried out systematically with the adoption of largely automated technology. Because computers cannot know standards of discourse until they are given calculation formulae, corpus analysis is rule-governed, meaning that manual categorization and its attendant probability of inconsistency have largely become a thing of the past.

In a way, CDA could be seen as a ‘top-down’ approach in which one sets out in search of power in discourse, and traces these to linguistic realisation. Conversely, CL could be seen as a ‘bottom-up’ approach in which one begins with a raw collection of texts and attempts to link observed patterns to contextual influences. An oscillation between these two allows for an informed analysis of triangulation.

As CDA performed by van Dijk and Wodak has a historical base in Cognitive Linguistics, it is important to note that CL has a powerful statistical measure which models cognitive representations of semantic fields or associated lexis: collocation. The use of collocation

to explore the construal of social actors (and semantic preference for identity groups) will be discussed in the next chapter.

Here, Critical Discourse Analysis will be mobilised as the overarching ideological framework for analysis of the context of culture. The rich tradition attached to CDA, particularly in regards to media discourse analysis, will guide this research from the theoretical conception through to corpus-based statistical representation.

## **Chapter 3: Data and methods**

This dissertation focuses on analysis of two custom-collected specialized ('opportunistic') corpora, or a small collection of tightly defined, related texts (McEneary, Xiao, & Tono, 2005). In this chapter, I describe the: methods of collecting and cleansing this data (section 3.1), resulting corpora (3.2), methods of analysis (3.3) and finally, tools and techniques utilized (3.4).

### **3.1 Data collection and cleansing**

#### **3.1.1 Factiva search builder**

All articles were gathered online from *Factiva*, a news service operated by Dow Jones<sup>9</sup>. In the 'search builder' interface, users are able to query the engine for articles using "industry, subject, region and company indexing terms, as well as source and language" in combination with free text variables (Dow Jones Factiva, 2009, p. 6). *Factiva* was selected over other, similar, digital news collections (such as *Nexis UK*) because of this highly-customizable search function in addition to its relatively large collection of American publications and ability to export files directly to XML<sup>10</sup> format (this has now been disabled). The engine purports to offer over 20,000 sources in 22 languages, a feature which could allow for further contrastive studies beyond the scope of this thesis.

There is one major caveat in working with digital databases of news; it is important to note here that while the *Factiva* database is large, it is not exhaustive. Within the Hurricane Katrina sampling frame (section 3.1.2), 24 publications met inclusion criteria, and two of these publications were represented by fewer than 100 articles each within the entire database. The AIDS/HIV sampling frame included 40 candidate publications, 12 of which were entered in the *Factiva* database but contained one or zero articles for the target time period. There are certainly a great deal more American print publications than these, and though the most widely-circulated are present in high frequencies, the corpora are limited in overall size and potential for generalisations by the content of *Factiva*. Many publications may not yet have digitised back issues, or might even have chosen to selectively exclude items from digitisation efforts in cases where content could

---

<sup>9</sup> <http://www.dowjones.com/factiva/>

<sup>10</sup> XML (Extensible Markup Language) encodes metadata about texts in a format readable by both machines and humans,

cause offense to modern sensibilities. This is of particular pertinence to the AIDS/HIV corpus, where articles published in the early 1980s might not yet have been (or may never be) digitised. However, the large size and distribution across publications in the resulting corpora can guarantee that the essence, at least, of the reportage of the time has emerged.

The specialized custom corpora employed in this dissertation are made up of a huge variety of text types (editorials, letters to the editor, front page, etc.) all belonging to the same genre: American disaster/epidemic reporting. Because this is a specialized corpus designed by an all-inclusive sample of data based on extra-linguistic features (genre, time period) and linguistic attributes (inclusion of nodes indicating any given text's 'aboutness' of Hurricane Katrina or HIV/AIDS), methods of random sampling to ensure representativeness and balance will not be discussed, and instead a report of the sampling frame will precede description of the completed corpora.

### **3.1.2 Sampling frame**

The sampling unit selected for this work was a single article (including headline, by-line, dateline, source name, publication date, and publication section (if available)). The sampling frame consists of *all* articles from major American print publications meeting the following inclusion parameters:

- **Hurricane Katrina**
  - **Dates:** 1 August 2005 – 31 August 2006
    - Designed to include the entirety of storm reporting leading up to touchdown in August 2005, and inclusive of a frequency spike reflecting anniversary reporting in August 2006
  - **Search terms:** *Katrina* AND (*hurricane* OR *storm* OR *flood* OR *disaster*)  
*Katrina* was included for recall, gathering all articles containing mention of the storm nomenclature; addition of another 'storm marker word' greatly increased precision, excluding cases where Katrina is the name of another participant and the storm is not the topic of reporting
  - **Sources:** Major American print publications
    - A *Factiva* category of "sources covering general news and business news that are considered key publications in their region by virtue of circulation or reputation" (Dow Jones Factiva, 2009)

- *Publications excluded for non-print format:* Newsweek Web Exclusive, The Wall Street Journal (Online and Print)
  - *Publications excluded for lack of affiliation with print source:* Dow Jones Business News, Dow Jones News Service
- **AIDS/HIV**
- **Dates:** 1 January 1981 – 31 December 2009
    - This was designed as a diachronic study of three discrete decades (1980s, 1990s, and 2000s); selection dates reflect the natural terminus of the 2000s, and the unfortunate lack of data in *Factiva* meeting search criteria in the year 1980
  - **Search terms:** Adapted to cover naming conventions of the times
    - 1981: *gay* OR *homosexual* OR *opportunistic* OR *cancer* OR *plague* OR *Kaposi's*
    - 1982: *gay*: *gay* OR *homosexual* OR *opportunistic* OR *cancer* OR *plague* OR *Kaposi's* OR *G.R.I.D.* OR *GRID* OR *A.I.D.S.* OR *AIDS* OR *H.I.V.* OR *HIV*
    - 1983 onward: *A.I.D.S.* OR *AIDS* OR *H.I.V.* OR *HIV*
  - **Sources:** Major American print publications
    - *Publications excluded for non-print format:* Barron's (Online and Print), Barron's Online, Dow Jones Online News, Forbes ASAP, Forbes FYI, MarketWatch, Newsweek Web Exclusive, The Wall Street Journal (Online and Print), Washington Post.com
    - *Publications excluded for non-news content:* Dow Jones News Service - Ticker, WSJ Guides
    - *Publications excluded for lack of affiliation with print source:* Dow Jones Business News, Dow Jones News Service
    - *Publications excluded for non-American content:* Winnipeg Free Press, Forbes Global

### **3.1.3 Corpus annotation and markup**

Articles qualifying for inclusion were exported with XML mark-up, which conserved all of the metadata stored in the *Factiva* database. While the option to export with XML mark-up (not available in *Nexis UK*) was extremely helpful, *Factiva* does not allow for case-specific searches or secondary searches (e.g. *katrina AND (storm OR flood)*), and its duplicate filter was not always completely thorough. Therefore, a series of PHP<sup>11</sup> scripts were written to:

1. Separate individual texts from each 100-batch set

---

<sup>11</sup> PHP: Hypertext Preprocessor: a general-purpose scripting language enabling large batch operations to multiple files through the terminal.

2. Compile a database of metadata (including headlines, by-lines, datelines, source names, publication dates, and where available, publication sections) from the existing XML mark-up
3. Verify appropriateness of the articles by:
  - a. Discarding non-print titles or non-American publications
  - b. Discarding 'false positive' articles:
    - i. For the Katrina corpus: articles not containing *hurricane* OR *storm* OR *flood* OR *disaster* in the body, in addition to the search term *Katrina*
    - ii. For the AIDS/HIV corpus: articles only containing search terms in lowercase (for instance, *foreign aids*)
4. Generate unique text identification numbers based on publication name, date, and *Factiva* accession numbers
5. Overwrite texts with duplicate text identification numbers not filtered by *Factiva's* duplicate blocker
6. Cleanse data
  - a. Strip extraneous data (i.e. photo captions) by targeting content in the article body only
  - b. Replace special characters
7. Output cleansed texts in files named with unique identification numbers, and containing XML mark-up headers

Texts were then indexed into the CQPweb corpus analysis system, which annotates data using the seventh version of the UCREL Constituent Likelihood Automatic Word-tagging System (CLAWS) and the UCREL Semantic Analysis System (USAS), assigning parts-of-speech and semantic tags to each text (Hardie, 2012). Further description of this tool will be provided after a brief description of the corpora, below.

## **3.2 Description of the corpora**

### **3.2.1 Hurricane Katrina**

The total size of the Katrina corpus is 36,736,679 words in 41,964 texts from 24 publications. The most highly represented publication in this corpus by a large margin is the *New Orleans Times-Picayune* with 16,538 texts, making up 39.41% of all of the texts in the corpus (see Table 3-1 below for article counts from all publications). Given that the city of its publication was ground zero for the storm, this is not a surprising finding, but one that must be taken into account in the analyses. Though the data will be skewed



toward reportage from the editorial perspective of the *Times-Picayune*, the advantage is that the voice from the area most affected will be very well represented amidst the texts from other cities.

Source name	Article count
Atlanta Journal - Constitution	2,463
Barron's	147
Boston Globe	1,506
Chicago Sun-Times	1,189
Christian Science Monitor	411
Denver Post	645
Forbes	51
Houston Chronicle	6
New York Daily News	927
New York Post	1
New York Times	3,252
News & Observer	893
Newsweek	142
Philadelphia Daily News	3
Philadelphia Inquirer	2
Pittsburgh Post-Gazette	1,916
San Francisco Chronicle	838
Seattle Post-Intelligencer	710
St. Louis Post-Dispatch	1,967
St. Petersburg Times	1,749
New Orleans Times-Picayune	<b>16,538</b>
USA Today	1,442
Wall Street Journal	1,792
Washington Post	3,374
Grand Total	41,964

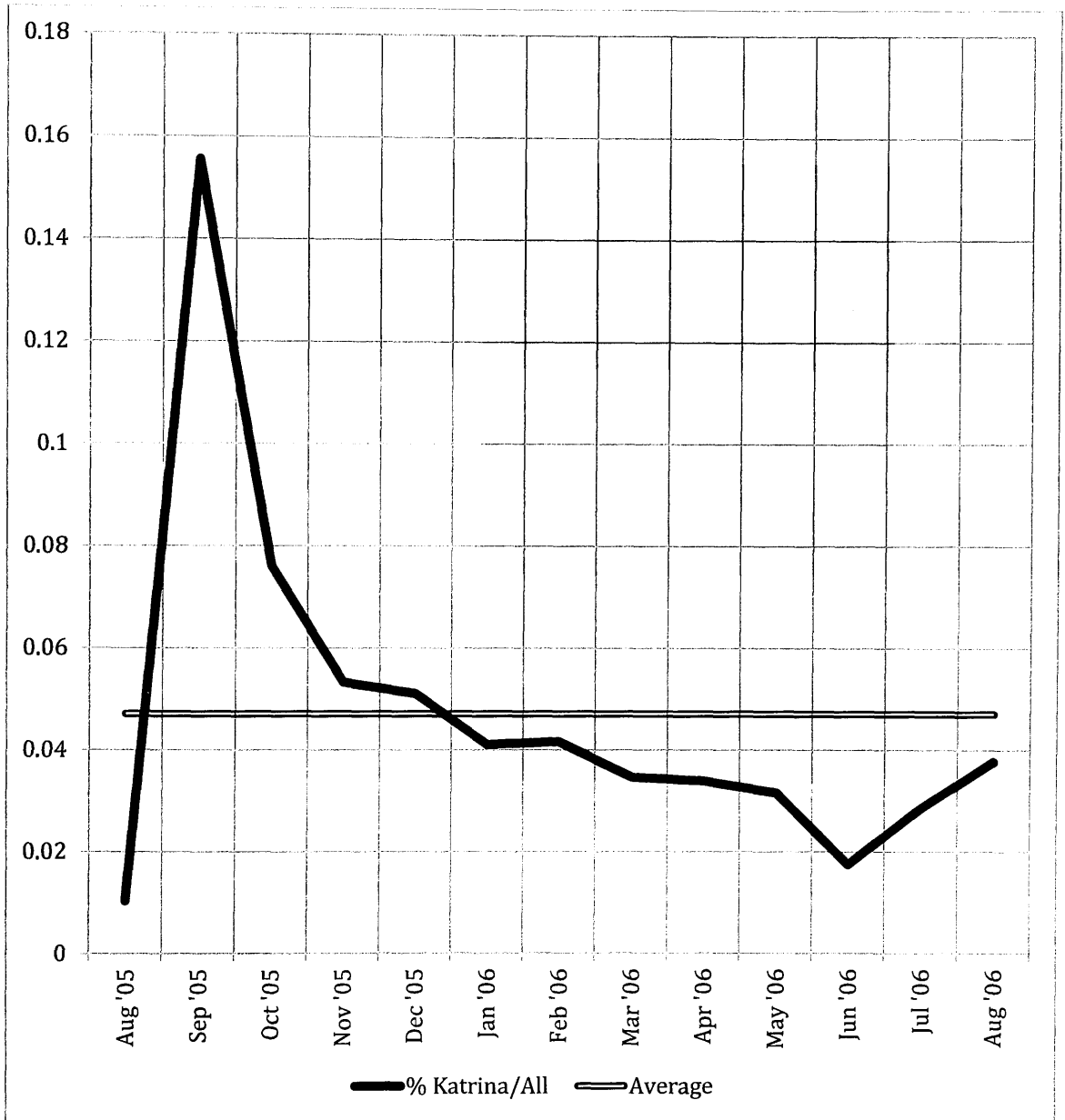
**Table 3-1: Count Hurricane Katrina-related articles by source name**

As a way of determining the proportion of articles about Katrina in terms of all of the news stories published in the same newspapers during the period under consideration, I carried out a search of articles in the Factiva database August 2005 – August 2006 and used the search term *the OR be OR to OR of OR and* as at least one of these highly frequent words will appear in a very large proportion of all articles indexed into *Factiva*. This resulted in 879,665 articles (see Table 3-2 below). Therefore, articles containing reference to Hurricane Katrina accounted for approximately 4.77% of American news articles over the 12-month period.

Month	Katrina articles	All articles	% Katrina/All
August 2005	684	66,196	1.03%
September 2005	10,302	65,685	15.68%
October 2005	5,423	70,547	7.69%
November 2005	3,640	67,272	5.41%
December 2005	3,420	66,339	5.16%
January 2006	2,800	67,427	4.15%
February 2006	2,752	65,614	4.19%
March 2006	2,604	74,196	3.51%
April 2006	2,444	71,638	3.41%
May 2006	2,155	67,811	3.18%
June 2006	1,790	67,375	2.66%
July 2006	1,812	63,162	2.87%
August 2006	2,518	66,403	3.79%
Total	42,344	879,665	4.81%
Average			4.83%

**Table 3-2: Count of Hurricane Katrina-related articles shown as a representative percentage of all articles in the Factiva database, August 2005 – August 2006**

In viewing the number of articles relating to Hurricane Katrina as a percentage of the total number of articles available in *Factiva* for a given month (see Figure 3-1 below), the relative frequency clearly spikes from August 25<sup>th</sup>, and peaks in the month following the storm. While coverage remains above the average level through December 2005, relative frequency decreases steadily until anniversary coverage, and never experiences another spike of this level.



**Figure 3-1: Visual representation of Hurricane Katrina articles in proportion to all Factiva articles by year**

This pattern of an initial spike followed by steady deterioration (or reverse-] distribution) fits neatly with expectation of reports marked by frequency, amplitude, and unexpectedness, detailed by Galtung and Ruge (1965, pp. 53–55) and remarked upon in the previous chapter. Contrast this to the alternate pattern in the AIDS/HIV corpus size distribution, below.

### **3.2.2 AIDS/HIV**

The total size of the AIDS/HIV corpus is 161,144,924 words in 166,576 texts from 26 publications. The publication with the highest frequency of texts containing the search

terms for this corpus is the *New York Times*, comprising 17.44% of the total number. This could reflect the salience of AIDS/HIV reporting in New York City (a location particularly affected during the early years of the epidemic), much like the predominance of Hurricane Katrina reporting in the New Orleans Times-Picayune was an indicator of cultural importance surrounding that event in that place.

Source name	Article count
Atlanta Journal - Constitution	1,846
Barron's	469
Boston Globe	13,669
Chicago Sun-Times	13,019
Christian Science Monitor	1,305
Denver Post	4,915
Forbes	203
Houston Chronicle	18
New Orleans Times-Picayune	8,795
New York Daily News	3,283
New York Post	11
New York Times	<b>29,050</b>
News & Observer	1,175
Newsday	4,091
Newsweek	638
Philadelphia Daily News	143
Philadelphia Inquirer	296
Pittsburgh Post-Gazette	6,136
San Francisco Chronicle	21,339
Seattle Post-Intelligencer	5,838
St. Louis Post-Dispatch	9,570
St. Petersburg Times	1,059
USA Today	10,790
Wall Street Journal	7,034
Washington Post	21,883
WSJ. The Magazine from The Wall Street Journal	1
<b>Grand Total</b>	<b>166,576</b>

Table 3-3: Count of AIDS/HIV-related articles by source name

All articles in the Factiva database 1981–2009 including any of the search terms *the OR be OR to OR of* numbered 19,783,018, with AIDS/HIV representing an average of 0.76% of the articles per year (see Table 3-4 below for figures by year).

Year	AIDS/HIV articles	All articles	% AIDS/All
1981	4	104501	0.00%
1982	23	105,952	0.02%
1983	179	102,030	0.18%
1984	275	192,495	0.14%
1985	2357	287,612	0.82%
1986	3419	358,494	0.95%
1987	9924	720,564	1.38%
1988	9254	801,528	1.15%
1989	8006	743,105	1.08%
1990	8055	756,215	1.07%
1991	9279	742,062	1.25%
1992	12286	784,337	1.57%
1993	9866	771,947	1.28%
1994	8147	808,620	1.01%
1995	7680	775,840	0.99%
1996	7555	779,127	0.97%
1997	7398	838,485	0.88%
1998	7198	930,605	0.77%
1999	6213	932,600	0.67%
2000	5975	880,278	0.68%
2001	6296	864,689	0.73%
2002	4588	849,308	0.54%
2003	6050	905,221	0.67%
2004	4757	864,392	0.55%
2005	4917	804,446	0.61%
2006	5219	814,277	0.64%
2007	5247	826,725	0.63%
2008	3673	768,100	0.48%
2009	3080	669,463	0.46%
Total	166,920	19,783,018	0.84%
Average			0.76%

Table 3-4: Count of AIDS/HIV-related articles shown as a representative percentage of all articles in the Factiva database, 1981 – 2009

Several peaks are represented in discourse related to AIDS/HIV over the span of several decades. The largest of these were in 1992 (1.57% of all articles) and 1987 (1.38%). In all of the years from 1985 to 1998 inclusive, AIDS/HIV represented a higher-than-average proportion of all articles in the Factiva database. In Figure 3-2 below, multiple spikes in relative frequency of AIDS/HIV articles can be seen. This pattern is markedly different from the Hurricane Katrina chart, and is a result of the event's relatively low frequency (unfolding over the span of decades) coupled with heightened consonance and continuity (Galtung & Ruge, 1965).

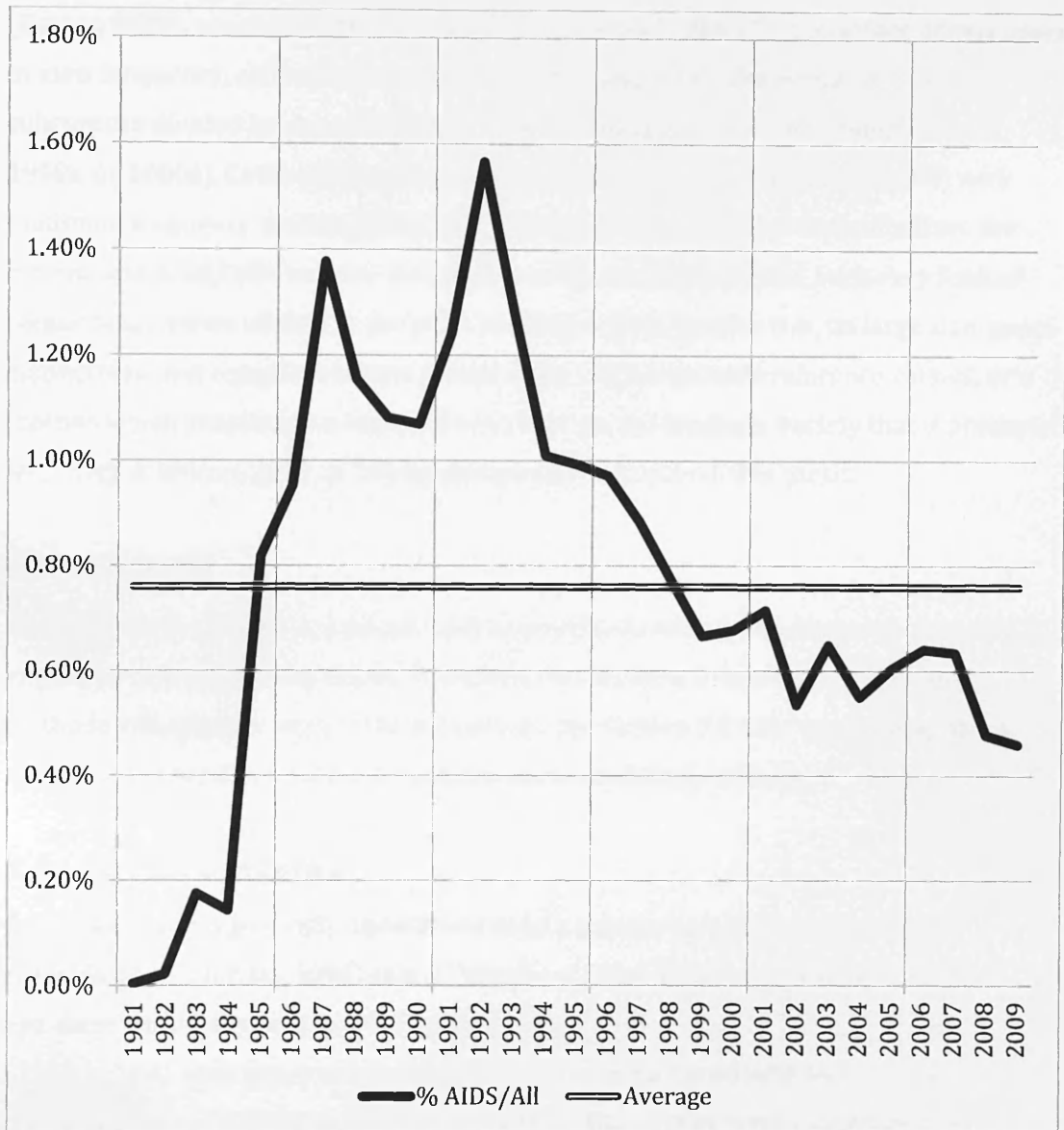


Figure 3-2: Visual representation of AIDS/HIV articles in proportion to all Factiva articles by year

The differing patterns of relative frequency in the Hurricane Katrina and AIDS/HIV corpora correspond to the differing criteria of newsworthiness realised by each event. This finding underlines the necessity of subjecting two case studies to shared methods and parallel analysis, in the hopes of reducing discovering patterns that can be generalised to a greater portion of news discourse.

### 3.2.3 Corpus of Contemporary American English

Due to the specialised nature of the corpora collected for the purpose of this thesis and the limitations in breadth of article access discussed above, at points I will also compare linguistic behaviour of search terms to The Corpus of Contemporary American English (COCA), a 425-million-word corpus of American English sampled from 1990-2011

(Davies, 2009), accessible through an online interface<sup>12</sup>. The COCA interface allows users to view frequency, collocation, and word list information on the entire corpus, or subcorpora divided by genre or decade/year (for my purposes: newspapers, 1980s, 1990s, or 2000s). Collocates may be viewed sorted by mutual information (MI) with minimum frequency thresholds, making this highly compatible with results from the Katrina and AIDS/HIV corpora. However, user rights are restricted, with very limited concordance views offered at the point of query return. Despite this, its large size, genre distinctions, and comparable time period make COCA a suitable reference corpus, or a "corpus which constitutes a standard reference for the language variety that it presents" (McEnery & Wilson, 2001, p. 14) for comparative analysis in this thesis.

### **3.3 Methods**

As described in Chapter 2, a mixed method (corpus-based critical discourse analysis) is employed throughout this thesis. Therefore, two sections follow below, detailing methods falling under each of these headings. See section 3.3.1 for corpus linguistics methods, and section 3.3.2 for critical discourse analytical methods.

#### **3.3.1 Corpus Linguistics**

A 'corpus' today is generally understood to be a principled collection of machine-readable texts (Sinclair, 1996) that is "usually of a size which defies analysis by hand and eye alone within any reasonable timeframe" (McEnery & Hardie, 2012, p. 2). Leech (1992, p. 116) adds that corpora are nearly always structured with *intent*; they "are rarely haphazard collections of textual material: They are generally assembled with particular purposes in mind, and are often assembled to be (informally speaking) *representative* of some language or text type". The corpora here are representative of news texts in a given time period, on the topic of (or at least making reference to) Hurricane Katrina or HIV/AIDS.

Leech (1992, p. 107) describes the main focuses of modern CL as "linguistic performance, rather than competence; linguistic description rather than linguistic universals; quantitative, as well as qualitative models of language; a more empiricist, rather than a rationalist view of scientific inquiry." Many researchers make a distinction between corpus-based versus corpus-driven approaches (Tognini-Bonelli, 2001), or (very simply) the use of a corpus as a large repository of examples by which one might

---

<sup>12</sup> <http://corpus.byu.edu/coca/>

test a hypothesis versus adopting a *tabula rasa* approach to the data, allowing salient features to surface computationally. In the chapters to follow, I oscillate between a corpus-driven approach to identify search terms (i.e. Chapters 4, 5, 7, and 8), and a corpus-based approach to explore naming strategies that I know to be of salience in the corpus, as a member of the society the texts have been produced in (i.e. Chapters 6 and 9). At no point, however, do I abandon intuition altogether; as a critical discourse analyst, it is important to remember that my own identity (see section 1.1) and existing ideologies are central to my manner of interpreting the data, though I endeavour to be as objective and scientific as possible in identifying and describing patterns.

There are four techniques/concepts associated with corpus linguistics that I make consistent use of throughout this thesis. These are analyses of frequency and distribution (section 3.3.1.1), collocation (3.3.1.2), concordance lines (3.3.1.3) and semantic preference and discourse prosody (3.3.1.4).

### **3.3.1.1 Frequency and distribution**

A frequency list is a list of all of the types in a corpus ordered by the frequency of occurrence (Hunston, 2002, p. 68), usually decreasing. Frequency lists are employed in this thesis sorted by part-of-speech, e.g. lists are generated of plural common nouns rather than all types in a given corpus. This allows me to identify the most- and least-frequent naming strategies of a particular part of speech with speed and accuracy. The frequency (or count of occurrences of a given feature or naming strategy) is also often mentioned as a point of comparison, to ascertain relative over- or underuse of one naming strategy versus another.

Here, distribution is taken to mean the equal or unequal dispersion of a feature across different portions of the corpora (Baker, Hardie, & McEnery, 2006, p. 62). Distribution throughout texts, or across publications in a corpus, is sometimes detailed in this thesis to demonstrate: 1) the frequency of evenly spread or concentrated use of one naming strategy compared to another, or 2) the preference for or aversion to certain features by a given newspaper. However, much of the focus of distribution in this thesis is in the form of distribution of collocates across semantic categories.

### **3.3.1.2 Collocation**

Collocation (particularly semantic collocation—see section 3.3.1.4) is the basis of analysis and description of naming strategies in each chapter. Collocation can be defined as a relationship of co-occurrence between a word or phrase and another item (Baker et



al., 2006, pp. 36–38). Calculation of collocates exposes patterns in the text that might not be accessible to intuition, and allows researchers to quantify and describe common and uncommon lexical constructions in a given text. Detailed analysis of an item’s collocates can expose its associations and connotations, and can be useful “in that they help to summarize the most significant relationships between words in a corpus” (Baker, 2006, p. 118). It is through analysis of the most common patterns comprising a word or phrase’s immediate context that we can gain insight into the meaning of the word in use—or, to quote the oft-quoted Firth, “You shall know a word by the company it keeps” (1957, p. 11).

This relationship is calculated statistically: “Words are said to *collocate* with one another if one is more likely to occur in the presence of the other than elsewhere” (McEnery & Hardie, 2012, p. 240). Calculation of collocation involves searching for all instances of the search term or node word and counting the frequency of all other items occurring within a set span. The spans utilized in this thesis are one item to the left to one item to the right (+/-1) in Chapters 4 and 7 where [adjectival] attribution is the main focus, and three items to the left to three items to the right (+/-3) in Chapters 5, 6, 8, and 9.

I sometimes refer to ‘false collocations’, by which I mean words that have been indicated as having a statistically significant relationship due to repetitive proximity, but that are not part of the same semantic or conceptual unit when analysed at the concordance line. For instance, in the string “AIDS victims; suspected carriers”, this instance of ‘suspected’ might be counted as a collocate of ‘victim’, though it modifies ‘carriers’ and occurs on the other side of a boundary. For this thesis, false collocations of the semantic sort are briefly described and usually discarded.

There is a multitude of statistical measures available and in use for the purpose of calculating collocation, and it is widely acknowledged that “different methods of calculating collocation tend to yield different results – the interplay between frequency and saliency needs to be taken into account” (Baker, 2006, p. 119). For this reason, not one but two statistical measures are used in each chapter of this thesis: Mutual Information (MI) and log likelihood (LL).

#### **3.3.1.2.1 Statistical thresholds**

Mutual Information is the main statistical measure of this thesis, and the first method used of identifying collocates. This score is an effect size measure, or a measure of the strength of a collocation, comparing the actual co-occurrence of items within a corpus to the mathematical rate of co-occurrence if the words were ordered entirely randomly. MI

is calculated by dividing the number of instances of a co-occurring word within the span of the node ('Observed') by the number of instances that could be expected in the same span, based on the word's frequency in the entire corpus ('Expected') and converting this to a base-2 logarithm (Barnbrook, 1996; Hunston, 2002). The higher the MI score is, the stronger the collocational relationship between items. This statistical measure favours lower-frequency, open class words that tend to offer more insight into the sociocultural use of language. The minimum MI score threshold of 3 adopted here is that recommended by Hunston (2002).

A second way of calculating collocation is to use a log likelihood (LL) test as a confidence indicator, or a measure of the certainty that one can have that results are not occurring due to chance (Rayson, Berridge, & Francis, 2004). Used on its own, LL tends to favour higher-frequency, closed class words, as the higher frequency overall of these items lend themselves to greater confidence of use. The cut-off for LL value in this thesis is 10.83, or 99.9% confidence, with 0.1% chance of error,  $p < 0.001$ . It was decided to use both MI and LL tests to ensure both strong collocations that were not likely to occur due to change.

Yet in actuality, the combination of a first pass of significance testing with MI scores along with high minimum frequency thresholds rendered the second pass of LL testing redundant. No collocates passing MI score and lower frequency thresholds failed to pass LL thresholds, and in fact, most collocates were far in excess of the minimum 10.83 LL value. I found the best use of LL to be diagnostic of an appropriate 'floor' of minimum frequency of collocation before calculation of MI. However, both MI scores and LL values appear alongside each collocate throughout the thesis for the readers' reference.

### **3.3.1.3 *Concordance line analysis***

Though collocation is employed as the 'first pass' tool in this thesis, the work is ultimately grounded in critical discourse analysis, a theoretical framework that entails close, qualitative analysis. One tool of corpus linguistics that allows simultaneous quantitative and qualitative examination is concordance analysis. A concordance is "a list of all of the occurrences of a particular search term in a corpus, presented within the context in which they occur –usually a few words to the left and right of the search term" (Baker et al., 2006, p. 43).

Though the texts in the Katrina and AIDS/HIV corpus are sometimes (necessarily) short, meanings can be quite dense; journalists often present multiple 'voices' and views of an

issue or constructions of a social group within the same paragraph. To allow these senses to be equally represented, I make use of an extended concordance line format that is somewhat unconventional. Firstly, concordance lines appear in a separate typeface to delimit these from the body of the thesis, and are numbered with sequences restarting in each chapter. The search term, or node word(s), is not centred but boldfaced for ease of location. Collocates are italicized; additional text being discussed as a point of illustration is emphasized with underline. See Figure 3-3 below for an example.

1. Sven Jesper Nyeboe, a fashion designer for many years, died of *complications* of **AIDS** yesterday at his home in Manhattan. (*New York Times*, 31 May 1989)

**Figure 3-3: Sample concordance line format**

Close examination of concordance lines (or individual instances of the search term in extended context) can help the researcher to get a better understanding of what is typical of structures of text in a corpus, and can expose patterns in phraseologies or meanings surrounding search terms that are not made evident during the quantitative stage of analysis. In this thesis, concordance analysis is used primarily for:

- a. quantification of association with other 'deviant' social groups (e.g. *gay* with *homeless*) when these naming strategies might be too infrequent or fall too far from the node to individually surface in collocation analysis;
- b. disambiguation of polysemous words (e.g. *treat* a patient physically with drugs vs. *treat* a person behaviourally with kindness), as in Kennedy (1991), Partington (1998), and Hunston (2002);
- c. exploration of surprising or counterintuitive patterns in collocation figures.

Each stage of analysis in every section is underpinned by concordance analysis, and nearly all discussion sections are accompanied by examples from the corpora.

#### **3.3.1.4 Semantic preference and discourse prosody**

Two additional concepts referenced in this thesis are further functions of collocation: semantic preference and semantic prosody. Semantic preference is considered to be an unusual and significant relationship between a node and a semantic field, whereas semantic prosody is the attitude conveyed through collocation with negative or positive collocates in context (Sinclair, 1998).

More specifically, semantic preference is “the relation between a lemma or word form and a set of semantically related words” (Stubbs, 2001, p. 65), or collocation with a semantic field rather than with a single word. Analysis of semantic preference is useful in a corpus-based critical discourse analysis, as this “denotes aspects of meaning which are independent of speakers” (Baker, 2006, p. 87) but are (assumed to be) shared across the majority of a given culture. By analysing the semantic preferences of the most frequent naming strategies in my corpora, I am able to explore the hidden meanings denoted by their usage. In this thesis, semantic preference is calculated largely through automated categorization of the collocates of a given naming strategy into semantic fields using the USAS semantic tagging system (further discussion to come in section 3.3.3.2).

Semantic prosody—or ‘discourse prosody’ in Stubbs (2001)—relates to the overall positive, negative, or neutral semantic connotation of a search term: “words or phrases are said to have a negative or positive semantic prosody if they typically co-occur with units that have a negative or positive meaning” (McEnery & Hardie, 2012, p. 136). These connotations are not easily recognized through intuition, though once identified may often confirm instinctive reactions to words or phrases, and have the capacity to surprise. This is because “semantic prosodies are mainly engaged at the subconscious level” (Tognini-Bonelli, 2001, p. 114). This expression of evaluation becomes intrinsic to the word or phrase’s meaning.

Deciphering lists of collocates to derive information about semantic preference and semantic prosody necessitates the incorporation of methods of interpretation. In this thesis, I use Critical Discourse Analytical methods and techniques to inform close analysis.

### **3.3.2 Critical Discourse Analysis**

Critical discourse analytical (CDA) methods used here are founded mainly upon: 1) the Discourse Historical Approach as summarised by Reisigl and Wodak (2009); and 2) Theo van Leeuwen’s (van Leeuwen, 2008) descriptions of the methods of exploring construal of social actors from a social perspective.

#### **3.3.2.1 Referential/nominal strategies**

The most frequent ways in which persons and groups of people are named and referred to linguistically in the corpora serve as the structural backbone of the thesis; each analysis chapter in turn deals with differing naming/referential/nominal strategies.

I adopt the objective of **nomination** defined by Reisigl and Wodak (2009, p. 94) as “discursive construction of social actors, objects/phenomena/events and processes/actions”, particularly as belonging to in-groups or out-groups. Reisigl and Wodak (2009, p. 94) note that the word strategy is used for both referential strategies and predicational strategies (below) as these are “more or less intentional plan[s] of practices (including discursive practices) adopted to achieve a particular social, political, psychological or linguistic goal”. There are myriad ways to construct and represent social actors, “such as membership categorization devices, including tropic reference by biological, naturalizing and depersonalizing metaphors and metonymies, as well as by synecdoches in the form of a part standing for the whole (*pars pro toto*) or a whole standing for the part (*totum pro parte*)” (Reisigl & Wodak, 2001, p. 45).

Methods for determining the naming strategy to investigate in each analysis chapter differ slightly, and detailed explanation of each method would be ill suited to this section. Therefore, descriptions of methods can be found at the start of each analysis chapter, directly preceding a description of the frequency of the search terms. Each method, however, is frequency-based, and each naming strategy is explored in the context of its collocates, more specifically collocates indicating predicational strategies.

### **3.3.2.2 Predicational strategies**

The discursive strategy of **predication** qualifies social actors and their actions, assigning (more or less) positive or negative judgments and appraisals. Predicational strategies are realized by the following devices (Reisigl & Wodak, 2009, p. 94):

- stereotypical, evaluative attributions of negative or positive traits (e.g. in the form of adjectives...relative clauses, conjunctive clauses, infinitive clauses and participial clauses or groups)
- explicit predicates or predicative nouns/adjectives/pronouns
- collocations
- explicit comparisons, similes, metaphors and other rhetorical figures (including metonymies, hyperboles...and euphemisms)
- allusions, evocations, presuppositions/implicatures, etc.
- *topoi* (formal or more content-related)

Analysis of predicational strategies reveals the “traits, characteristics, qualities, and features” attributed to social actors (Reisigl & Wodak, 2009, p. 94) in a given context. In

particular, I will be looking at the attributions conveyed through adjectives, collocations, metaphors (see section 3.3.2.5) and *topoi* (3.3.2.3 below).

### 3.3.2.3 *Topoi*

The concept of *topos* (plural: *topoi*) is drawn from classical argumentation theory, and is adopted by Reisigl and Wodak (2009, p. 74) to mean:

“[P]arts of argumentation which belong to the obligatory, either explicit or inferable premises. They are content-related warrants or ‘conclusion rules’ which connect the argument or arguments with the conclusion, the claim. As such, they justify the transition from the argument or arguments to the conclusion” (Kienpointner, 1992, p. 194).

In other words, a “*topos* is both a place of common knowledge (cf. commonplace) and a conclusion rule or warrant that may be invoked in discourse interpretation” (Grue, 2009, p. 309). Reisigl and Wodak specifically connect the argument with a conclusion, e.g. under the *topos* of advantage, “if an action under a specific relevant point of view will be useful, then one should perform it” (Reisigl & Wodak, 2001, p. 75). Wodak (2001) cites a rare authentic example explicitly indicating argumentation; the ‘*topos* of culture’ and ‘*topos* of threat’ are apparent in this appeal: “The greatest damage that one can do to a people is to put the identity, cultural heritage, and the opportunities of its young people negligently at stake. This is why we have introduced the ‘Austria first’ petition. In order to guarantee Austrians their right to a fatherland”(2001, p. 77). Here, the claim (damage to culture) warrants support of ‘Austria first’. However, I find the definition’s inclusion of the presence (however explicit or inferable) of a clear argumentation strategy—or indeed the general concept of an entailing conclusion rule—rather problematic when this concept is applied to large collections of data. Žagar points out that Wodak (as well as other critical discourse analysts, including Richardson, Pelinka, Reisigl, Kryzanowski, and Baker) presented “a fixed list of names of *topoi*, without any explanation of their functioning”, which has the effect of “referring to *topoi* or evoking them or simply mentioning them, which mostly seems to serve the purpose of legitimating the (already) existing discourse and/or text analysis, but gives little analytical- or theoretical-added value in terms of argumentation analysis” (Žagar, 2010, p. 9).

As will be revealed shortly, in my data, explicit conclusions for patterns of categorization of social actors are certainly absent; the conclusion is nearly always inferred. One example from my data (section 4.3.4) of the *topos* of threat/numbers is as follows: “He would later tell her that at one point there weren’t nearly enough officers at the

Superdome to manage 30,000 homeless, hungry people.” (*Boston Globe*, 25 February 2006). While the warrant isn’t explicit in this concordance line, readers with shared common knowledge will likely be able to infer that a lack of police officers, in combination with a large number of people in dire circumstances, is a threat to social order. Later in the text, police officers are reported as having fired shots in the Superdome, presumably having gained reader warrant to act under extreme conditions. Ivie (1980, cited in Richardson, 2004, p. 280) defines *topoi* as “reservoirs of generalized key ideas from which specific statements or arguments can be generated” (emphasis added), and I would like to focus on the former portion of this definition rather than the latter. It is my hypothesis that within a shared cultural context, arguments can be left unsaid, and are implied (and often interpreted and understood) by the mere invocation of *topoi*. Here, I use *topoi* to mean common categorization schemes of social actors through which sociocultural meanings can be inferred. These are “recurrent and repetitive motifs or *leitmotifs*” (Žagar, 2010) or “intellectual themes” more related to the concept of “literary *topoi*” (Curtius, 1990) than Aristotelian or Ciceronian *topoi*.

Richardson and Wodak (2009, p. 49) offer a list of “the most common *topoi* that are used when writing or talking about ‘others’”:

- Usefulness, advantage
- Uselessness, disadvantage
- Definition, name-interpretation
- Danger and threat
- Humanitarianism
- Justice
- Responsibility
- Burdening, weighting
- Finances
- Reality
- Numbers
- Law and right
- History
- Culture
- Abuse

It is these that I will refer to as key places or themes of construction, rather than key warrants leading to explicit conclusions.

#### **3.3.2.4 Passivation/activation**

Another crucial aspect in analysis of representation in this thesis is the exploration of the roles afforded to social actors. Throughout the thesis, I code representation of social actors as agents/actors or patients/goals in respect to specific actions within the

concordance lines. The placement of one group of people as completing or receiving actions represents functional choices on the part of the text producers (Halliday, 1985) and indicates traces of cultural ideology underlying their creation. The role of agent/patient is not always coded according to the grammatical position of the social actors in a given clause, but rather by the sociosemantic representation of the effect of the action (van Leeuwen, 2008). For instance, in the clause “victims are receiving aid from charities”, victims are the agents receiving, but the act of receiving is in itself a passive one (it is in fact the charities who are actively giving). I use two specific terms to describe habitual construal of social actors in one role or the other: “[a]ctivation occurs when social actors are represented as the active, dynamic forces in an activity, *passivation* when they are represented as ‘undergoing’ the activity, or as being ‘at the receiving end of it’” (van Leeuwen, 2008, p. 14). When social actors are passivated, I additionally disambiguate between concordance lines where they are subjected (“treated as objects in the representation”) versus those where they are beneficialized (“form a third party which, positively or negatively, benefits from the action”) (van Leeuwen, 2005, p. 15).

In describing the meaning and consequence of activation or passivation of a given social actor, it is sometimes necessary to take into account the process type in question. For these descriptions I draw upon the classification system from Systemic Functional Grammar (Eggins, 2004; Halliday, 1985), with major process types as follows:

- **Material:** processes of doing, “usually concrete, tangible actions” which result in physical changes (Eggins, 2004, p. 215)
- **Mental:** processes of feeling, thinking, or perceiving (Eggins, 2004, p. 225)
- **Behavioural:** processes that are halfway between material and mental processes, which relate to “action that has to be experienced by a conscious being”, which are “typically processes of physiological and psychological behavior” (Eggins, 2004, p. 233)
- **Verbal:** processes of verbal action, including those which only include “symbolic exchanges of meaning”, e.g. verbiage from a non-sentient being (Eggins, 2004, p. 235)
- **Existential:** processes of existing (Eggins, 2004, p. 238)



- **Relational:** processes of being assigned attributes or relationships between terms, usually on either side of some form of the lemma 'be' (Eggins, 2004, p. 239)

Habitual activation or passivation of a given social actor in the context of a favoured process type represents a functional choice. Cases such as these will be highlighted throughout the thesis.

### **3.3.2.5 Cognitive semantic approach to metaphor**

In this thesis, I often refer to the conceptual or cognitive metaphor theory pioneered and developed by Lakoff and Johnson (M. Johnson, 1987; Lakoff & Johnson, 1980; Lakoff & Turner, 1989; Lakoff, 1979, 1987) to describe the methods by which one conceptual domain is construed, realized, and understood in terms of another. This occurs as a result of a 'mapping' between a source domain (from which we draw information) and a target domain (onto which this information is projected by way of better understanding it). Generally speaking, target domains tend to be more abstract and conceptual, whereas source domains tend to be more concrete or physical. Common uses of metaphor within a given society tell us about the relationships this society might find easily comprehensible or helpful in helping to understand the context around them. Usual pairings of the semantic domains of the source and the target indicate information about the supposed shared attributes or qualities of both concepts.

The role of metaphor "can be a semantic one that is concerned with stretching the resources of the linguistic system to accommodate change in the conceptual system but it can also serve as a stylistic resource for conveying authorial evaluation" (Charteris-Black, 2004, p. 8). The latter use of metaphor is of particular interest in the analysis of the representation of social actors. However, identifying nonliteral meanings in corpora containing tens or hundreds of millions of words is challenging enough, before adding the additional layer of seeking only nonliteral meanings that convey authorial evaluation. Therefore, identification of metaphors relies upon semantic categorization of collocates and investigation into unexpected domains or parts-of-speech. Where necessary, I have performed detailed qualitative analysis—for instance, on each concordance line of *wave*—to disambiguate literal and nonliteral uses of words (a *wave of flood water* overtopping the levee versus a *wave of refugees* displacing to a town). The tools used to accomplish this, and to interface with and analyse the Hurricane Katrina and AIDS/HIV corpora will now be introduced.

### **3.3.3 Tools**

Two web-based corpus analysis tools, both developed at Lancaster University, have been utilized to access and manipulate data: CQPweb (section 3.3.3.1) and USAS (section 3.3.3.2).

#### **3.3.3.1 Corpus Query Processor online (CQPweb) and Constituent Likelihood Automatic Word-tagging System (CLAWS)**

CQPweb, the web-based analysis system used here, was selected over alternate concordancer software packages such as WordSmith or AntConc because it is simply more powerful; a string search over tens (or indeed, hundreds) of millions of words in CQPweb takes a fraction of a second, and users enjoy high flexibility in complex search features, including part-of-speech tag (POSTag) searches (Hardie, 2012). The complex metadata query system offered by CQPweb is also critical in dealing with over 200,000 texts from large variety of dates and publications. Users may, for instance, restrict their queries to publication time or source, and view frequency, distribution, concordance, or collocation data from that selection in isolation. With this functionality, I am able to view the corpus as whole or to spontaneously define subcorpora that might then be compared against one another using the tool's keyword analysis function.

One integral suite of programs included in CQPweb is the Constituent Likelihood Automatic Word-tagging System (CLAWS), which automatically assigns a grammatical tag to each item within an indexed corpus. The version used on both the Katrina and AIDS/HIV corpora is CLAWS7. CLAWS7 tagging is modelled upon: 1) probability data, 2) a lexicon, 3) a suffix list, and 4) an idiom list. In the first instance, "CLAWS assigns potential word-tags using a number of rules based on the ending and orthography of the word, and then uses a Hidden Markov Model method for estimating the most likely word-tag in each context" (Wynne, 1996). CLAWS7 also references a lexicon containing approximately 12,000 words, each of which is associated with between one and six possible (or candidate) tags. In the event that CLAWS encounters an item not contained within the lexicon, a basic morphological analysis is undertaken to attempt to match a list of common or predictable word endings and associated candidate tags to that item. A built-in idiom list accounts for multi-word expressions for the purpose of disambiguation. As will be demonstrated several times through the thesis, CLAWS is not error-free, but has tested at 96-97% accuracy in part-of-speech tagging (Leech, Garside, & Bryant, 1994), which is adequate for my first pass of analysis. The second stage of automated tagging is performed by another tool: USAS.

### 3.3.3.2 UCREL Semantic Annotation System (USAS) and Wmatrix

USAS (*UCREL Semantic Annotation System*) is a framework developed at Lancaster University for automatic semantic tagging of input texts. The tagset consists of 21 major discourse fields (see Figure 3-4 below) and 232 subdivisions arranged in loose correspondence to the Longman Lexicon of Contemporary English (McArthur, 1981). “The semantic tags show semantic fields which group together word senses that are related by virtue of their being connected at some level of generality with the same mental concept. The groups include not only synonyms and antonyms but also hypernyms and hyponyms” (Archer, Wilson, & Rayson, 2002, p. 1). While CQPweb incorporates USAS tagging, users must choose between POS tags and semantic tags (semtags) in search and display functions. Therefore, I maintained POS flexibility on CQPweb and employed another tool developed at Lancaster University for USAS tagging: Wmatrix (Rayson, 2003).

USAS broad category	Description
A	General and abstract terms
B	The body and the individual
C	Arts and crafts
E	Emotion
F	Food and farming
G	Government and public
H	Architecture, housing and the home
I	Money and commerce in industry
K	Entertainment, sports and games
L	Life and living things
M	Movement, location, travel and transport
N	Numbers and measurement
O	Substances, materials, objects and equipment
P	Education
Q	Language and communication
S	Social actions, states and processes
T	Time
W	World and environment
X	Psychological actions, states and processes
Y	Science and technology
Z	Names and grammar

Figure 3-4: Structure of USAS major discourse field distinctions (from Archer et al., 2002, p. 2)

Wmatrix is another ‘fourth generation’ (Rayson, 2003) web-based analysis system also developed at Lancaster University (Rayson, 2008), which likewise utilises the UCREL CLAWS and USAS systems. While it is far less robust than CQPweb, its extended multi-word-expression treatment and ability to toggle between part-of-speech and semantic tags make this more flexible for analysis of frequent semantic fields and the tokens composing them. Wmatrix is capable of tagging words, parts of speech, and semantic

fields through comparison to an internal lexicon of both single-word and multi-word expressions and consideration of surrounding context in uploaded texts. At the time of writing, it is unable to tag entire corpora of the size employed for this research, but is able to offer intelligent tagging for isolated wordlists—such as collections of thousands of statistically significant collocates of the Katrina and AIDS/HIV naming strategies. When tagging items isolated from their extended context, Wmatrix is not able to perform with the precision realised across entire corpora, and instead returns a list of candidate senses for each lexical item ranked by frequency, informed by “frequency-based dictionaries, past tagging experience, and intuition” (Rayson, Archer, Piao, & McEnery, 2004, p. 10).

I uploaded collocate lists for each of the node words into Wmatrix, and manually disambiguated (Baker et al., 2006, p. 58) the semtags the program assigned to every collocate for greater assurance of accuracy. Many tagging errors were simply due to the specialised nature of the corpus (e.g., “new” was tagged *T3: Time: Old, new and young* because Wmatrix has not yet been taught the bi-gram *New Orleans*) other inaccuracies did arise from the lack of context surrounding the collocates. Particularly when Wmatrix was ambiguous in assigning a USAS category, I considered the suggested semtags in the program’s preferred order and made decisions heavily informed by references to the concordance lines in CQPweb. Items with category Z semantic category (names and grammar) are not considered in this thesis, as I am not concerned with individual people or places (proper nouns) or grammatical patterning. However, items tagged Z99 (unmatched) have all been manually disambiguated and re-sorted into more appropriate semantic categories.

Operating within the established framework and categorisation structure allowed for a certain degree of automation in addition to objectivity. Though each collocate was eventually tagged with one of the 232 subdivisions defined by USAS, collocates have been grouped more broadly by POSTag and into their major discourse fields (Figure 3-4 above) for ease of comparison and reference (see Appendix A for the full table).

### **3.4 Stages of analysis and study matching**

The chapters that follow are a matched pair of three ordered analyses. These are as follows:

1. **Chapters 4 and 7:** Initial exploration of the corpora, in response to research question 2. A ‘baseline’ sense of construal of social actors is achieved through

analysis of the most frequent referential strategy in the Katrina corpus (*people*) and AIDS/HIV corpus (*people*, modified to be *people with AIDS/A.I.D.S./HIV/H.I.V./GRID/G.R.I.D.*)

2. **Chapters 5 and 8:** Contrastive analysis of referential strategies, in response to research question 3. In these chapters, the next most-frequent naming strategies are compared and contrasted on the basis of semantic preference. Referential strategies considered in Chapter 5 (Hurricane Katrina) include *resident*, *evacuee*, *victim*, and *survivor*. Chapter 7 (AIDS/HIV) referential strategies include *patient*, *victim*, *sufferer*, and *carrier*.
3. **Chapters 6 and 9:** Diachronic analysis of the construal of social actors, in response to research question 4.

Each analysis is an exploration of referential strategies and associated predicational strategies informed by semantic collocation. Where possible, I will include further discussion on the construal of social actors as related to *topoi*, cognitive metaphors, activation/passivation, comparisons, allusions, and conjunctions.

Micro methods have been developed on an *ad hoc* basis to adapt to the challenges of using two corpora so dissimilar in size, and to respond most adroitly to the research questions posed in Chapter 1. Methods specific to each chapter are introduced in a special methods section at the chapter's start.

Finally, this thesis was designed to gain critical perspectives on the construal of social actors in the wake of a natural disaster and the spread of an epidemic, but equally to experiment with and ultimately establish a robust method for large-scale semantic collocation analysis. As such, methods do not remain static between paired chapters, but develop in Chapters 7-9 in response to perceived limitations in Chapters 4-6. Detailed explanations and justifications for these alterations will be found in the AIDS/HIV analyses chapters, and a final evaluation of the methods established is included in the conclusion Chapter 10.

## ***Chapter 4: Attributes of ‘people’ affected by Katrina***

This chapter is the first of the analysis chapters in the thesis, and the first in a series of three based on the Hurricane Katrina corpus. A brief introduction and orientation to the research question being addressed in this section, as well as a link to the AIDS/HIV companion chapter, follow in section 4.1 below. The methods used for identifying the naming strategy under investigation and for calculating its collocates are detailed in sections 4.2.1 and 4.2.2 respectively. In-depth analyses of these collocates, categorized by USAS broad semantic category, can be found in section 4.3. Results are summarized in section 4.4.

### ***4.1 Introduction***

The main research question addressed in this thesis is how American media discourse negotiates the concepts of risk and blame in ‘Acts of God’ reporting, particularly in relation to attributes of identity such as ethnicity, class, and sexuality. Considering a) the relatively large size of the corpora under investigation; b) the metamorphosing nature of disaster/disease reporting over months/years; and c) the sheer number of naming strategies contained within the texts, it is advantageous to establish a ‘baseline’ of sorts.

This is accomplished by identifying the most common naming strategy in the corpus overall and investigating each of this item’s adjectival collocations in the first chapter of each corpus study, thereby demonstrating the most attributes most commonly associated with social actors broadly in a given corpus before beginning nuanced analysis of less common naming strategies. This chapter (and its corresponding HIV/AIDS chapter 7) addresses the following actionable research question (emphasis given to the most prominent portions):

**What are the most frequent naming strategies employed in Hurricane Katrina and AIDS/HIV reporting, and what can corpus-based discourse analysis of associated attributes, transitivity, and cognitive metaphors tell us about the in- and/or out-group qualities attached to the people that the naming strategies represent?**

The method developed to explore these questions is detailed below. I shall describe the technique of deriving the most frequent naming strategy in the corpus before moving to description of calculation of collocates for this section of the work.

## 4.2 Methods

### 4.2.1 Identification of the naming strategy

All of my texts have been part-of-speech tagged using the CLAWS tagger and CLAWS7 tagset, offering a search flexibility that made identification of high-frequency 'human' words relatively straightforward. A simple query for “(\_NN|\_NN2)” (translating as *common noun, neutral for number OR plural common noun*) outputs a list of all such items in the corpus, which can be ranked by frequency. I have decided to limit my search to common nouns either plural or neutral for number as the object of this thesis is to discover generalized ideologies associated with certain identity groups, not to investigate single instances of one woman, person, child, etc. The list of plural and number neutral common nouns has been manually reviewed for words indicating humanness. The single most frequent naming strategy in a given corpus forms the basis of analysis for this Katrina chapter and AIDS/HIV Chapter 7, whereas a collection of less-frequent (but still prominent) naming strategies are compared and contrasted in Chapters 5 and 8.

With a raw frequency of 68,353 instances over 23,102 texts of the Katrina corpus (or 1860.62 instances per million words), *people* is the most common plural naming strategy in the corpus. Due to its profusion, a case-specific search was necessary to exclude proper names (e.g. “The Young People’s Choir”) and reduce skew in collocation. While this will eliminate cases of sentence-initial *people* where it refers to specific human actors rather than proper names of organisations, this was not observed in the first couple of hundred concordance lines, and has thus accepted as a reasonable bypass. The case-sensitive search for “people” (lowercase) returns 62,934 matches in 22,266 texts (1713.11 instances per million words), to reassure the reader that not many items have been lost with capitalization. The ultimate aim of this chapter is to investigate the manner by which human actors are modified by collocates, and this will be the most successful automated ‘way in’. A list of collocate lemmas was generated using the method detailed below, and will be referred to throughout the chapter.

### 4.2.2 Calculation of collocates

Collocates have been calculated using the statistical significance thresholds standard across the entire thesis:  $MI \geq 3$ ,  $LL \geq 10.83$ . This chapter focuses on the attribution portion of predication strategies, and the window span has therefore been restricted very tightly to +/-1, with collocates limited to the adjectival part of speech category. This

allows a view in to the most frequent attributes of *people* while also downsampling (disregarding items from larger spans and other parts of speech) to a wieldy number of collocates.

In generating collocates, CQPweb offers users the option to retain or discard elements of annotation, including simple tag, lemma, tagged lemma, part-of-speech tag, and semantic tag. By including all of these, the resultant collocation database is highly flexible. I have, for instance, been able to view collocates as part-of-speech tagged lemmata, and can limit results to specific word classes. Due to the large size of the Katrina corpus and the frequency of the node word, even these restrictions (in addition to a minimum frequency of 5 for collocates) allow 599 results. To narrow the collocate count down to a manageable figure, I increased the minimum frequency of collocates to 20, which offers 84 statistically significant collocates for analysis.

Secondary collocation (where the collocates of collocational pairs *adj. + people* are calculated) was investigated where possible, given the frequency of primary collocation. However, given the restricted nature of existing pairs, and the relatively low frequency of co-occurrence at the first tier, statistically significant secondary collocations produced through this method tend to include closed-class words and little else of note. Therefore, I have thinned a (random, reproducible) sample of concordance lines and conducted a manual analysis of each to determine features where needed.

The results of these collocation calculations can be found below in section 4.3, arranged by broad USAS semantic category. A full collocation table (containing information about text distribution and lists of all candidate semtags) can be found in Appendix B.

### **4.3 Analysis**

Collocates of *people* have been grouped below, grouped by USAS broad category and listed in descending order of log likelihood value. Categories containing larger numbers of collocates can be considered more salient to the construction of *people*, namely: I (money and commerce in industry); N (numbers and measurement); and S (social actions, states and processes).



USAS broad category	Significant collocate lemma
A: General and abstract terms	<b>missing</b> [LL=631.119, MI=4.748], <b>ordinary</b> [LL=513.669, MI=5.553], <b>desperate</b> [LL=332.694, MI=4.144], <b>qualified</b> [LL=221.906, MI=4.88], <b>familiar</b> [LL=168.319, MI=3.456], <b>innocent</b> [LL=132.515, MI=4.739], <b>wonderful</b> [LL=91.297, MI=3.03]
B: The body and the individual	<b>disabled</b> [LL=361.326, MI=4.667], <b>sick</b> [LL=138.427, MI=3.511], <b>ill</b> [LL=88.027, MI=3.401]
F: Food and farming	<b>hungry</b> [LL=74.073, MI=3.519]
H: Architecture, housing, and the home	<b>homeless</b> [LL=889.252, MI=4.533]
I: Money and commerce in industry	<b>poor</b> [LL=2396.688, MI=4.567], <b>low-income</b> [LL=501.994, MI=4.873], <b>needy</b> [LL=226.881, MI=4.535], <b>wealthy</b> [LL=151.385, MI=3.851], <b>rich</b> [LL=142.941, MI=3.174], <b>unemployed</b> [LL=96.626, MI=4.227], <b>middle-class</b> [LL=80.188, MI=4.074], <b>working</b> [LL=57.242, MI=3.047]
L: Life and living things	<b>dead</b> [LL=340.703, MI=3.271]
M: Movement, location, travel and transport	<b>displaced</b> [LL=854.464, MI=4.057]
N: Numbers and measurement	<b>many</b> [LL=10524.116, MI=4.283], <b>most</b> [LL=2276.501, MI=3.668], <b>1000</b> [LL=913.794, MI=4.775], <b>100000</b> [LL=837.912, MI=5.636], <b>100</b> [LL=782.617, MI=3.749], <b>200</b> [LL=646.891, MI=4.184], <b>10000</b> [LL=644.037, MI=5.007], <b>180000</b> [LL=612.254, MI=5.882], <b>200000</b> [LL=552.332, MI=5.576], <b>20000</b> [LL=552.332, MI=5.316], <b>1300</b> [LL=537.559, MI=6.144], <b>2500</b> [LL=527.907, MI=4.583], <b>500000</b> [LL=479.329, MI=5.991], <b>500</b> [LL=455.851, MI=3.881], <b>fewer</b> [LL=427.884, MI=4.075], <b>3000</b> [LL=412.094, MI=4.781], <b>300</b> [LL=404.615, MI=4.099], <b>50</b> [LL=390.271, MI=3.083], <b>400</b> [LL=357.679, MI=4.255], <b>30000</b> [LL=349.258, MI=5.133], <b>700</b> [LL=292.481, MI=4.675], <b>1200</b> [LL=286.391, MI=5.282], <b>150000</b> [LL=283.335, MI=5.428], <b>25000</b> [LL=282.846, MI=5.095], <b>1500</b> [LL=266.459, MI=4.64], <b>300000</b> [LL=244.303, MI=5.313], <b>250</b> [LL=236.445, MI=4.371], <b>6000</b> [LL=234.524, MI=4.709], <b>7000</b> [LL=230.341, MI=5.012], <b>1800</b> [LL=222.854, MI=5.744], <b>600</b> [LL=217.665, MI=4.074], <b>5000</b> [LL=212.999, MI=4.279], <b>4000</b> [LL=207.456, MI=4.51], <b>40000</b> [LL=204.769, MI=4.925], <b>12000</b> [LL=195.977, MI=5.136], <b>8000</b> [LL=194.506, MI=4.832], <b>400000</b> [LL=189.619, MI=5.22], <b>14000</b> [LL=183.631, MI=5.566], <b>70000</b> [LL=175.275, MI=5.511], <b>50000</b> [LL=174.927, MI=4.562], <b>80000</b> [LL=163.921, MI=5.382], <b>150</b> [LL=152.707, MI=4.083], <b>2000</b> [LL=148.968, MI=4.912], <b>600000</b> [LL=148.495, MI=5.422], <b>60000</b> [LL=145.779, MI=4.829], <b>75000</b> [LL=135.696, MI=5.73], <b>18000</b> [LL=127.965, MI=5.532], <b>1100</b> [LL=117.618, MI=4.719], <b>120</b> [LL=117.331, MI=3.96], <b>1700</b> [LL=111.288, MI=5.138], <b>250000</b> [LL=99.661, MI=4.749], <b>9000</b> [LL=96.273, MI=4.797], <b>65</b> [LL=69.663, MI=3.047], <b>15000</b> [LL=45.348, MI=5.102]

S: Social actions, states and processes	<b>black</b> [LL=1994.959, MI=4.112], <b>american</b> [LL=1825.694, MI=3.535], <b>white</b> [LL=454.07, MI=3.205], <b>stranded</b> [LL=287.235, MI=4.668], <b>encouraging</b> [LL=150.466, MI=4.031], <b>iraqi</b> [LL=101.934, MI=3.521], <b>resilient</b> [LL=90.751, MI=4.3], <b>generous</b> [LL=64.439, MI=3.119], <b>gay</b> [LL=54.185, MI=3.027]
T: Time	<b>young</b> [LL=3192.527, MI=4.759], <b>elderly</b> [LL=507.24, MI=4.157]
X: Psychological actions, states and processes	<b>talented</b> [LL=74.989, MI=3.766], <b>smart</b> [LL=73.41, MI=3.154]

**Table 4-1: The significant adjectival collocate lemmas of *people* grouped by USAS semantic category**

A brief glance at the most statistically significant collocates is telling; the items with the highest log-likelihood values include *young*, *poor*, *black*, and *homeless*, all ‘risk’ groups salient in Hurricane Katrina reporting for various reasons, to be explored below.

#### **4.3.1 USAS category A: general and abstract terms**

The first category of collocates to be investigated is USAS A: General and abstract terms. Due in part to the wide remit of this category, collocates of *people* in the Katrina corpus grouped under this heading are both numerous and diverse. Adjectival collocates vary from expected descriptions of *people*’s unknown whereabouts (*missing*) in the wake of a natural disaster, to perhaps less standard reports of their mental/emotional states (*desperate*) and appraisals of their characters or places within society (*ordinary*, *innocent*, *wonderful*).

##### **A category adjectival collocates of *people***

**missing** [LL=631.119, MI=4.748], **ordinary** [LL=513.669, MI=5.553], **desperate** [LL=332.694, MI=4.144], **qualified** [LL=221.906, MI=4.88], **familiar** [LL=168.319, MI=3.456], **innocent** [LL=132.515, MI=4.739], **wonderful** [LL=91.297, MI=3.03]

**Table 4-2: USAS A category collocates of *people*, listed in order of descending log likelihood value**

*Missing people* is hardly a strange collocation in a corpus of disaster reporting; it occurs 133 times, most often (84 cases) in a formulaic description of “Find Family National Call Center...the official 24-hour coordination center for information on storm-related **missing people**.” In the remaining 49 cases, concerned with quantifying the numbers of *missing people* and the lengths to which government agencies and loved ones are going to locate them, only one case stands apart:

1. State Medical Examiner Louis Cataldie has estimated that dozens, perhaps hundreds, of bodies will never be found. They disappeared with the floodwaters, lost in the marshes or at the bottom of the Gulf of Mexico. Other **missing people** may be exploiting that status, eager to start over because of past run-ins with the law or problems with family. (*Pittsburgh Post-Gazette*, 27 August 2006)

**Figure 4-1: Sample concordance line of *missing people***

This is the only case where a *missing* person is given agency and negatively judged as committing malicious behaviour in the wake of the storm. Indeed, another collocates—*qualified*—refers simply to the appointment of *qualified people* (not people affected by Katrina, or PAKs) to assist in rebuilding and city re-planning efforts. *Familiar* collocates with *people* in a whole range of contexts with no clear patterns, for instance people *familiar* with talks, with areas, with governmental departments. It is rather with the other collocates that we find more frequent appraisals of PAKs.

Of those remaining (*ordinary, desperate, qualified, familiar, innocent, wonderful*), *ordinary* certainly seems to be the least value-laden. The top collocates of *people* ranked by MI in the COHA corpus are *circumstances, citizens, conversation, routine, and rules*. The appearance of *citizens* in the list of *people's* collocates in COCA is a possible parallel to *people* in the Katrina corpus, but concordance analysis shows a great deal of difference in their treatment. In general American English, the *ordinary citizen* goes about their usual tasks, whereas in a corpus of disaster reporting, *ordinary people* are marked, having recently undergone extraordinary events. Occurrence of this collocation generally falls under two categories: 1) failure of the government or governmental officials to care for the 'ordinary people' affected by Katrina, or 2) heroism or stoicism on the parts of 'ordinary' PAKs operating under extreme circumstances. In line 2 below, the governor comes to the realization that *ordinary people* rely upon the aid of the government, which, in my opinion, should perhaps not be a "new understanding". In line 3, the administration is accused of not caring about *ordinary people*—the theme of care will be revisited shortly in this chapter. *Ordinary people* join corporations and church groups in line 4, "giving generously" to charities in the wake of Katrina; whereas in line 5, *ordinary people* who put themselves into direct danger to save others are celebrated publically. Line 3 shows a minority pattern in these concordance lines—the marking of *ordinary people* not as direct victims of an absentee state service, or as an unlikely hero, but as a marked sufferer failed by their government. It is interesting that *ordinary* has been used in line 3 at all, as this appears to close the distance between the average reader and the distant other. Generally speaking, the occurrence of *ordinary* is unnecessary, and a case of marking the typical. In investigating these cases, it instead

appears that *ordinary people* are being marked out as receiving atypical treatment, falling foul of normal circumstances, or behaving with unusual intrepidity. Therefore, *ordinary* is likely included to contrast with these counterpoints within the sentence or story.

2. “The governor may be showing a new understanding that there are times when **ordinary people need their government’s help** to cope with problems that they can not solve alone,” said Noah Berger, executive director of the Massachusetts Budget and Policy Center, which monitors the effects of fiscal policy on people’s lives. (*Boston Globe*, 8 September 2005)
3. The revealing comments by members of the Bush team and the slow response to the destruction of New Orleans show that the administration doesn’t care about ordinary people. (*Atlanta Journal-Constitution*, 18 September 2005)
4. By moving back and forth between panoramic views and up-close zoom shots of individuals, Mr. Brinkley is able to convey both the larger arc of the tragedy that engulfed the Gulf Coast and the intimate fallout that that catastrophe had on the lives of ordinary people -- many of them black and most of them poor and lacking the resources to evacuate on their own. (*New York Times*, 16 May 2006)
5. Corporations are giving generously, as are thousands of church groups and **ordinary people**. (*Atlanta Journal-Constitution*, 14 September 2005)
6. After the governor introduced **ordinary people who rescued tens of thousands of hurricane victims via boat and school bus**, they leapt to their feet once more. (*New Orleans Times-Picayune*, 18 September 2005)

Figure 4-2: Sample concordance lines of the collocation of *ordinary* and *people*

While *innocent people* collocates with a good margin above the minimum frequency (28 instances), it is a small minority of these that refer to PAKs; all of them are reflected in the lines below. Like *ordinary* above, it seems unnecessary to mark PAKs as *innocent* in most scenarios, and the occurrence of this collocation may indicate that there is a type of natural disaster victim who is not blameless. In the first three sample concordance lines provided below, *innocent* PAKs are compared to non-innocents, including Guantanamo detainees encountered by a National Guard working in post-Katrina New Orleans, as well as contrasted to them in lines 8 and 9, when they are set apart from other inhabitants of the Astrodome<sup>13</sup>, who are targets for police activity and considered “murderers and rapists”. In the second half of all concordance lines arising from the collocation of *innocent* with *people*, the association seems to function in a related way to the occurrence of *ordinary* above. The marked use (suffering of the clearly non-criminal, non-deserving, sudden other) tightens the distance between subject and reader, creating

---

<sup>13</sup> The Houston Astrodome and New Orleans Superdome were repurposed as ‘shelters of last resort’ where little or no food or water, and no utilities, were made available to evacuees as a matter of emergency survival protection. Over 46,000 PAKs were housed between these two sites in the weeks after the storm, and conditions were unsanitary and dangerous (Barnshaw & Trainor, 2010)

an emotional association in line with other lexical items in these lines, including “dignity” and “desolate”. The *innocent people* are evoked in line 12 to justify breaking disaster preparedness protocol. While this collocation is quite rare, and very infrequently refers to PAKs, its use does mark PAKs as somehow more *innocent* than other types of victims (perhaps the long-term homeless, penniless, or desolate).

7. “With Guantanamo, at least there's some reason people are down there. These are all **innocent people**.” (*Wall Street Journal*, 2 September 2005)
8. Dozens of officers tried patrolling outside around the convention center, but, according to Lt. Melvin Howard, the crowds and darkness made it difficult and dangerous to work inside. Police could not use flashlights without giving away their position and becoming possible targets, Howard said. Nor could they open fire, if confronted, without the risk of killing **innocent people**. (*Washington Post*, 15 September 2005)
9. In their view, the whole thing, beginning with the levee break and ending with the cramming of thousands of **innocent people** into what they were sure were death chambers with murderers and rapists, was a setup. (*New York Times*, 9 October 2005)
10. Living in a shelter or in the Astrodome in Houston is not a life with dignity, and it may be months before the many **innocent people** who lost their homes to Katrina can figure out what they will do with their lives. (*New York Times*, 2 September 2005)
11. I hope their nightmares are a continual replay of this, the greatest catastrophe to hit our United States. Think of the **innocent people** - homeless, penniless, desolate, sick at heart and mind, and in dire need. (*St. Petersburg Times*, 13 September 2005)
12. Three days after Brown told senators that he went straight to the White House and did not call Chertoff the day of Katrina’s landfall because it would “have wasted my time,” Chertoff said: “There is no place for a lone ranger in emergency response.” He added that the cost “is visited on too many **innocent people**.” (*Washington Post*, 14 February 2006)

**Figure 4-3: Sample concordance lines of the collocation of *innocent* and *people***

In contrast to this discourse of innocence are the discourses of criminality surrounding another collocate: *desperate*. Concordance lines are divided here. In many cases, *desperate people* are construed as being capable of anything (as in line 13 below), even when in other circumstances (e.g. being more wealthy, better educated, and living in higher-class neighbourhoods than the majority of PAKs) they would not be ‘nasty’ (line 14). The sight of so many starving *desperate people* leads to one journalist likening the New Orleans area to a Third World country—a parallelism that will be revisited at length in Chapter 6. In the other half of concordance lines, however, the depiction of *desperate people* as looters or “potential thieves and thugs” is problematized, and the attributed group is reframed as victim rather than potential wrongdoer (lines 16 and 17 below).

13. These are desperate people and *desperate people can do desperate things*. (*Atlanta Journal-Constitution*, 2 September 2005)
14. "You take the finest group of suburbanites in America, with great wealth and education, and you put them in a situation without food or water for 24 hours," he said. "I don't care who you are – you're going to get nasty. *Desperate people are dangerous people*." (*New Orleans Times-Picayune*, 19 September 2005)
15. The scenes of starving, *desperate people pleading for help* while uncontrolled gangs of thugs stormed hospitals and shot at rescue workers were not from some faraway *Third World country* with limited resources and weak political infrastructure, but from the United States, the richest and most powerful nation in the world. (*Boston Globe*, 18 September 2005)
16. When I look at the images of this "looting" more closely, I see *desperate people scavenging whatever food and water they can find* because they are hungry, and no one has come to help them. (*Boston Globe*, 10 September 2005)
17. New Orleans officials have criticized the blockade as a callous move to cut off *desperate people perceived as potential thieves and thugs* looking to terrorize West Bank communities. (*New Orleans Times-Picayune*, 26 February 2006)

Figure 4-4: Sample concordance lines of the collocation of *desperate* and *people*.

*Wonderful* is the most explicitly positive collocate in this group, approvingly evaluating *people* 37 times in the -1 position. However, concordance analysis shows that *wonderful people* are not PAKs, but are more often those helping PAKs, namely from other cities of settlement (e.g. Atlanta in line 18 below, or Chicago in line 19). This shows that while the voices incorporated into reportage from PAKs (line 18) or local journalists (line 19) are appreciative of care received from others, the positive semantic prosody indicated by this collocation in fact does not reflect upon PAKs themselves.

18. New home, business, hope in Atlanta. I just want to thank the *wonderful people of Atlanta* for their help after what my family went through with Hurricane Katrina. (*Atlanta Journal-Constitution*, 30 August 2006)
19. The *wonderful people of Chicago* are among those who have touched us with their compassion. (*New Orleans Times-Picayune*, 29 September 2005)

Figure 4-5: Sample concordance lines of the collocation of *wonderful* and *people*.

A very similar pattern of focus on the areas around PAKs is observed in the collocates of the next category: USAS category X.

### 4.3.2 USAS category X: psychological actions, states and processes

Though the two collocates categorized in USAS X—*talented* and *smart*—do carry a semantic sense of psychological states, these are also very closely related to judgments of one's character or worth, as in USAS A collocates *wonderful* and *qualified*. For this reason, category X is discussed out of alphabetical order and can be found below.

### X category adjectival collocates of *people*

**talented** [LL=74.989, MI=3.766], **smart** [LL=73.41, MI=3.154]

Table 4-3: USAS X category collocates of *people*, listed in order of descending log likelihood value

Once more, concordance lines arising from the collocation between *people* and both *talented* and *smart* tend to describe not PAKs, but those around them, e.g. employees of the Army Corps of Engineers, politicians, and rebuilding experts. People affected by Katrina are not qualified with these words, and in some cases, people from the Gulf region are explicitly excluded.

20. "Each time one hits, it's a major disruption and a loss of time," said President Tim Hinkley. "At this point, it's really tough for us to attract talented people to Biloxi." (*St. Louis Post-Dispatch*, 30 November 2005)
21. "There may be *smarter people* than me, but they ain't in Louisiana," Long once said. (*New Orleans Times-Picayune*, 16 October 2005)

Figure 4-6: Sample concordance lines of *talented* and *smart* collocating with *people*

In line 20 above, the act of attracting people to Biloxi indicates that there are few (or none) resident there. Therefore, the joking bit of bravado in line 21 strikes close to home; the stereotype of the "dumb south" is played out in jest, but also in earnest reporting. In light of the next category, those who are resident in the south are construed as being very much in need of assistance in rebuilding infrastructure.

### 4.3.3 USAS B and L categories: *disabled, sick, ill and dead*

Given that each text within the corpus contains at least one reference to Hurricane Katrina, it is unsurprising that collocates of *people* might belong to USAS categories B (the body and the individual) and L (life and living things); impact on health, wellness, and livelihood of residents is certainly a topic of utmost concern during and after a natural disaster. However, the exact collocates appearing in these categories—and their uses—are somewhat less predictable.

### B category adjectival collocates of *people*

**disabled** [LL=361.326, MI=4.667], **sick** [LL=138.427, MI=3.511], **ill** [LL=88.027, MI=3.401]

Table 4-4: USAS B category collocates of *people*, listed in order of descending log likelihood value

The USAS B category collocate with the highest log likelihood score—*disabled*—does not indicate an identity created as a by-product of the storm. This collocate appears only in the -1 position in all 78 of its occurrences, eliminating the possibility of constructions such as '*people disabled* by the storm'. Instead, this collocate appears most frequently as the ways in which *disabled people* were evacuated, cared for, or (sometimes) neglected

by others. Of 78 concordance lines, 73 can be analysed for transitivity—the remaining five feature existential constructions such as the example in line 22 below. Of those remaining, *disabled people* are only the agents in 16 (21.92%) of concordance lines, mostly undertaking non-material processes such as living, trusting, relying, and needing help. They are sometimes depicted in mass counts, metaphorically indicating number as in the example of ‘pockets’ in line 22 below, or the more threatening process of several hundred ‘crowding’ in protest, as in line 23. In 36 concordance lines, or 49.32% of cases, *disabled people* are direct objects of another agent, sometimes appearing as the patients in agentless passive constructions as in line 24 below, and in the remaining 21 (28.77%), they are beneficiaries of an action, as in the services provided in line 25. Over half the time (40, 54.79%), *disabled people* are linked together in lists of ‘others’, including senior citizens, minorities, women, and other vulnerable social groups. Examples are included in lines 24 and 25, and indicate that certain identities are linked in post-storm discourses of helplessness and need.

22. “The problem is that if a disaster like Katrina happened in the Bay Area, I think it would be even worse than it was in New Orleans because there are so many pockets of **disabled people** here,” said Craig Gates, a former San Francisco police officer who lost the use of his legs in a motorcycle accident. ” (*San Francisco Chronicle*, 7 October 2005)
23. Several hundred **disabled people**, many in wheelchairs, crowded into the hallways and private suites of House and Senate office buildings yesterday to protest proposals to slow federal spending for Medicaid, the health program for the poor. (*Pittsburgh Post-Gazette*, 20 September 2005)
24. Elderly and disabled people were hit hard in the aftermath of Hurricanes Katrina and Rita. (*USA Today*, 30 December 2005)
25. Health officials also are struggling to prepare to provide services to homeless people, non-English-speaking immigrants, **disabled people**, the elderly and other vulnerable populations. (*USA Today*, 21 February 2006)

**Figure 4-7: Sample concordance lines of the collocation of *disabled* and *people***

*Sick* and *ill* are near-synonyms, and occur with somewhat similar statistical significance scores. For the most part, these are both construed in expected ways—without agency, social power, and etcetera.

*Sick* collocates with *people* in the given spans 45 times, 36 of which are direct attributive adjectives in non-existential clauses. In 32 cases (88.89% of concordance lines) *sick people* are the patients or beneficiaries of actions of others. They are the agents in only four clauses, one of which is mental (‘see us come back’), one behavioural (‘died’) and two material (‘began to return’ and ‘try to cross’) – both of which are marked in some way, as the start/intent of the process or attempt at one, without the guarantee of



completing it. This, in combination with lack of agency overall, indicates absence of power in construal of *sick people*, which is expected.

There are, however, a handful of unexpected results, illustrated by examples below. *Sick people* are also construed as negatively impacting their environs, and constructed through the scope of their emotional effect on well people. For instance, they are the contaminants in lines 26 and 27 below; in 26, the cause of bacteria leading to a building being shut down, and in line 27, metaphorically described as rubbish or storm debris, where they 'litter' an open garage. In concordance line 28 below, it is the *sick people* who are receiving no help, but it is the doctor who is foregrounded as enduring the 'misery' of watching this. Even as patients, *sick people* are framed through the experiences of their carers, and this feature—in combination with their lack of agency—means that it is quite difficult to explore their experiences in the corpus. This will be revisited in the AIDS/HIV portion of the thesis, particularly in Chapters 8 and 9.

26. Because of the bacteria which comes from years of **sick people** and the flooding, "it would be very difficult to put anybody in that building for a long time with an impaired immune system," Finn said. (*New Orleans Times-Picayune*, 1 September 2006)

27. The open garage, transformed into a makeshift ICU, was littered with **sick people**. (*Atlanta Journal – Constitution*, 16 May 2005)

28. Like so many, Dr. Gregory Henderson endured days of filthy flood water, close-by gunfire and the misery of watching **sick people** get no help. (*News & Observer*, 4 September 2005)

**Figure 4-8: Sample concordance lines of the collocation of *sick* and *people***

The near-synonym of *sick*, *ill*, is extremely similar, if a bit more active. In 29 qualified concordance lines resulting from this collocation, six (20.69%) place the *ill people* in the agent position, about double as often proportionally as with *sick people*. The actions associated with them are much the same, however, with material processes made up of movement ('made it to the metro area', 'relocated'), behavioural indicating the cessation of life ('died'), and relational processes ('unable to abide by the mandatory evacuation rule', 'are least able to fend for themselves', 'lack the contacts or the money for treatment') indicating several deficits in ability to gain power in society at any point.

Given that 'die' has appeared in concordance lines as behavioural processes undertaken by both *ill* and *sick* people, it should come as little surprise that *dead* also collocates with *people*. Though this is categorized as USAS L: life and living things, I will discuss it here with B: the body and the individual given this parallel.

## L category adjectival collocates of *people*

dead [LL=340.703, MI=3.271]

Table 4-5: The sole collocate of *people* from the USAS L category

Though the Katrina corpus does have a relatively restricted sampling frame, the genre of the texts contained does affect results. For instance, the appearance of the phrase *dead people* or *people dead* in a news text (even one that mentions 'Katrina' and a storm word elsewhere) is not necessarily describing storm casualties. Therefore, I have manually reviewed the 123 concordance lines and categorised only those 70 that directly pertain to Hurricane Katrina, disregarding the others. Proportions of pre- and post-modification are roughly equal, with 40 occurrences of *dead people* and 30 of *people dead*, indicating a slight preference for thematising the attribute over the actual being. What is probably more interesting is that only 18 of these 70 occurrences are what I class as 'reporting' cases of *people dead*, for instance, quantifying deaths (see sample concordance line 29 below) or reporting military missions (line 30).

29. Katrina left more than 1,700 **people dead** in Louisiana and Mississippi. (*USA Today*, 30 August 2006)
30. "But most of the houses we are looking at are empty," Oregon National Guard Staff Sgt. James Lindseth, 33, said as his platoon, inspecting for **people dead** or alive, worked its way through dank and broken homes that had been in the water a few days ago. (*Washington Post*, 25 September 2005)

Figure 4-9: Sample concordance lines of the collocation of *dead* and *people*

For the most part, *dead* collocates with *people* in *vox populi* statements from witnesses or PAKs, as in sample lines 31 and 32. These are used to liken to situation in the post-Katrina Gulf coast to a 'Third-World environment', whether from the perspective of quoted and attributed individuals or from the editorial voice (line 33). *Dead people* is extremely emotive, and in conjunction with Hurricane Katrina reporting, is used far less frequently to report statistics or search types, but to convey (often horrific-sounding) conditions to the reader.

31. An evacuee identified as Joyce West added, "There were **people dead on the highway, there were people just dropping from exertion, from the heat. No water, no food.**" (*Boston Globe*, 5 September 2005)
32. "There are fires burning in the distance, there are gunshots, there are naked babies who need formula and there's no power. I'm in a Third-World environment, it smells like dead people and cops are driving by and not trying to help anybody. Nothing made sense." (*Atlanta Journal-Constitution*, 29 November 2005)

33. But the pictures of dead people left uncollected on the streets, armed looters ransacking shops, survivors desperate to be rescued, racial divisions -- these were truly out of sync with what we'd imagined the land of the free to be, even if we had encountered homelessness and violence on visits there. (*New Orleans Times-Picayune*, 1 August 2006)

**Figure 4-10: Sample concordance lines of the collocation of *dead* and *people***

In a way, this is the culmination of patterns seen in the B category collocates. *Disabled, sick* and *ill* people are construed as objects to be helped or moved, or as beneficiaries of aid, but ultimately helpless and somewhat harmful to the immediate well people and to the larger society. *Dead people* are mostly discussed in terms of the range of the psychological damage done to those still living, rather than the causes or counts of deaths after Katrina. Again, this difficulty in investigating those most affected by the circumstances framing the studies will be revisited later in this thesis.

#### **4.3.4 Collocates from USAS F and H categories: hungry and homeless**

Containing just one collocate each, the USAS categories F and H have been combined here not just to avoid superfluous subsections, but also due to shared concordance lines. Several times throughout analysis, *hungry* and *homeless* appear alongside one another in concordance lines beside *people*, as will be demonstrated in part with illustrative examples below. I begin with the less-frequent collocate: *hungry*.

##### **F category adjectival collocates of *people***

**hungry** [LL=74.073, MI=3.519]

**Table 4-6: The sole collocate of *people* from the USAS F category.**

The collocation between *hungry* and *people* is a relatively infrequent one, just overtopping the minimum frequency of 20 at 24 instances. For the most part, it is PAKs who are *hungry*, and frequent discourses do recur in concordance lines featuring this collocation, including the *topos* of numbers and threat/danger (see line 34), where *homeless, hungry people* are enumerated and construed as a threat to social order, unable to be “managed” by police. The disaster state and ensuing shortage of food is construed as somehow being un-American, as evinced in line 35 below. “The contract of American citizenship” does not entail aid (of the financial, moral, or sustenance sorts), an indictment symptomatic of the risk society. It is only in a sparse three concordance lines that a parallel is drawn between Katrina’s temporarily *hungry* and the long-term *hungry* (line 36), who are also American, but do not express the same outrage at lack of contractual fulfilment in the corpus.

34. He would later tell her that at one point there weren't nearly enough officers at the Superdome to manage 30,000 homeless, hungry people. When she finally did get through to him, he thinking it was headquarters on the line shouted: "They're shooting! They're shooting!" (*Boston Globe*, 25 February 2006)
35. What has not been noticed is that the people with the most articulate understanding of what the contract of American citizenship entails were the poor, abandoned, *hungry people* huddled in the stinking darkness of the New Orleans convention center. "We are American," a woman at the convention center proclaimed on television. (*New York Times*, 25 September 2005)
36. Ewing, of the Maryland Food Bank, said the post-Katrina images of desperation reminded him of the less-publicized plight that hungry people face daily. (*Washington Post*, 9 March 2006)

**Figure 4-11: Sample concordance lines of the collocation of *hungry* and *people***

A similar contrast between 'long-term' and 'sudden' social states comes in the form of the next point of analysis. The only adjective from the USAS H category (Architecture, housing and the home) that collocates with *people* in the Katrina corpus is *homeless*. Given Katrina's physical destruction of properties, both commercial and residential, the topic of homelessness was sure to arise.

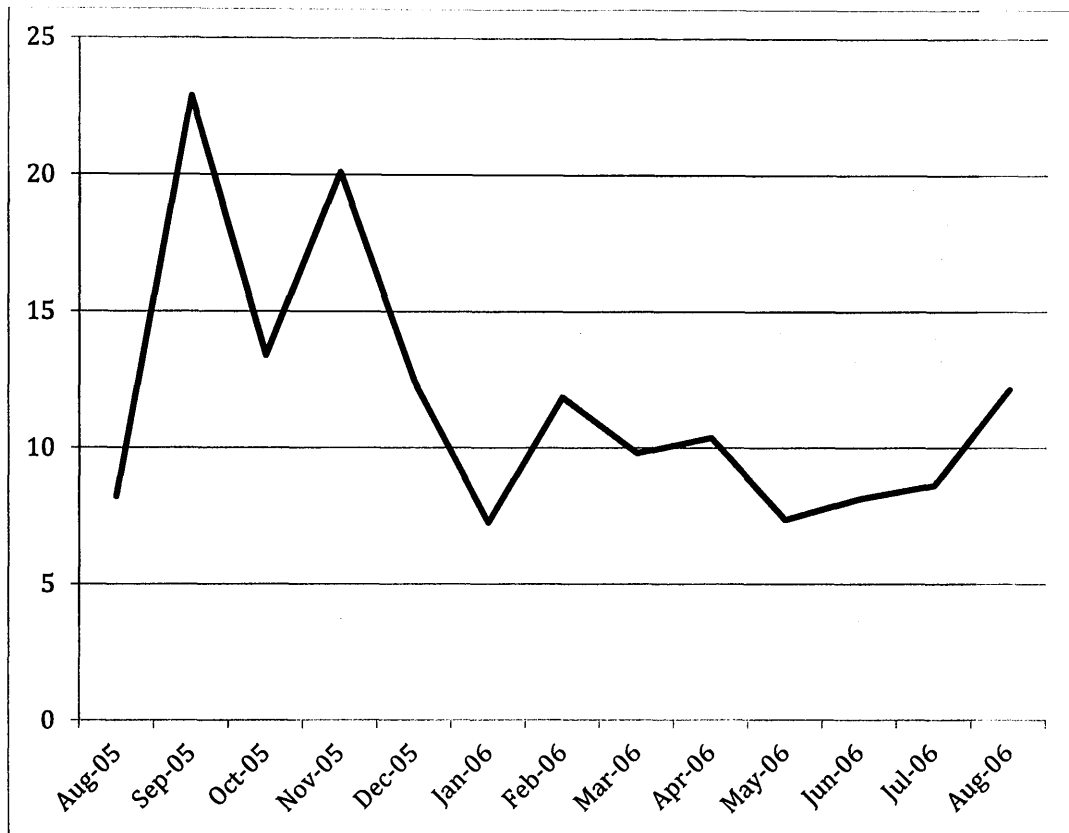
<b>H category adjectival collocates of <i>people</i></b>
<b>homeless</b> [LL=889.252, MI=4.533]

**Table 4-7: The sole collocate of *people* from the USAS H category**

However, homelessness has been a grave social problem in America far predating Hurricane Katrina, and of course extending far beyond it. In the first chapter in a volume on homelessness in America, Cunningham and Henry summarise homelessness statistics in the year of Hurricane Katrina:

Using data gathered by local communities, the Homelessness Research Institute at the National Alliance to End Homelessness found that approximately 744,313 people experienced homelessness [in America] in January 2005. This study revealed how many people were homeless at a point in time; the reality is that two to three times as many people (2.5 to 3.2 million) experience homelessness over the course of a year. (2008, p. 2)

There was, therefore, a possibility that *homeless* collocates with *people* in the Katrina corpus simply because this is a (frequent) identity in America in general, regardless of the climate of reporting in the period studied. However, in the Katrina corpus, the subject of homelessness does not follow a static pattern that would correspond to stasis of the social problem (see Figure 4-12 below), and in fact roughly mirrors the salience of Katrina reporting in general, when considered in frequency per million. This discounts the possibility that *homeless* collocates with *people* irrespective of Hurricane Katrina.



**Figure 4-12: Instances of *homeless* in each month of the Katrina corpus, standardised to frequency per million**

This collocation is much more frequent than that between *hungry* and *people*, occurring 200 times. In these lines, three dominant patterns emerge, with roughly equal shares. The first is the marked identity of the “suddenly” *homeless*, made so by the storm itself (see line 35 below). The sudden adjustment to this status is often construed as psychologically disorienting (line 36), often leading to desperation (see the section above). This is in contrast to the long-term *homeless*, who are set apart through the use of adjectives such as “genuine” (37) and “actual” (38), as though the temporarily homeless are not genuinely or actually homeless, or cannot claim ownership of this identity. The “sudden *homeless*” are often described in the context of water metaphors, as in lines 41 and 42 below, where they are a “wave” to be “absorbed”, drawing in turn upon threat and bursting metaphors. This threat is compounded more explicitly in other concordance lines, for instance where an “influx of newly homeless people turned Baton Rouge into a nightmare town” on the merit of asking residents where they could stay (line 43). The “sudden homeless” are contrasted frequently to the long-term (or “genuine”) homeless, even when these two identities blend. In concordance line 44 below, the speaker is herself (temporarily) homeless, but describes herself as being “like” *homeless* people, rather than adopting this identity or moniker.

37. The group that so heroically coordinated the entire relief effort, including housing 157,000 suddenly homeless people and employing tens of thousands of others, was the Chicago Relief and Aid Society. (*Wall Street Journal*, 21 September 2005)
38. But Houston residents and New Orleanians alike said they were beginning to sense fatigue from the city called upon with hardly a warning as a refuge for tens of thousands of freshly homeless people who arrived in shell shock, not knowing even the day of the week. (*New Orleans Times-Picayune*, 5 December 2005)
39. With TV and the press focused on politicians spitting at each other amid a natural catastrophe, average people in communities across the nation, aided by many can-do private companies, opened their doors to the genuine homeless people of New Orleans. (*Wall Street Journal*, 23 December 2005)
40. Foster said even actual homeless people were chipping in to support Lewis and her family. (*Atlanta Journal-Constitution*, 24 September 2005)
41. Clarence Moore, a children's pastor at Atlanta Vineyard Church in Norcross and the director of Cross Purpose Ministries, said he has expanded his normal homeless outreach ministries, trying to find accommodations for this new wave of homeless people. (*Atlanta Journal-Constitution*, 10 September 2005)
42. It would be scandalous if the president failed to drop those plans at a time of huge budget deficits, when the country must absorb hundreds of thousands of homeless people. (*Seattle Post-Intelligencer*, 9 September 2005)
43. But then the influx of newly homeless people turned Baton Rouge into a nightmare town, she says. "They were stopping you all the time, to see if you knew somewhere they could stay." (*Newsweek*, 12 December 2005)
44. "We're like homeless people. I don't know what to do. I don't know what type of life this is," said Frank, 38, sitting dejectedly on a mattress in the Golden Gate Motor Inn in Sheepshead Bay, Brooklyn, where she and her daughter were sent by the Red Cross last month. (*New York Daily News*, 5 October 2005)
45. While many chronically homeless people struggle with mental illness or drug addiction, the people recently left homeless by Hurricane Katrina also underscore the need to better respond to people suddenly left without resources, he said. (*Pittsburgh Post-Gazette*, 10 October 2005)
46. Why do our own people presume that homeless people are the same as criminals? Katrina taught us that we are all just a day away from being homeless ourselves. (*New Orleans Times-Picayune*, 26 December 2005)

Figure 4-13: Sample concordance lines of the collocation of *homeless* and *people*

Comparison to the persistently homeless is not always in vain, however. Some publications took the opportunity to highlight the struggles of this group through comparisons to PAKs and the sudden homeless, as in line 43 and 44, where improvements to social services and heightened empathy are publicized, rather than a stark 'temporary' vs. 'long-term' contrast.

### 4.3.5 Collocates from USAS I category: ‘low-income’ and ‘moderate/high-income’ sets

The semantic domain of ‘Money and commerce in industry’ shows close connections to *people* in the primary search, and is the category with the third highest frequency of adjectival collocates (after N: numbers, and S: social actions, states and processes). In fact, *poor* has the fourth highest log likelihood value (following *many*, *these*, and *young*), and is therefore one of the most salient collocates overall. All collocates from this category can be seen in Table 4-8 below.

I category adjectival collocates of <i>people</i>
<b>poor</b> [LL=2396.688, MI=4.567], <b>low-income</b> [LL=501.994, MI=4.873], <b>needy</b> [LL=226.881, MI=4.535], <b>wealthy</b> [LL=151.385, MI=3.851], <b>rich</b> [LL=142.941, MI=3.174], <b>unemployed</b> [LL=96.626, MI=4.227], <b>middle-class</b> [LL=80.188, MI=4.074], <b>working</b> [LL=57.242, MI=3.047]

Table 4-8: USAS I category collocates of *people*

Eight types populate this category, and can be divided into negative or positive subcategories, as seen below in Table 4-9. All of the ‘negative’ types collocate with greater frequency and strength than each of the ‘positive’ collocates, which gives an early indication of the discourses surrounding those who were directly impacted by Hurricane Katrina.

Category	Negative collocates of <i>people</i>	Positive collocates of <i>people</i>
I: Money and commerce in industry	<i>poor, low-income, needy, unemployed</i>	<i>middle-class, rich, wealthy, working</i>

Table 4-9: Adjectival collocates of *people* from USAS Broad Category I: Money and commerce in industry

Calculation of secondary collocation for all of the markers of income was not possible, as many of these (for instance, *wealthy* with *people*) have frequencies too low at the primary level to make secondary analysis very meaningful. Therefore, new aggregate searches of all of the ‘low income’ items and ‘moderate/high-income’ items were carried out as follows:

1. “The low income set”:  
 ⇒ ( *poor* | *low-income* | *low income* | *needy* | *unemployed* ) *people*
2. “The middle/high-income set”:  
 ⇒ ( *middle-class* | *middle class* | *rich* | *wealthy* | *working-class* | *working class* ) *people*

Collocation has been calculated on these aggregate sets, with minimum frequency set at 5 (to account for dramatically decreased frequency), spans of +/-3, and statistical

significance as standard. Collocates of these sets will be discussed in sections 4.3.5.1 and 4.3.5.2 below.

#### 4.3.5.1 Low-income set

Search terms from the low-income set returned 684 matches in 549 texts; the frequency breakdown of the individual search terms can be found below in Table 4-10. *Poor people* is the most common attribution of this set, making up nearly one in four of all results, and occurring 18.62 times per million words.

No.	Search result	No. of occurrences	Percent
1	<i>poor people</i>	512	74.85%
2	<i>low-income/low income people</i>	103	15.04%
3	<i>needy people</i>	48	7.02%
4	<i>unemployed people</i>	21	3.07%

Table 4-10: Frequency of nodes from the 'low-income set'

Secondary collocates of the low-income set can be found in Table 4-11 below. Many of the themes that occur here relate to those collocating more generally with *people* – *homeless* links back to section 4.3.4, and the consequence of a semantic preference for category N (numbers and measurement) will be more fully discussed in section 4.3.7. After briefly covering some new collocates, in this section I will focus on unexpected themes, and those thus far undeveloped in this chapter: semantic preference for one ethnicity (as in the S category), and collocation with the opposite attribute (with *rich* in the I category).

USAS broad category	Significant collocate lemma
A: General and abstract terms	<b>affect</b> [LL=26.797, MI=3.75]
H: Architecture, housing, and the home	<b>homeless</b> [LL=32.381, MI=4.722], <b>house</b> [LL=16.323, MI=3.252]
I: Money and commerce in industry	<b>subsidy</b> [LL=47.327, MI=6.294], <b>pay</b> [LL=45.12, MI=3.197], <b>poor</b> [LL=24.606, MI=3.535], <b>rich</b> [LL=23.428, MI=4.767]
N: Numbers and measurement	<b>thousand</b> [LL=61.423, MI=4.125], <b>fewer</b> [LL=43.156, MI=5.293], <b>mostly</b> [LL=35.735, MI=4.603], <b>count</b> [LL=35.489, MI=5.053],
O: substances and objects generally	<b>heating</b> [LL=64.073, MI=6.036]
S: Social actions, states and processes	<b>help</b> [LL=138.701, MI=3.589], <b>black</b> [LL=44.903, MI=3.83], <b>minority</b> [LL=33.896, MI=5.481], <b>assistance</b> [LL=27.36, MI=3.507],

Table 4-11: Secondary collocates of the 'low-income set'

Concordance lines for *affected* show expected discourse patterns, construing low-income *people* as being *affected* by circumstances either literally out of their control (the storm)



or described as figuratively so (legislation). A number of collocates—*subsidy, pay, heating, house, help, and assistance*—relate to bills and programs designed to assist PAKs to deal with household bills following the event (see line 47 for an illustration).

However, more notably, these lines show the first evidence of a phenomenon that I have deemed 'deviancy doubling'. This follows in the research tradition of Lloyd (1995), who studied society's harsh treatment of criminally violent women, who were deemed to be judged for both their transgressions from the law and from expectations of feminine behaviours. Consider line 48 below, where elderly and infirm people are likewise *affected*. Linking of two disparate 'them' social actors by showing commonalities in their experiences indicates that the identities construed are both considered to occupy 'out' groups in the cultural context of the text. This is not classical negative attribution of the out-group, but negativity through association. Deviancy doubling creates exponential social distance between the construed groups and the 'us' group of readers, who find decreasingly little room to relate with greater numbers of alienating others.

47. Now the bill must be returned to negotiators for more fine-tuning, as Democrats and moderate Republicans press for more funding for disease research and *subsidies to help low-income people pay their heating bills*. (*Washington Post*, 20 November 2005)
48. Although gruff and intense with his troops, Honore has displayed uncommon compassion with many of the elderly, infirm and *poor people affected* by the storm. (*New Orleans Times-Picayune*, 19 September 2005)

Figure 4-14: Sample concordance lines of low income set

More interesting perhaps are patterns of collocation and deviancy doubling with other social groups. The low-income group is repeatedly linked with minorities, most often *black* people. Explicitly, these groups are marked as having borne the brunt of the storm (which is statistically correct), but more interestingly, this is incorporated as evidence of governmental failure—both failure to practice compassion befitting a political candidate (line 49) and failure to hold all human lives in the same regard (line 50). Racial and financial minority groups are hereby linked as an accusation of institutional failure to support those perceived as most at risk in everyday life, risk only exasperated by the storm. This evidences what I call 'social sequestering', or the division of 'us' and 'them' groups into clear risk categories, by which the 'them' group bears the totality of risk, and the 'us' group is exonerated from risk (and relieved from risk society anxiety).

49. With the political fallout of Hurricane Katrina hovering over the proceedings, amid evidence that poor people and minorities have suffered the most in the disaster, Republicans also sought to portray Bush's nominee as a former lawyer who had at times displayed compassion for the dispossessed. (*Washington Post*, 14 September 2005)

50. Was the disaster the fault of a weakened system of levees and urban dikes? Or should we blame a political system that holds *black lives* and *poor people in low regard*? (*Pittsburgh Post-Gazette*, 30 August 2006)

Figure 4-15: Sample concordance lines of low-income set linked to other social groups

Deviancy doubling isn't the only pattern observed here; it is not only other minority (or relatively powerless) groups that the low-income set is linked to. As I highlighted earlier, collocation with the opposite set—the moderate/high-income set—also occurs. This will be discussed in the section for this set, below.

#### 4.3.5.2 Moderate/high-income set

The moderate/high-income set is much more rare in the corpus overall, occurring with less than 17% of the frequency of the low-income set. This could indicate two things: 1) that low income is a more salient topic in Hurricane Katrina reporting; and/or 2) that moderate/high-income is the unmarked, normative state, and marking of the low income identity is indeed overlexicalisation (or at the very least, conjecture) in many cases.

No.	Search result	No. of occurrences	Percent
1	<i>rich people</i>	44	40.74
2	<i>wealthy people</i>	26	24.07%
3	<i>middle-class/middle class people</i>	21	19.44%
4	<i>working-class/working class people</i>	17	15.75%

Table 4-12: Frequency of nodes from the 'moderate/high-income' set

Collocation was calculated as with the previous set, though the lower overall frequency of these items has affected the number of collocates produced. Only four can be found in Table 4-13 below, one of which (*affect*) is shared with the set above, and another (*poor*) points back to it.

USAS broad category	Significant collocate lemma
A: General and abstract terms	<i>affect</i> [LL=54.46, MI=6.326], <i>only</i> [LL=39.085, MI=4.502], <i>very</i> [LL=27.256, MI=4.656]
I: Money and commerce in industry	<i>poor</i> [LL=69.577, MI=6.433]

Table 4-13: Secondary collocates of the 'moderate/high-income set'

Though it is a collocate of both the low-income and moderate/high-income sets, *affect* is imbued with a slightly different prosody in this case. As indicated in line 51 below, *middle-class people* are also *affected*—and therefore passivized—by something outside of their control (the Bush administration's emergency plan) but the outcome is actually positive: the plan works for *only* them and those of a higher income bracket. This is in

contrast to the low-income group, who are negatively passivized in such constructions. Similarly, financial loss for the *very wealthy* is construed not as immediate, but as a consequence many decades in the making—the loss of inheritance for heirs (line 52).

51. Former President Bill Clinton, asked by President Bush to help raise money for the victims of Hurricane Katrina, offered harsh public criticism of the Bush administration's disaster-relief effort on Sunday, saying, "You can't have an emergency plan that works if it *only affects middle-class people* up." (*St. Petersburg Times*, 19 September 2005)
52. HURRICANE KATRINA may have cost *very wealthy people* a lot of money. Perhaps it would be more accurate to say that it may have cost their heirs a lot of money. The cost came not from the direct effects of the disaster. It came because the hurricane's impact on the poor people who remained in New Orleans made it politically unattractive for the Senate to vote on repealing the estate tax this week. (*New York Times*, 9 September 2005)

**Figure 4-16: Sample concordance lines of moderate/high-income set**

The final collocate is another pattern shared with the previous set—collocation between the two groups. Many of these concordance lines (illustrated by example 53 below) use *rich people* and *poor people* as a continuation of the indictment of governmental failure discussed above—indicating both the race and class divides in America through this construction even despite the use of the coordinating conjunction between. Two other patterns arise: the cry for help, and the rally for community. In the former (as in line 54), PAKs problematize the enforcement of defining barriers in reporting between the groups, iterating that this is "about people" rather than classism. In line 55, we see that some residents acknowledge that class divides exist in the New Orleans area, but feel as though this is what defines the region rather than what pulls it apart.

53. This historical moment is absolutely about the way the local and federal governments deal with *rich people* and *poor people*, and that's as clear as a cloudless sky if you're watching the news. Of course, Katrina's class issue was not the first horror that raked the conscience. Television gave us more than enough nightmares to go around -such as the fact that, if you turned on NBC or MSNBC, you could see people in their death throes. Television made the Gulf Coast's dire need palpable. Still, the race-and-class story is slapping us in the face. (*Seattle Post-Intelligencer*, 3 September 2005)
54. A weeping young woman held out her dehydrated-looking child and pleaded for help. "This is not about *rich people* or *poor people*," she said. "This is about people." (*New York Times*, 2 September 2005)
55. "The mix of *poor people* and *rich people* living close together creates our cultural milieu," Manning said. (*New Orleans Times-Picayune*, 22 October 2005)

**Figure 4-17: Sample concordance lines of low-income and moderate/high-income set collocating**

Analysis of concordance lines from this section show that many publications took the storm as a platform to point out class divisions in America, both to accuse the

government of acting upon classism in lack of broad-reaching aid networks, and to incorporate voices that ran counter to this, declaring that the rescue missions were about humanity rather than income. In section 4.3.8, this contrast is made even clearer. But first: *displaced*.

#### 4.3.6 Collocate from the USAS M category: *displaced*

There is only one statistically significant adjectival collocate from the USAS M category, and this is *displaced*. While *displaced* is an extremely interesting lexical item, and its representation of the M category does encompass the diaspora of PAKs in the days, weeks, months, and even years following Hurricane Katrina, *displaced people* as a bigram is not the only ‘way in’ to analysis of this naming strategy.

<b>M category adjectival collocates of <i>people</i></b>
<b><i>displaced</i> [LL=854.464, MI=4.057]</b>

Figure 4-18: The sole collocate of *people* from the USAS M category

In 15 of the 225 resulting concordance lines (6.7%), *displaced* is preceded by ‘the’ and treated as a collective noun, identifying all of Katrina’s *displaced* as one set, as in lines 56 and 57 below.

56. But it is also true that cities, institutions and families are chafing under the weight of the *displaced people* of Mississippi and Louisiana. (*New York Times*, 25 September 2005)
57. On Wednesday, a prime topic was the lingering aftermath of Hurricane Katrina in New Orleans and the *displaced people* taking their time leaving government-paid-for hotel rooms. (*Boston Globe*, 3 March 2006)

Figure 4-19: Sample concordance lines of *displaced people* preceded by ‘the’

There are relatively few secondary collocates of *displaced people* due to the low frequency of the primary collocation of *displaced* with *people*. However, these do occur with statistical significance, and can be grouped into USAS broad categories as below:

- USAS N category of numbers and measurement:
  - *thousand* [LL=143.15, MI=6.10]
  - *number* [LL=56.79, MI=4.52]
  - *many* [LL=43.96, MI 3.41]
- USAS H category of architecture, housing and the home:
  - *house* [LL=97.30, MI=6.09]
- USAS L category of life and living things:
  - *live* [LL=25.11, MI=3.58]

Secondary collocation with USAS semantic categories that are representative of all of the naming strategies under investigation in this chapter indicate that *displaced* is more than an attributive adjective in this corpus—it forms part of a bigram that is itself a naming strategy. Though this could have been included (and indeed, pilot studies have been run to test its suitability as a referential strategy), isolating instances where *displaced* appears on its own as a noun (e.g. ‘the displaced’, ‘Katrina’s displaced’) from cases where this is an adjective (e.g. ‘*displaced residents*’) could not be accomplished with precision. As such, the results that I come to regarding the alternative referential strategies—particularly *evacuee*, with which *displaced* shares the most concordance patterns—will represent findings on this naming strategy as well. Being inappropriate to discuss as an adjectival collocate, I thusly move on to the next category: USAS N.

#### ***4.3.7 Collocates from the USAS N category: many, most, fewer, and assorted figures***

The most highly populated USAS semantic category of collocates of *people* in the Katrina corpus is N: Numbers and measurement. In reviewing Figure 4-14 below, the pattern becomes clear. In the aftermath of natural disasters, media outlets might find it newsworthy or salient to report detailed information on the number of those impacted, most usually in the form of quantifiers.

The statistical calculation for collocation that is routinely used also has some bearing on the abundance of this type of collocate. Any given numeral—for instance, 180000 below—will be quite rare in the corpus, and may only appear in set contexts, i.e. “180000 people” in a certain news cycle. This would give this collocate much higher mutual information value and log likelihood values than actual import would indicate. In different data sets, the inclusion of dispersion data to downsample items not evenly distributed might be helpful. But in media texts of this sort (where the cycles following one large, catastrophic event are ever-changing), discounting change might actually remove many interesting findings. Therefore, I will touch upon this category briefly here, but focus on words that are more frequent in English overall (*many, most, and fewer*) rather than those whose statistical significance might have been influenced by relative rarity in the corpus as a whole.

### N category adjectival collocates of *people*

**many** [LL=10524.116, MI=4.283], **most** [LL=2276.501, MI=3.668], **1000** [LL=913.794, MI=4.775], **100000** [LL=837.912, MI=5.636], **100** [LL=782.617, MI=3.749], **200** [LL=646.891, MI=4.184], **10000** [LL=644.037, MI=5.007], **180000** [LL=612.254, MI=5.882], **200000** [LL=552.332, MI=5.576], **20000** [LL=552.332, MI=5.316], **1300** [LL=537.559, MI=6.144], **2500** [LL=527.907, MI=4.583], **500000** [LL=479.329, MI=5.991], **500** [LL=455.851, MI=3.881], **fewer** [LL=427.884, MI=4.075], **3000** [LL=412.094, MI=4.781], **300** [LL=404.615, MI=4.099], **50** [LL=390.271, MI=3.083], **400** [LL=357.679, MI=4.255], **30000** [LL=349.258, MI=5.133], **700** [LL=292.481, MI=4.675], **1200** [LL=286.391, MI=5.282], **150000** [LL=283.335, MI=5.428], **25000** [LL=282.846, MI=5.095], **1500** [LL=266.459, MI=4.64], **300000** [LL=244.303, MI=5.313], **250** [LL=236.445, MI=4.371], **6000** [LL=234.524, MI=4.709], **7000** [LL=230.341, MI=5.012], **1800** [LL=222.854, MI=5.744], **600** [LL=217.665, MI=4.074], **5000** [LL=212.999, MI=4.279], **4000** [LL=207.456, MI=4.51], **40000** [LL=204.769, MI=4.925], **12000** [LL=195.977, MI=5.136], **8000** [LL=194.506, MI=4.832], **400000** [LL=189.619, MI=5.22], **14000** [LL=183.631, MI=5.566], **70000** [LL=175.275, MI=5.511], **50000** [LL=174.927, MI=4.562], **80000** [LL=163.921, MI=5.382], **150** [LL=152.707, MI=4.083], **2000** [LL=148.968, MI=4.912], **600000** [LL=148.495, MI=5.422], **60000** [LL=145.779, MI=4.829], **75000** [LL=135.696, MI=5.73], **18000** [LL=127.965, MI=5.532], **1100** [LL=117.618, MI=4.719], **120** [LL=117.331, MI=3.96], **1700** [LL=111.288, MI=5.138], **250000** [LL=99.661, MI=4.749], **9000** [LL=96.273, MI=4.797], **65** [LL=69.663, MI=3.047], **15000** [LL=45.348, MI=5.102]

Table 4-14: USAS N category collocates of *people*, listed in order of descending log likelihood value

*People* collocates with *many*, *most* and *fewer* highly frequently: 2,554 times, 693 times, and 112 times, respectively. Therefore, secondary collocations have been calculated with a window span of +/-3 and a minimum collocate frequency of 5 to account for the lower overall frequency of the primary collocation. Given the nature of the primary collocation as quantification and a noun, the most interesting secondary collocates to consider are the verbs: what are so many *people* construed as doing or having done to them?

Interestingly, these patterns vary quite noticeably. *Many*, the most frequent of these items, offers the highest number of material verbs as secondary collocates. *Many people* are *stranded*, *displaced*, and *trapped* passively, but actively, they also *flee*, and *relocate*. By contrast, secondary collocates for *most* and *people* focus more on mental processes, e.g. *assume*, *imagine*, *realize*, *think*, *understand*, *believe*, and *like*. This makes them more passive in a social sense (they are construed as acting mostly within their own minds), but also less threatening (they are not being displaced *en masse* into another town or state). Secondary collocates of *fewer* are nearly all grammatical, indicating comparison, e.g. *than*, *far*, etc. The only two lexical items that collocate with *fewer people* with statistical significance are *need* and *come*, indicating both mental and material processes in action, for which *people* are activated in both instances.

I will not dwell long in this section, as quantification (found in high proportions of USAS category N collocates) is a pattern that occurs across nearly every one of the naming

strategies of ‘others’ that I explore in this thesis. See Chapters 5.5.1, 6.3.1, and 7.3.3 for much more detailed discussion of this pattern.

#### 4.3.8 Collocates from USAS S category: social actions, states and processes

After category N, category S contains the second highest number of collocates of *people*, marking social actions, states, and processes as a prominent component of the construal of human actors in the Katrina corpus (though noticeably less so than enumeration of these actors).

S category adjectival collocates of <i>people</i>
<b>black</b> [LL=1994.959, MI=4.112], <b>american</b> [LL=1825.694, MI=3.535], <b>white</b> [LL=454.07, MI=3.205], <b>stranded</b> [LL=287.235, MI=4.668], <b>encouraging</b> [LL=150.466, MI=4.031], <b>iraqi</b> [LL=101.934, MI=3.521], <b>resilient</b> [LL=90.751, MI=4.3], <b>generous</b> [LL=64.439, MI=3.119], <b>gay</b> [LL=54.185, MI=3.027]

Table 4-15: USAS S category collocates of *people*, listed in order of descending log likelihood value

Collocates from the S category fall under several subcategories, including identity types (e.g. *gay*, *Iraqi*) and personal qualities (e.g. *generous*, *resilient*). The most interesting feature that I spotted—and the one that I will begin this section with—is the occurrence of social-semantic antonyms, e.g. *black* and *white*.

##### 4.3.8.1 Black and White

The pre-Katrina population in New Orleans was approximately 70% black, which means that the delay of response and relief efforts for a disaster in this area would have certainly ‘disproportionately’ affected black people. However, the Press shows habitual marking of this group in reporting of an event that was not ethnically motivated or related (e.g. a hate crime). This could indicate ideological linkage between the storm and ethnicity, or perhaps needless overlexicalisation (and racism) on the part of the press. Further exploration might allow us to see whether this is the case.

*Black* collocates with *people* in a +/-1 span 515 times in 346 texts with a MI score of 4.11 and a log likelihood value of 1994.96. However, while *black people* and *people are black* are common 2- and 3-grams in the Katrina corpus, the high frequency of *black people* in particular can be partly attributed to Kanye West<sup>14</sup>, who was a powerful (non-institutional) voice following the storm.

<sup>14</sup> Kanye West (born 1977) is a music producer and Grammy Award-winning rapper who was raised in a middle-class family in suburban Chicago. His debut album *The College Dropout* was released in 2004, and sold over three million copies in the United States. The follow-up album, *Late Registration*, was released August 30, 2005 (three days before an appearance at a Hurricane Katrina benefit concert), and also sold over three million copies in America. American sales on each of his three later solo records have all

#### 4.3.8.1.1 “George Bush doesn’t care about black people”

During a live broadcast Hurricane Katrina benefit concert on September 2, 2005, the rapper infamously declared, “George Bush doesn’t care about black people”. This statement created a discourse entirely of its own: West quotations and direct derivatives account for 125 of the overall collocate frequency, and the pattern influences second-tier collocates for *black* collocating with *people*: the first is *care* [LL=1247.58, MI=8.82] and the fourth is *about* [LL=622.46, MI=4.84].

A search for the string *care about black people* returns 118 matches in 106 texts, and expanding the search by using the lemma of *care* brings up an additional seven references, for a total of 125 in 111 texts. Each of these 111 was published after West’s appearance at the benefit concert; no instances of *care/caring/cares about black people* predate his speech.

Of 41,964 texts in the corpus, only 1,080 were published between August 25, 2005 and September 1, 2005 (inclusive), or in the time between the storm’s first landfall in Florida and the September 2<sup>nd</sup> benefit concert. Because this window was so narrow, it’s important to look at a wider diachronic spread to increase confidence that West’s use of this phraseology led directly to an increase in its use overall.

Month/year	Words in category	Hits in category	Dispersion (no. files with 1+ hits)	Frequency per million words in category
<i>August 2005</i>	608,965	0	0 out of 684	0
<i>September 2005</i>	9,043,083	73	64 out of 10,219	8.07
<i>October 2005</i>	4,557,231	5	5 out of 5,360	1.1
<i>November 2005</i>	2,984,182	12	12 out of 3,583	4.02
<i>December 2005</i>	2,811,902	11	7 out of 3,385	3.91
<i>January 2006</i>	2,344,983	6	5 out of 2,770	2.56
<i>February 2006</i>	2,276,377	6	6 out of 2,731	2.64
<i>March 2006</i>	2,242,377	3	3 out of 2,572	1.34
<i>April 2006</i>	2,218,160	0	0 out of 2,429	0
<i>May 2006</i>	2,038,695	0	0 out of 2,143	0
<i>June 2006</i>	1,598,417	1	1 out of 1,779	0.63
<i>July 2006</i>	1,624,077	0	0 out of 1,804	0
<i>August 2006</i>	2,388,230	8	8 out of 2,505	3.35
<b>Total:</b>	<b>36,736,679</b>	<b>125</b>	<b>111 out of 41,964</b>	<b>3.4</b>

**Table 4-16: Frequency and dispersion of *care/caring/cares about black people* listed by month and year**

---

exceeded one million, and at the time of writing, West also owns and operates his own record label, GOOD Music. Source: [www.biography.com/people/kanye-west-362922](http://www.biography.com/people/kanye-west-362922)



Results shown in Table 4-16 reinforce the hypothesis that West was the direct originator of this phraseology. While no instances of *care/caring/cares about black people* are returned before his September 2<sup>nd</sup> speech, the phrase occurs 73 times in the same month, with a somewhat steadily decreasing frequency until August 2006, when it reappeared in yearly post-event summaries. We might then claim that he was somewhat responsible indirectly for the publicity of the ideology (belief in governmental neglect of this social group) that took place afterwards. West's statement was directly quoted 94 times in the corpus, with an additional 12 indirect quotes indicating attribution. While this does say something interesting about the sheer discursive force of a single sentence, I am equally interested in the 19 cases where the phraseology does not appear as a direct or indirect quote, but seemingly for another purpose. One of the clearest indicators that the phrase has been repurposed is the deletion of the auxiliary and negation, or a shift in the inflection of the verb *care* from simple present tense to present progressive. Concordance lines showing one of these two indicators are below:

58. Three days into rescue efforts, when it was clear Brown was over his head, Bush said he was doing 'a heck of a job.' That remark got people wondering if Bush cares about **black people**. Kanye West just said it out loud. (*Denver Post*, 13 September 2005)
59. There's no way to measure whether Bush really cares about **black people**. But one thing is clear: He knows how to care for his own. (*Denver Post*, 13 September 2005)
60. President Bush has been accused of not caring about **black people**. Especially poor ones. (*Atlanta Journal-Constitution*, 18 September 2005)
61. When asked about Kanye's recent controversial outburst in the wake of Hurricane Katrina, accusing President Bush of not caring about **black people**, Common said: "Kanye was speaking from the heart and he just said out loud what a lot of people were already thinking." (*Chicago Sun-Times*, 15 September 2005)
62. On Kanye West's criticism of George Bush for not caring about **black people**: it's not as easy as what Kanye West said, but I do know this: Bush has an unrealistic conception of black people in America. (*St. Louis Post-Dispatch*, 18 September 2005)
63. Asked in an ABC-Washington Post poll whether Bush let race and poverty affect the speed of the government's response, 76 percent of blacks and 24 percent of white respondents said "yes." The same poll asked whether Bush "cares about **black people**," and nearly seven in 10 blacks responded "no." (*San Francisco Chronicle*, 14 September 2005)
64. A Gallup poll found only 21 percent of blacks surveyed believe Bush cares about **black people**. (*San Francisco Chronicle*, 23 September 2005)

Figure 4-20: Sample concordance lines of alternate forms of the lemma *care*

Concordance lines 58 and 59 (from the same text) repurpose the quote to indicate authorial uncertainty as to whether the president does, in fact, 'care about black people'. The 60<sup>th</sup> concordance line shows a shift toward generalisation, by which the president is

accused in the passive voice, allowing for deletion of the agent, who is no longer just West. Concordance lines 61 and 62 repurpose the quote to expand upon the debate over racial division in America, while lines 63 and 64 show that the phrase was even incorporated into a poll – in which most respondents agreed with West’s sentiment.

These semantic differences tied to verb inflection highlighted the need for a more qualitative analysis of the entire data set. For each of the 125 concordance lines, a code describing the manner of invoking the phraseology has been assigned. The overall frequency is shown in the table to the right, and explanations and examples from each category are depicted below.

- a. **Direct Quote:** Quotation marks are used, and West is directly attributed.  
⇒ *Example:* On Friday, during an NBC telethon for Hurricane Katrina survivors, West said that “**George Bush doesn’t care about black people.**”
- b. **Indirect Quote:** Quotation marks are not used, though the phraseology is static. Attribution to West is either directly stated or can be inferred from context.  
⇒ *Example:* Bush appointee Michael Brown was so incompetent as the head of the Federal Emergency Management Agency during Hurricane Katrina that it was easy for many to take a rapper seriously when he said that **Bush didn’t care about black people.**
- c. **Repurposing:** Cases where the phraseology has been repurposed, here: reporting results of a question in a poll, or citing a sermon with the same name.  
⇒ *Example:* Repeated polls since Katrina have shown that blacks in overwhelming numbers believe that George Bush **does not care about black people.**
- d. **Generalisation:** West is not attributed; rather, the phrase/ideology is ascribed to a large, generalised population.  
⇒ *Example:* Many claim the **president doesn’t care about black people.**
- e. **Redirection:** George Bush is not the one accused of *not caring*; the verb is connected to another social actor, though the phraseology stays static.  
⇒ *Example:* But those with an interest in perpetuating the idea that the chief cause of black misfortune is an American culture that “**doesn’t care about black people**” decry racism whether it exists or not.
  - i. In one case, the redirection occurs with the object rather than the subject:

⇒ *Example: It's **not** that George Bush **doesn't care about black people**; he **doesn't care about poor people**.*

Direct quotations make up the majority of these cases, with indirect quotations occurring with the next highest frequency. However, over 15% of occurrences are *not* attributed to West himself, showing a wider adoption of the phraseology for various purposes, the most interesting of which for this research is generalisation, or attribution of the belief of one person to a larger, more amorphous group, or an indication of a developing ideology.

Use of West's phrase	Frequency
Direct quote	94
Indirect quote	12
Repurposing	7
Generalisation	6
Redirection	6
<i>Total</i>	<i>125</i>

**Table 4-17: Types of quotation and recontextualisation of Kanye West's quote**

George Bush himself seems to have responded to West's statement during a press meeting in the White House. In a December 2005 comment about the role that race may have played in response to Hurricane Katrina, Bush said, "One of the most hurtful things I can hear is, you know, 'Bush doesn't care about African-Americans' ... It's not true". I believe that the president revisiting this statement in an official comment, over three months after its initial utterance, underlines its cultural importance at the time. Also notable: Bush has changed the more colloquial, ideologically loaded 'black' to the more politically-correct 'African-American' in his refutation. (The next coming sections will explore the semantic differentiation between these two frequent descriptors.) His wife, Laura Bush, also contributed to the conversation in an interview (reprinted multiple times in the corpus) in which she states that, "I think all of those remarks are disgusting, to be perfectly frank, because of course President Bush cares about everyone in our country...I know what he's like and I know what he thinks and I know how he cares about people". Laura Bush's comments are interesting because she has removed racial descriptors from her language, which removes the marked construction. The occurrence of her quotes alongside those of her husband's indicated that a brief study into the construction of *care/cares/caring/cared about* would be interesting.

A review of the concordance lines for this search shows 40 instances of the following construction: [government actor, e.g. George Bush] + [optional negation] +

*care/cares/caring/cared about* + [social actors]. None of these cases predate Kanye West's speech, and a similar pattern as earlier (peaking in September 2005 before trailing off) can be observed in their occurrences. The following table shows the 'groups' that the government is construed as *caring* or *not caring* about.

Actor	Does care about	Does not care about
George Bush	everyone in our country (4); people (5);	african-americans (3); poor people (3); the plight of black people (2); everybody in America (2); blacks (1); Kanye West (1); people who look like me [African-American] (1); poor black people (1); the [poor] victims (1)
The president and the Republican party	all americans, regardless of color or income (1)	
Leaders	people (1)	americans (1)
The GOP		blacks (1)
Politicians		ordinary people (1)
The [Bush] administration		ordinary people (1); poor, non-white, sick, helpless people anywhere (2); the poor (1)
The government		people like them [residents of New Orleans] (1)
Those in charge		poor black people (1)
We, the rulers	the people (1)	
The country	the poor (1)	

**Table 4-18: Frequency of government actors depicted as *caring* or *not caring* about social actors**

As results in Table 4-18 show, government actors are portrayed as *not caring* about specific (mostly minority) groups, and at a far lesser frequency, *caring* mostly about ambiguous, generalised populations. George Bush is construed as *not caring about black people*, but also *poor* people, while the Bush administration additionally *does not care about ordinary people* (meaning people not from the privileged, mostly white, upper class) or the *sick and helpless*. Four constructions did not fit into these categories, as they indicated a hypothetical situation and a call to action. But it does seem that quotes such as "If Hurricane Katrina has taught us anything, it is that we need leaders who **care about all people, black and white, rich and poor, young and old**" imply that the current leaders do *not* care about 'all people' in this way.

West is widely quoted across multiple publications, though not always in full. *The Chicago Sun-Times* quoted West most frequently, with 17 instances in 15 texts

(20.72/million words in publication). *The New York Daily News* also reproduced this quote relatively frequently, featuring some iteration of it 10 times in 10 texts (17.63/million words in publication). While the portion of the quote directly attacking the president is clearly the most incendiary, it is interesting to note that only five times out of 125 did the Press reproduce the portion of West's quote pertaining to them: "I hate the way they portray us in the media. You see a black family, it says they're looting. You see a white family, it says they're looking for food." While the press seems eager to engage in allegations of racism on the part of the president (certainly in reproducing the most explicit accusation), journalists were largely unwilling to contribute to the discussion of whether the same sentiment had been reproduced by their own institution and in their own discourse.

The vilification of the government and governmental figures in the resulting media coverage, as well as the avoidance of fault on the part of the press, shows a redirection of discussion symptomatic of risk society discourse. The press can be seen as having avoided acknowledgement of a skewed representation of social actors in the press by favouring construction of a moral panic related to contention that the government is prejudiced against its own people. Having explored the high frequency of West's quote, its precession to a greater debate on the importance of race in American, and the variety of methods of (and possible motives for) rephrasing and reapplying the sentiment conveyed, I conclude that: 1) West's statement was a very explicit summation of what had previously been a marginal discourse in the press; and 2) served as a catalyst for wider consideration of Katrina rescue and recovery as racist and/or classist.

#### **4.3.8.1.2 Secondary collocations of *black + people* and *white + people***

With *care* and *about* featuring so highly on the secondary collocation list, it seemed likely that other discourses could be accessed by going through this list. Were there any overlapping identities or hidden meanings that West was making when he accused Bush of not caring about black people? This will be explored by way of contrastive analysis with the construal of *white people*.

This chapter is focused on predicational strategies, specifically attributes of social actors. Therefore, secondary collocates will be viewed through the scope of additional attributes associated with *black + people* and *white + people*. Significance thresholds have been maintained as throughout the thesis ( $MI \geq 3$ ,  $LI \geq 10.83$ ), with collocate lemma headwords restricted to adjectives and the minimum collocation frequency lowered to 5 to account for the much lower frequency of the primary collocation (515 hits overall for

*black + people*, 169 overall for *white + people*). Secondary collocates are listed in Table 4-19 below.

Category	Secondary collocates of...	
	<i>black + people</i>	<i>white + people</i>
I: money and commerce in industry	<b>poor</b> [LL=325.91, MI=6.41]	<b>rich</b> [LL=46.04, MI=6.96]
S: social actions, states and processes	<b>white</b> [LL=49.286, MI=4.61]	<b>black</b> [LL=78.46, MI=5.76]
T: time	<b>young</b> [LL=16.71, MI=3.30]	-

**Table 4-19: Secondary adjectival collocates of *black + people* and *white + people***

The first category (USAS I: money and commerce in industry) showcases a deep contrast between secondary collocates of *black + people* and *white + people*. *Black + people* only collocates with *poor*, whereas *white + people* collocates with the opposite: *rich*. The negative discourse around *poor black* people is addressed in sample line 65 below. News consumers readily believed false reports of “violent crime, rape and pillaging” in New Orleans because “the public is accustomed to riotous behaviour from black people in lower-class neighborhoods”, at least in reportage.

65. “It was all violent crime, rape and pillaging. But none of it was true.” To some in the public and in the news media, the images of barely checked violence in New Orleans, and its fleeing residents, seemed plausible. New Orleans is a violent city with an average of 200 homicides a year. The scenes of *poor black people engaged in lawlessness* after such events as the acquittal of the police officers who beat Rodney King or the 1977 New York City blackout are depressingly familiar, said writer and social critic Stanley Crouch. “The public is accustomed to riotous behavior from black people in lower-class neighborhoods,” he said. (*Washington Post*, 5 October 2005)
66. Howard Dean, chairman of the Democratic Party, is saying the *poor black people in New Orleans were not evacuated from the city because they are poor black people*. (*St. Louis Post-Dispatch*, 15 September 2005)
67. And when they look at New Orleans, they see glaring incompetence and racial injustice, where the *rich white people were saved* and the *poor black people were left to die hideous deaths*. They see some conservatives blaming the poor for not saving themselves. (*New York Times*, 7 September 2005)

**Figure 4-21: Sample concordance lines of *American* collocating with *people***

It is likely for this reason that these two naming strategies also collocate with one another. As in seen in concordance lines 66 and 67 above, the difference between *poor black* and *rich white* is explicitly specified as the reason that Howard Dean (66) and other citizens (67) believe certain PAKs were saved and others were not. Post-storm discourse of blame is completely shaped by the demographic of victims, and this is not performed quietly. Despite including statistical evidence that there are a greater number of poor whites, in the year following Hurricane Katrina, the press showed a much

stronger preference for reproducing a discourse of 'double deviancy' around the *poor* and *black*.

#### 4.3.8.2 *Iraqi and American*

Two additional collocates also appear to be somewhat antonymic, the indicators of nationality: *Iraqi* and *American*. The collocate *Iraqi* arises due to the United States' involvement in Iraq during the period of the corpus. As none of the 33 resulting concordance lines have any bearing on discourse after Hurricane Katrina, this has been discarded. However, *American* is very interesting, and bears deeper investigation.

The high instance of collocation between *American* and *people* was at first quite heartening. As the fifth highest collocate (LL+857.149) it buoys the top five, which otherwise consists of deviant out-group descriptors *poor*, *young*, *black*, and *homeless*. In a collection of exclusion, it seems to stand alone as marker of inclusion and belonging. To investigate this further, I generated 100 (random, reproducible) concordance lines for manual examination.

The majority share of the concordance lines by far is 53 instances (of the sample 100) within quotations, usually attributed to various politicians. Some of these are directly attributed to President Bush, for instance in line 68 below. Direct address to (or commentary on) the Bush administration constitutes 20 more concordance lines (e.g. 69).

68. Speaking to reporters at the White House, Bush vowed to follow up on the report to ensure the changes are made. 'I reminded our Cabinet that hurricane season begins in June and that we will be tracking the implementation of the recommendations in this report,' he said. 'We will learn from the lessons of the past to better protect the *American people*.' (*Boston Globe*, 24 February 2006)
69. Isn't it pathetic that Bush showed zero compassion or empathy when he finally made his announcement to the *American people*? (*San Francisco Chronicle*, 2 September 2005)

**Figure 4-22: Sample concordance lines of *American* collocating with *people***

It seems as though *American* collocates with *people* not as a marker of explicit inclusion, or even as a label for those *people* who were most directly affected by Hurricane Katrina. In particular, the set phrase 'the American people' is used almost exclusively by politicians, or by journalists addressing politicians. This phrase is not really a description of a collection of individual Americans, but is evoked as a rhetorical device abstracting responsibility to voters, akin to 'for the good of the nation'.

### 4.3.8.3 Gay

In this corpus, the collocation between *gay* and *people* is very infrequent, occurring just 22 times. However, this word is one of the core four explored in Chapter 9 (*homosexual, homosexuals, gay, and gays*), and the timespan of this corpus (2005-2006) is entirely contained within the AIDS/HIV corpus (1981-2010). Therefore, I defer the reader to later chapter for a much more detailed, nuanced discussion of this word.

### 4.3.8.4 Encouraging, Resilient, Generous

Character qualities *encouraging*, *resilient* and *generous* have been grouped together here for ease of analysis. Each of the 40 instances of *encouraging + people*, *encouraging* is a verb that has been mis-tagged as an adjective. Therefore, this item does not meet the criteria for inclusion in this chapter and has been discarded.

*Generous* displays the same characteristics that now seem to be characteristic of post-Katrina reporting; this is an attribute of those assisting PAKs, explicitly from areas outside of the Gulf Coast. Nashvillians are described as “a generous people” (line 47), with the use of an article indicating that this is a known and constant status.

Interestingly, line 48 shows both the pattern of good qualities (here, generosity) coming from around the country beyond the Gulf, but also a condemnation of the lack of such support by the government—another indication of risk society discourse.

70. Even by Southern standards, Nashvillians are a *generous people*, and the Nashville Symphony was quick to offer the Louisiana musicians an opportunity to play together. (*Wall Street Journal*, 6 October 2005)
71. *Generous people* from around the country answered the call for help after Katrina and Rita, but they certainly expected that the government would lead the way with organized and efficient recovery effort. (*News & Observer*, 3 January 2006)

Figure 4-23: Sample concordance lines of *generous* collocating with *people*

*Resilient* is the only collocate from this small subcategory that directly refers to PAKs, and it is operationalized as a predicational strategy by both the PAKs themselves, and by others who were more further removed from the circumstances. Three times, *resilient people* is preceded by the indefinite article, marking the subject as a separate tribe altogether. This distinction is heightened by the definition illustrated in line 72 below, where the “resilient people” are being described as being uniquely shaped by their particular geography. In inward-facing predication (line 73), these *resilient people* further assign themselves agency, e.g. “we make things happen”, whereas when this



attribute is assigned from the outside perspective (as in lines 74 and 75), other adjectives are listed, e.g. “phenomenal”, “friendliest”, “genuine”, and “good”.

72. We are a resilient people, shaped by the land and waters that gave birth to us. (*Pittsburgh Post-Gazette*, 7 September 2005)
73. “Jazz Fest is necessary,” she says, “We’re resilient people. We make things happen. People said we shouldn’t have a Mardi Gras or Jazz Fest, but we need to feel normal again. That’s our peace. And music is the biggest part of the equation here. It’s my heart, my soul, my gift from God.” Stranded in the Ninth Ward for days after the storm, Neville is slowly repairing her flood-ruined house while living in LaPlace, La. (*USA Today*, 1 May 2005)
74. My heart sank when I watched New Orleans and its phenomenal and resilient people brutally treated to a large dose of Mother Nature. (*USA Today*, 12 September 2005)
75. “The people down here are the friendliest, most resilient people,” he said. “It’s overwhelming. They’re genuine, good people.” (*San Francisco Chronicle*, 27 September 2005)

Figure 4-24: Sample concordance lines of *resilient* collocating with *people*

This shows that while *resilience* is a positively valued quality assigned to PAKs both by PAKs and by others, PAKs construct this attribute as tribal and view it as an opportunity for agency, whereas others continue the pattern of passivity and description. Passivisation is another feature of the next collocate—*stranded*—as well.

#### 4.3.8.5 *Stranded*

The attribute of being *stranded*, or abandoned without help, often alone, is one that was likely to crop up in disaster reporting. In the 62 concordance lines of this collocation, the themes of numbers and counts recur, with *stranded people* enumerated frequently throughout. Some other echoes of previous findings also reappear in this context; in line 76 below, the attribute of being *stranded* is explicitly linked with certain (“marginalized”) social groups, and linked to issues of “racial volatility” in America. The lack of aid that leads to finding oneself *stranded* is unambiguously attributed to lack of organization and failure in duty of care on the part of the government (risk society at work). In line 77, governmental mismanagement leads to “disbelief”, and the president is criticized for failing to recognize the severity of the situation in light of “tone-deaf optimism” regarding rebuilding. In response to these condemnations, the government is depicted then as attempting to “shift blame” in line 78. This was unsuccessful, as of course many officials and departments have now been singled out for catastrophic organizational failures (as in line 79).

76. It cuts to the issues of failure in the face of disaster, the limits of compassion, racial volatility, and the harsh realities of the American system that renders some of its communities -- like the marginalized **people stranded** and furious in New Orleans -- only as strong as they are able to make themselves. (*San Francisco Chronicle*, 4 September 2005)
77. The disbelief was about **people stranded** for days on rooftops, abandoned in a sports stadium, unprotected in hospitals, drowned in nursing homes. The disbelief was that while "this" happened, FEMA fiddled, a ship sat idle in the Gulf of Mexico with 120 sailors and 600 hospital beds, and the president expressed the tone-deaf optimism that Trent Lott would rebuild a "fantastic" new home from his rubble: "I'm looking forward to sitting on the porch." (*Boston Globe*, 9 September 2005)
78. The White House and Bush's apologists, stung by criticism that federal response was lacking, went in search of everything from the political registration of the New Orleans mayor to 100 school buses swamped in a lot, to shift blame for **people stranded** days after the storm. (*Pittsburgh Post-Gazette*, 11 September 2005)
79. Mr. Chertoff failed to designate a single person to oversee the federal response to Hurricane Katrina, the Government Accountability Office said in a preliminary report to Congress. That led to confusion and a lack of decisive action in the storm 's aftermath, the GAO said. It was the first time an investigation has singled out Mr. Chertoff and DHS in the delayed rescue efforts that left tens of thousands of **people stranded** for days in the flooded city without food or water. (*Wall Street Journal*, 2 February 2006)
80. Afterward he canoed into the flooded city, paddling through a cemetery to check his and his neighbors' homes. He helped rescue dozens of **stranded people**. "It was quite a surrealistic experience," he said. (*Atlanta Journal-Constitution*, 19 February 2006)

Figure 4-25: Sample concordance lines of *stranded* collocating with *people*.

These are all now-familiar signs of risk society, where the failure of the government to protect people from suffering or sudden, cataclysmic events is negatively evaluated, particularly in contrast to the individual 'hero story' (see line 80) where a single person acts in the absence of institutional aid to deliver assistance to fellow residents.

#### 4.3.9 Collocates from USAS T category: young and elderly

Interestingly, the only two collocates from the USAS T category to collocate with *people* in the given spans in the Katrina corpus are opposing identity indicators: *young* and *elderly*. In a corpus of natural disaster reporting, it is intuitive that the very old and the very young might be construed as risk groups, having perhaps restricted physical mobility, access to capital, or the power to independently evacuate for the short- or long-term. However, these two groups are construed quite differently in the corpus.

### T category adjectival collocates of *people*

**young** [LL=3192.527, MI=4.759], **elderly** [LL=507.24, MI=4.157]

Figure 4-26: USAS T category collocates of *people*, listed in order of descending log likelihood value

The *elderly* lemma collocates show a high instance of deviancy doubling, as previously charted in collocates of *poor*. This echoes Mautner's (2007) research on this particular identity group, indicating similarities in construction in American and British contexts. The top lemma collocates for *elderly people* here are *sick* (MI 8.38, LL 77.08), *poor* (MI 5.87, LL 49.42), *die* (MI 5.56, LL 40.34), and *kill* (MI 5.52, LL 28.51). The two former collocates place *elderly people* in specific risk groups, made more evident by the processes undergone in the latter two collocates—*dying* and being *killed*.

*Young* is not marked in this way, though both *young* and *elderly* people may be considered to be at special risk during natural disasters. By contrast, top lemma collocates of *young people* include *educate* (MI 6.50, LL 42.24), *opportunity* (MI 4.97, LL 99.33), *encourage* (MI 4.51, LL 30.44), and *involve* (MI 4.08, LL 22.63). This positivity around the naming strategy *young people* is dominated by texts from The New Orleans Times-Picayune, which accounts for two of the concordance lines below, but over 90% of concordance lines resulting from this collocation. Even in acknowledging the struggles that could be encountered in dealing with *young people*, publications are generally optimistic, explaining that the best ways of helping teenagers to avoid crime and drugs is to “engage [them] in healthy opportunities” (line 81) and “encourage [them] to reject the gansta ethos” (line 82). *Young people* are constructed as critical future members of society who must be given opportunities to “contribute to the fabric of our community” (line 83), get involved “in their country and in patriotism” (line 84) and possibly to “take up careers in government” (line 85).

81. About four decades ago a prominent mental health expert from Kansas said the best way to deal with teen crime and drugs is to **engage young people** in healthy opportunities. (*New Orleans Times-Picayune*, 16 February 2006)
82. He lived the hard life they rap about, and though it was too late to keep him off death row, he ultimately rejected the gangsta ethos and **encouraged young people** to do the same. (*Boston Globe*, 13 December 2005)
83. It is important that we offer our **young people** the **opportunity** to contribute to the fabric of our community. (*New Orleans Times-Picayune*, 21 May 2006)
84. I further believe that **getting** our qualified **young people involved** would promote an invaluable educational experience which would incur an involvement in their country and in patriotism. (*St. Petersburg Times*, 1 July 2006)

85. “We need to leverage the occasion of Katrina’s anniversary to focus new energy to the issue of improved government performance,” said Max Stier, president of the partnership, a nonpartisan group that seeks to revitalize federal service and encourage young people to take up careers in government. (*Washington Post*, 18 August 2006)

**Figure 4-27: Sample concordance lines of *young* collocating with both the white and *black* sets.**

While age is a salient attribute in the corpus of Hurricane Katrina reporting, not all ages are framed in the same discourse of risk. *Elderly people* are deviancy doubled with the *sick* and *poor*, and are construed as succumbing to death, whereas *young people* are at risk not of natural disasters, but falling prey to social pressures to take part in criminal activity. The actuality of such activities (using drugs, committing crimes) does not collocate with *young people* in the same way as found in Teo (2000), which is not to say that different naming strategies for the same group (e.g. *youths*, *teens*) might not have this result. But this was not marked out as a statistically significant naming strategy in this corpus, and will instead be marked as interesting fodder for future work. More limitations can be found in the next section—the summary.

## **4.4 Summary**

### **4.4.1 Of Findings**

*People* are surrounded by competing discourses: they are on the one hand *innocent* victims of the storm, but when *desperate*, they might be driven to criminal acts. Both *black* and *white* collocate with *people*, indicating that ethnicity was marked and salient in these texts. *Black people* additionally collocates with *poor*, and *white people* collocates with *rich*, a fact that was used to provide reasoning behind the failure of the (allegedly negligent, classist/racist) government to successfully evacuate the most needy citizens.

PAKs in general are the beneficiaries of aid and assistance, and are construed as being in multiple risk groups: storm survivors, *disabled*, *sick*, *ill*, *dying*. This contributes to a stigma of charity, and also removes the prospect of agency—PAKs are never seen to help or assist one another, but rather to passively receive this from individual citizens or charity organisations from around the country.

One unexpected challenge in this chapter was locating discourses around PAKs themselves at all. Much of the conversation occurring after Hurricane Katrina focuses on the government, on institutional forces, and on the individuals from outside of the Gulf region who assisted PAKs. Positive judgment is most often pointed outward, and form

predicational strategies for residents of cities that took in PAKs rather than PAKs themselves.

#### **4.4.2 Of methods**

Utilisation of a custom-collected, part-of-speech tagged corpus has lent great power to the speedy identification of the most frequent naming strategy. Analysis of this item has allowed me to establish a baseline of semantic preference for social actors throughout the corpus, to which I might compare other, less frequent naming strategies.

Collocates have also been part-of-speech tagged, and lemmatized by headword. This allowed me to thin results to adjective view only, and to collapse items that might have been relatively infrequent on their own (e.g. 'poorer') into a more frequent headword (e.g. 'poor') to take into account a wider range of types simultaneously. The automatic semantic tagger USAS was extremely helpful for objective categorization of collocates into groups. Both of these methods will be carried forward into all further chapters.

While narrowing the collocation span to +/-1 has kept results to a manageable minimum and exposed only the most frequent attributions of the naming strategy under investigation, there is an obvious limitation to this methodological choice. Such a restricted span results in the exclusion of less obvious prosodies occurring in colligation patterns such as people [copula] [adjective]. As a result, in the next chapter (chapter 5, on an assortment of less-frequent naming strategies in the Katrina corpus) I will resort to a more orthodox +/-3 span, given greater flexibility with more infrequent nodes.

Further, focus upon attribution was an interesting starting point for this, the primary analysis chapter of the Katrina corpus and of the thesis as a whole. However, much of the meaning made in the texts has been disregarded with collocates of other parts of speech. Therefore, Chapter 5 will be organized by part of speech of collocate headwords, including adjectival, nominal, and verbal categories of association with four naming strategies—*resident*, *evacuee*, *victim* and *survivor*.

## **Chapter 5: Semantic preference of frequent Hurricane Katrina naming strategies**

### **5.1 Introduction**

While those who followed the news during the aftermath of Hurricane Katrina might have a good general idea of popular naming strategies for people affected by Katrina (PAKs) employed during this time (perhaps *victims, Louisianans, evacuees*), the determination of search terms from a personal/experiential perspective could unfairly skew results toward constructions created using the more memorable, controversial terminology (*refugees, looters, the-black-and-poor*). It is imperative, then, to primarily allow the contents of the corpus to define common naming strategies while secondarily referring to intuition and memory for supplemental search terms.

In tandem with HIV/AIDS Chapter 8, this chapter addresses research question 3:

**What are the most frequent naming strategies employed in Hurricane Katrina and AIDS/HIV reporting, and what can corpus-based discourse analysis of associated attributes, transitivity, and cognitive metaphors tell us about the in- and/or out-group qualities attached to the people that the naming strategies represent?**

The general structure follows as before, with sections detailing how naming strategies/search nodes were selected coming immediately after the introduction, followed by a description of frequencies and dictionary definitions of naming strategies in section 5.3. Systematic analysis of collocates of each naming strategy begins in section 5.4, which leads in to discussion of metaphors and associated semantic prosody in section 5.5. The chapter ends with a summary of findings both methodological and discursive in nature.

### **5.2 Method: Targeting search terms**

To gain a basic and broad understanding of the social actors present within the corpus, two simple search strategies were used: nominal frequency lists and possessive collocations.

### 5.2.1 Nominal frequency lists

While perusal of frequency lists as the first point of analysis may offer insights into the strategies employed in smaller corpora, the Katrina corpus consists of 306,592 types and 36,736,679 tokens<sup>15</sup>, making manual analysis of the complete frequency list implausible. However, as the corpus was part-of-speech (POS) tagged in CQPweb using the UCREL CLAWS7 tagset (<http://ucrel.lancs.ac.uk/claws/>), it was possible to generate frequency lists more conducive to targeting frequent naming strategies. As in the previous chapter, lists of nouns were generated using CQPweb's POS tag search (which includes CLAWS7's 22 categories of noun), and it was these lists that were manually inspected for common 'people' nouns. Plural common nouns and common nouns neutral for count were particularly productive of naming strategies for humans; the former includes 10,713 *residents* (1.07% of all plural common nouns), 5,838 *victims* (0.58%), 5,150 *evacuees* (0.52%) and 2,882 *citizens* (0.29%), while the latter includes 68,353 *people* (53.19% of all common nouns neutral for number) – explored in Chapter 4.

No.	Search result	No. of occurrences	Percent of common plural nouns in the corpus
1	<b>officials</b>	14681	1.47%
2	<b>residents</b>	10713	1.07%
3	homes	9765	0.98%
4	<b>students</b>	9150	0.92%
5	<b>children</b>	9072	0.91%
6	<b>members</b>	8793	0.88%
7	prices	8590	0.86%
8	<b>police</b>	7256	0.73%
9	schools	6898	0.69%
10	services	6576	0.66%
11	companies	6262	0.63%
12	<b>workers</b>	6090	0.61%
13	areas	5989	0.6%
14	<b>victims</b>	5838	0.58%
15	hurricanes	5694	0.57%
16	others	5559	0.56%
17	things	5553	0.56%
18	<b>evacuees</b>	5150	0.52%
19	<b>families</b>	4983	0.5%
20	efforts	4706	0.47%

Table 5-1: List of the top 20 frequent common plural nouns with human actors in boldface

Though a future study into the representation of some of the actors above (e.g. police, officials, and families) would be extremely interesting, the scope of this research is

<sup>15</sup> CQPweb's token count does include punctuation marks, accounting for the difference between token count and overall word count of the corpus.

limited to media representation of those directly affected by the storm. Outside of a manual cleanse of each of the concordance lines surrounding these nodes, it would be difficult to distinguish discussion between results between, for instance, *families who* fled the storm, and news bulletins offering advice for *families of* missing New Orleans residents.

Further, while *citizens* was #45 in the frequency list of plural common nouns (above), inspection of the concordance lines revealed that a disproportionate number of these occurrences appear as part of lobbyist or civil rights group names, such as the *National Citizens' Coalition for Nursing Home Reform* or the *Division for Citizens with Developmental Disabilities*, or in the bi-gram *senior citizens*. As the geo-political nature of this naming strategy is held largely in common with the term *residents*, and as *residents* is the second most frequently used person-based plural noun in Table 5-1 the latter will be used for investigation into this type of referent. Similarly, *senior citizen* has been--and will continue to be--discussed with *elderly* (the much more frequent, related naming strategy) where relevant.

### **5.2.2 Possessive collocations**

The customised POS frequency lists do turn up the most frequent patterns of human referents, but it seemed that these findings could be best complemented by strategies targeting words of lower frequency and higher relevance. To target only those humans directly impacted by Hurricane Katrina, I again searched for collocates of the genitive case of Katrina (both the genitive 's and *of* prepositional phrases). Both sets of searches were limited first to +/-4 to allow for a broad range of intervening syntax and use of parallel conjunctions; 'human' results are reflected in tables 5-2 and 5-4 below. Because news media discourse is ultimately defined by shortened, elided textual features lacking in overly complex relational structures, I finally analysed "Katrina's" collocates exclusively at the +1 position (table 5-3) and "of Katrina" collocates exclusively at -1 position (table 5-5) to reveal the n-grams occurring tightest to the node phrase. Collocates in a +/-4 span for both search terms included *victim(s)*, *survivor(s)*, and *evacuee(s)*.



No.	Word	Total no. in whole corpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Log-likelihood value
8	victims	11,394	8.998	312	286	1618.048
49	survivors	2,838	2.241	32	30	110.931
141	dead	3,960	3.127	17	17	29.804
145	casualties	269	0.212	6	6	28.595
189	evacuees	10,140	8.007	23	22	18.541
228	victim	1,106	0.873	6	6	12.848

Table 5-2: Top 'human' collocates of *Katrina's* in the +/-4 span

No.	Word	Total no. in whole corpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Log-likelihood value
4	victims	11,394	1.5	266	243	2246.784
30	survivors	2,838	0.374	26	24	169.695
69	kids	5,655	0.744	13	13	49.861
74	dead	3,960	0.521	11	11	46.16
93	evacuees	10,140	1.335	11	11	27.088

Table 5-3: Top 'human' collocates of *Katrina's* in the +1 span

No.	Word	Total no. in whole corpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Log-likelihood value
3	victims	11,394	7.302	381	365	2284.396
14	survivors	2,838	1.819	79	73	443.839
16	evacuees	10,140	6.499	100	93	360.914
47	victim	1,106	0.709	15	14	63.158
99	refugees	1,546	0.991	9	9	23.73

Table 5-4: Top 'human' collocates of *of Katrina* in the +/-4 span

No.	Word	Total no. in whole corpus	Expected collocate frequency	Observed collocate frequency	In no. of texts	Log-likelihood value
4	victims	11,394	1.217	266	259	2360.824
14	survivors	2,838	0.303	35	34	263.725
34	victim	1,106	0.118	14	14	106.1
50	faces	1,957	0.209	11	11	65.643

Table 5-5: Top 'human' collocates of *of Katrina* in the -1 span

These three terms, in addition to *residents* (targeted above) form a group of node words of high statistical significance and varying degrees of frequency. It is these naming strategies that will form the focus of the following chapter.

### 5.3 Exploring naming strategies

In this section, the naming strategies will be first introduced by way of dictionary definitions, followed by information about frequency within the *Katrina* corpus and

publication preference. Afterwards, the collocational profiles of the various items will be discussed in turn.

### 5.3.1 Dictionary definitions

It is not my intention to state that these naming strategies are interchangeable (in fact, my aim is the opposite), or to hypothesize that journalistic choice for one word over another was dictated by malice, or even a conscious access to the more subtle meanings underlying them. I include a selection of dictionary entries below to demonstrate two points: 1) the diversity of identities construed in what might be conceived by many as ‘conventional’ readings of these words; and 2) the absence of a [natural] “disaster” as a cause of adopting any of these defined naming strategies, save for *survivor*.

	Longman Dictionary of Contemporary English Online <sup>16</sup>	Oxford English Dictionary Online <sup>11</sup>	Collins Cobuild Dictionary Online <sup>11</sup>
<i>Resident</i>	Someone who lives or stays in a particular place	A person who resides permanently in a place; a permanent or settled inhabitant of a town, district, etc.	The residents of a house or area are the people who live there
<i>Evacuee</i>	Someone who is sent away from a place because it is dangerous, for example because there is a war	A person who has been evacuated	An evacuee is someone who has been sent away from a dangerous place to somewhere safe, especially during a war
<i>Victim</i>	Someone who has been attacked, robbed, or murdered; someone who suffers because of something bad that happens or because of an illness	One who is reduced or destined to suffer under some oppressive or destructive agency	A victim is someone who has been hurt or killed
<i>Survivor</i>	Someone who continues to live after an accident, war, or illness; someone who manages to live normally in spite of many problems	A person, animal, or plant that outlives another or others; one remaining alive after another's death, or after some <u>disaster</u> in which others perish	A survivor of a <u>disaster</u> , accident, or illness is someone who continues to live afterwards in spite of coming close to death; a survivor of a very unpleasant experience is a person who has had such an experience, and who is still affected by it

Table 5-6: Dictionary definitions of *resident*, *evacuee*, *victim* and *survivor*

<sup>16</sup> Accessed on 28 November 2011.

It is interesting to note, then, that *survivor* is by far the least frequent naming strategy employed in the Katrina corpus. In exploring the collocates, we will see if this exaggerated link between *survivors* and disaster (and *residents* and the home, *evacuees* and diaspora, *victim* and suffering) hold true. But first, I provide some additional information about the incidence of all naming strategies, including those more frequent.

### **5.3.2 Incidence of naming strategies**

Herein, the incidence of naming strategies will be discussed in terms of overall occurrence, usage patterns over time, and observed preference of publications.

*Resident/s* was the most frequently occurring naming strategy found overall, with 23,523 matches in 10,862 different texts. From those strategies arising from genitive searches, *victim/s* was the most popular, occurring 12,284 times in 7,427 texts. *Evacuee/s* follows closely in frequency (11,052) but had much lower dispersion (4,516 different texts). *Survivor/s* is relatively uncommon, appearing only 3,357 times in 2,262 texts.

When viewing occurrence of nodes over time (see Fig. 5-1 below), it becomes apparent that *resident/s* occurs with the highest frequency not only overall, but also for each month of reporting. *Resident* is the only node to experience a trough in the month of September 2005; all others peak in the weeks following Hurricane Katrina before showing a general trend of decline. Contrary to this pattern, *resident* falls precipitously from August 2005 to September 2005, and remains at its lowest frequencies through November 2005 before beginning to climb once more. *Victim* and *evacuee* experience steep climbs followed by sharp falls over September 2005. All of these naming strategies become less salient over time, though the patterns with which they fall out of fashion do differ.

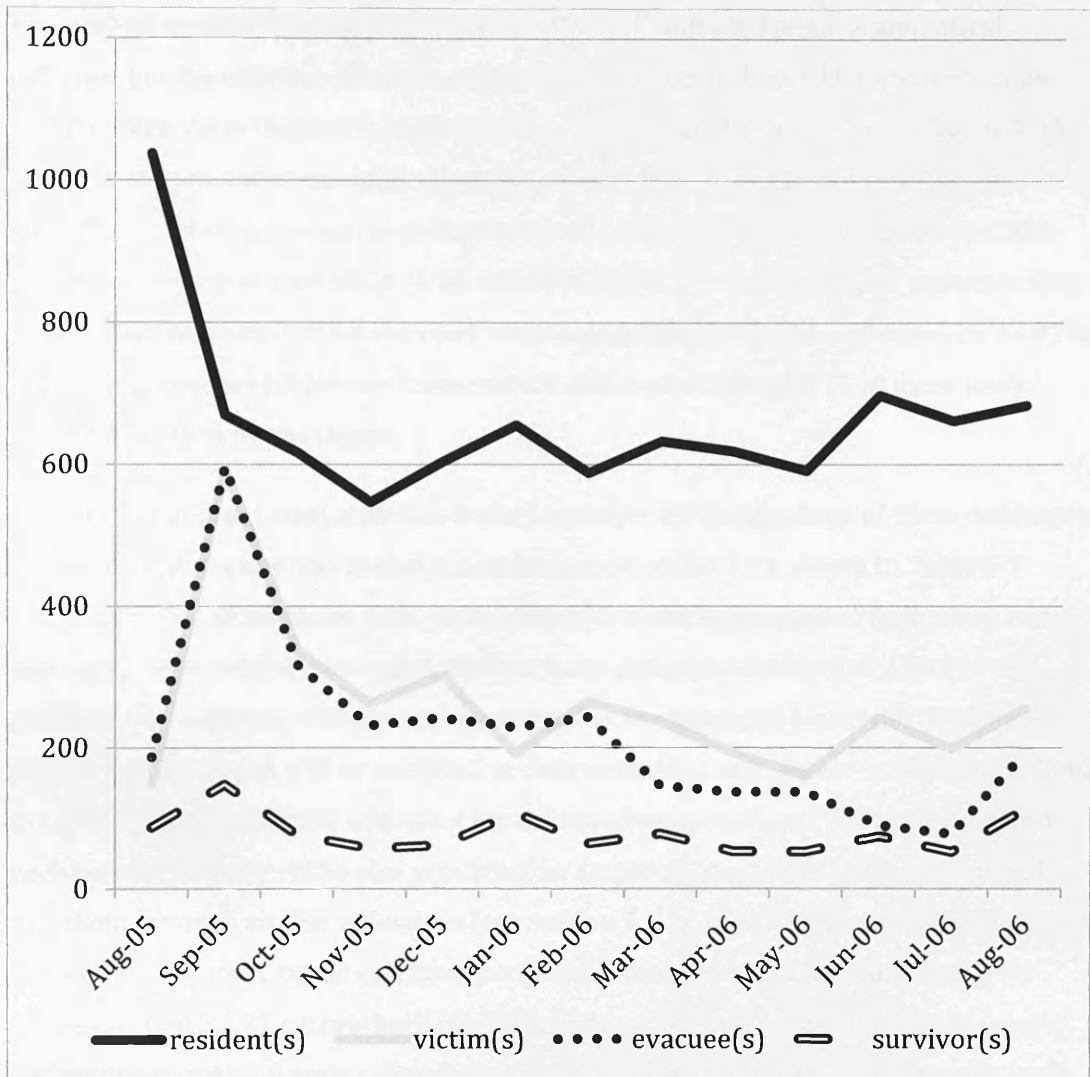


Figure 5-1: Frequency of naming strategies over time, expressed in frequency/million

#### 5.4 Method: Downsampling collocates

This section details the method by which a section of collocates have been selected for close examination.

Collocates were limited to those occurring at least 20 times within the corpus, and with mutual information scores of at least 3 and log likelihood scores over 10.83, or within the 99.9<sup>th</sup> percentile ( $p < 0.001$ ). Even with fairly rigorous requirements such as these, 175 lemmas of high statistical significance were returned for the four naming strategies under investigation (two of these are prepositions from the Z category and will not be discussed in this chapter). To make sense of the patterns emerging from such a broad selection of results, it became necessary to group collocates together by classes.

The part-of-speech (POS) tagging within CQPweb illuminated some grammatical patterns, but the addition of another classification system allowed for greater insight into meaning. As in Chapter 4, each collocate was assigned a candidate tag by the USAS semantic tagger, and these were manually disambiguated. In 123 out of 173 total cases (71.10%) of adjective, noun, or verb lemma collocates, the first tag assigned by USAS was deemed the most accurate. In 26 cases (15.03%), another candidate tag other than the first one was selected for the most salient position. Finally, 24 collocates, or 13.87%, needed tags assigned manually from outside of the candidate list; 15 of these were *Unmatched* (Z99) by the tagger.

To ensure that the closest attention could be paid to the most salient of these collocates, I have employed a method of downsampling based on the data shown in Table 5-7 below. For each of the three main parts of speech under investigation (adjective, noun, and verb), relevant collocate lemmata have been categorized into their USAS broad category. The contents of the semantic categories containing at least 10% of collocates of a given part of speech will be analysed in their entirety, and appear in boldface in Table 5-7. It is my belief that this will allow for the broadest coverage of discourses around each naming strategy while also providing an empirical, mathematically determined threshold for analysis. For adjectives (see section 5.5.1), this contains USAS categories M (movement, location, travel and transport) and N (numbers and measurement); for nouns (section 5.5.2), categories S (social actions, states and processes) and W (world and environment); whereas verbs (section 5.5.3) showed a larger spread, having large percentages of collocates across categories H (architecture, housing and the home), M, S, and X (psychological actions, states and processes).

USAS Broad Semantic Category of ADJECTIVE Collocate Lemmata

	A	B	E	G	H	I	L	M	N	P	Q	S	T	W	X	Z	Total
resident	2	1	1	0	0	1	0	6	7	0	0	2	3	2	0	2	28
victim	1	0	1	0	0	1	0	1	0	0	0	1	1	0	0	0	6
evacuee	0	0	0	0	0	0	0	0	19	0	0	1	1	0	0	0	21
survivor	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	2
Total	3	1	2	0	0	2	0	8	27	0	0	4	5	2	0	-	
% all POS collocates	5.6%	1.9%	3.7%	0.0%	0.0%	3.7%	0.0%	14.8%	50.0%	0.0%	0.0%	7.4%	9.3%	3.7%	0.0%	-	

USAS Broad Semantic Category of NOUN Collocate Lemmata

	A	B	E	G	H	I	L	M	N	P	Q	S	T	W	X	Z	Total
resident	1	2	0	0	2	0	0	1	1	0	2	1	0	0	0	14	24
victim	2	0	3	1	0	1	0	0	0	0	0	6	0	7	0	1	21
evacuee	1	0	0	0	2	0	0	2	1	0	0	0	0	1	0	3	10
survivor	0	2	1	0	0	1	1	0	0	0	2	3	0	1	0	1	12
Total	4	4	4	1	4	2	1	3	2	0	4	10	0	9	0	-	
% all POS collocates	8.2%	8.2%	8.2%	2.0%	8.2%	4.1%	2.0%	6.1%	4.1%	0.0%	8.2%	20.4%	0.0%	18.4%	0.0%	-	

USAS Broad Semantic Category of VERB Collocate Lemmata

	A	B	E	G	H	I	L	M	N	P	Q	S	T	W	X	Z	Total
resident	0	0	0	0	0	0	0	7	0	0	2	2	0	0	3	0	14
victim	0	1	0	0	2	0	0	0	0	0	0	6	0	1	1	0	11
evacuee	1	0	0	0	3	0	0	7	0	1	1	4	0	1	0	0	18
survivor	1	0	0	0	1	0	0	0	0	0	0	3	0	0	1	0	6
Total	2	1	0	0	6	0	0	14	0	1	3	15	0	2	5	-	
% all POS collocates	4.1%	2.0%	0.0%	0.0%	12.2%	0.0%	0.0%	28.6%	0.0%	2.0%	6.1%	30.6%	0.0%	4.1%	10.2%	-	

Table 5-7: Collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*, categorized by part of speech and USAS broad semantic categories

## 5.5 Analysis

We now move in turn through the collocates representing the most proportionally substantial USAS semantic categories, by part of speech. This section will begin by adjectival collocates, followed by nominal collocates, and finally verbal collocates. Some comments will be made after these pertaining to metaphorical uses of some words before I summarise results and evaluate this method in section 5.7.

Please note that in tables of this section, collocates (of the same part of speech and USAS semantic tag) shared between two or more naming strategies will appear in italics for emphasis.

### 5.5.1 Adjectival collocations

Cross-calculating lemma collocations on the basis of POS tags and semtags resulted in a list of 56 adjectival collocates of statistical significance organized into 16 of the 20 major USAS fields (with Z discounted). Adjectival collocates per node are as follows: *resident* (27), *victim* (6), *evacuee* (21), and *survivor* (2). A full list of all collocates presented by POS and broad USAS category are presented in Appendix C (*resident*), Appendix D (*victim*), Appendix E (*evacuee*), and Appendix F (*survivor*). This section focuses on discussion of frequent adjectival collocates as members of semantic categories M and N, representing 16.36% and 49.09% of adjectival collocates respectively.

#### 5.5.1.1 Category M: movement, location, travel and transport

As the distribution in Table 9 below shows, adjectival M category collocates are disproportionately favoured by *resident(s)*, which has 7 such, and disfavoured by *evacuee(s)*, which has none at all. *Victim(s)* and *survivor(s)* have only one collocate apiece here, and this is shared both between the two and with *resident(s)*, making *displaced* the most widely dispersed collocate in the group, and my starting point of analysis.

Naming strategy	M category adjectival collocates of naming strategy
resident(s)	<i>displaced</i> [LL=696.079, MI=5.431], <i>remaining</i> [LL=175.485, MI=3.391], <i>returning</i> [LL=140.039, MI=5.051], <i>low-lying</i> [LL=130.319, MI=3.922], <i>uptown</i> [LL=127.895, MI=3.326], <i>evacuated</i> [LL=99.733, MI=4.911]
victim(s)	<i>displaced</i> [LL=442.75, MI=3.946]
evacuee(s)	-
survivor(s)	<i>displaced</i> [LL=118.396, MI=3.931]

Table 5-8: USAS M category adjectival collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*

Concordance lines arising from the collocations of the various naming strategies with *displaced* are extremely numerous – 654 for *resident(s)*, 122 for *victim(s)*, and 33 for *survivor(s)*. As such, I have employed the method of secondary collocation here to glean some information about recurrent connotations with these identities. Statistical cut-offs are as ever (MI>3, LL>10.83), though the minimum frequency has been lowered from 20 to five to allow for the decreased number of initial search query results.

Secondary collocates of the collocation between *displaced* and *residents* belong to the semantic fields of housing (*house, housing, address, trailer*), movement and location (*gulf, scatter, coast, Rouge, return, Bernard, Louisiana*), government (*vote*), numbers and measurement (*thousand, hundred, fill, many*), and social roles (*black, worker*). Secondary collocates for the collocation of *victim(s)* with *displaced* belong only to the semantic fields of housing (*house*) and numbers and measurement (*thousand, many*). For *displaced* with *survivor(s)*, secondary collocates include those of housing (*house*) and movement and location (*arrive, city*). Therefore, the availability of housing and the responsibility for placing *displaced* PAKs into it is of paramount concern across these three naming strategies. The *topos* of numbers features again here (as was discussed in the previous Chapter, and will be shortly revisited in section 5.1.1.2), as does the repeated semantic field of movement and location (see section 5.5.3.2 for more on this), construing PAKs as large, threatening and homogenous masses moving in on various *cities* and places. *Resident* has the most complex secondary collocational profile, linked quite closely to those adjectives reviewed in the previous chapter, including one of race (*black*) and another indicating income (*worker*). Refer back to the sections 4.3.5 and 4.3.8 for discussion of the impact of these associations. As *displaced* appears with such a variety of secondary semantic preferences and in a range of contexts – not to mention that it also collocates as a verb later in this chapter – I will not dwell too long in this section for risk of extrapolation. Rather, I turn to unique collocates of *resident* that can be demonstrated to have more clear meanings in concordance.

*Residents* who have *remained* in New Orleans or other areas despite recommendations for evacuation are construed as dangers to themselves and to others, perpetrating public health risk by staying in contaminated areas; see lines 1 and 2 below for illustrative examples. The same holds true for *returning residents* (e.g. line 3 below) who had gone back to see their homes before clearance was granted through municipalities.



1. Widespread gas leaks and fires, contaminated floodwater and disease-bearing mosquitoes have rendered the city unsafe. He begged the city's remaining residents to leave, estimating that 5,000 to 10,000 people are still trying to hold on and make ends meet in the city. (*New Orleans Times-Picayune*, 7 September 2005)
2. City medical officials said no major health epidemics have yet emerged, but they warned remaining residents not to make skin contact with the stagnant water. (*Wall Street Journal*, 9 September 2005)
3. The Health and Hospitals document was nearly identical in tone and message to a news release posted Wednesday on the city's official Web site, in which returning residents are warned that they are "entering the city of New Orleans at your own risk," because "there are still many health and safety issues." (*New Orleans Times-Picayune*, 28 September 2005)

**Figure 5-2: Sample concordance lines of *remaining* and *returning* collocating with *residents***

The opposite is true for *residents* who *evacuate* rather than *remain*. The collocation between *resident(s)* and the adjective *evacuated* is not a frequent one, occurring just at the minimum threshold of 20 times. This might be expected, given that the identity construed might be better (or more succinctly) conveyed with the use of the single-word naming strategy *evacuee*. However, *evacuated resident* is a relatively empowered identity, particularly in comparison to *evacuee* (as I will soon demonstrate). Whereas *evacuees* have left but may never return, *evacuated residents* retain their residency, and are courted throughout the concordance lines as displaced citizens. In concordance line 4 below, they are planned participants in public meetings; in line 5, state and local officials are attempting to "persuade" them to return. The temporariness of their displacement allows for some power of "belonging" to a place (that they currently do not live in) to endure.

4. St. Bernard Parish officials on Friday turned their attention to the parish's reconstruction, announcing plans to hold a public meeting with evacuated residents in Baton Rouge next week, as Hurricane Katrina's waters continued to recede and recovery teams collected more bodies. (*New Orleans Times-Picayune*, 10 September 2005)
5. As evidence, the state and local officials cite an array of stalled bills and policy changes they say are crucial to rebuilding the city and persuading some of its hundreds of thousands of evacuated residents to return, including measures to finance long-term hurricane protection, revive small businesses and compensate the uninsured. (*New York Times*, 22 November 2005)

**Figure 5-3: Sample concordance lines of the collocation between *evacuated* and *resident(s)***

*Low-lying* and *uptown* are adjectives modifying nouns of location, such as "neighborhoods" or "areas", and does not directly refer to *resident(s)*. Therefore, these have been discounted for the purpose of my analysis.

Adjectival collocates have not proven particularly insightful in this section, as the attribute of being on the move is somewhat expected in the given context. Turning instead to the next section – numbers and measurement – will provide results less self-evident in nature.

### 5.5.1.2 Category N: numbers and measurement

Three of the four naming strategies investigated here are modified with adjectives of number or measurement, a pattern corresponding closely to findings from previous studies (e.g. Gabrielatos & Baker, 2008; KhosraviNik, 2010; Krzyzanowski & Wodak, 2009). This is an interesting pattern that has been proven to be typical of out-group construction, so little more will be said here other than that an overwhelming appearance of numbers and measurements in connection with the groups of (American) people explored here does indicate an othering strategy at work, namely of aggregation, collectivization, and de-humanisation (KhosraviNik, 2010). It should be noted that two naming strategies in particular – *resident(s)* and *evacuee(s)* – attract the wealth of collocates in this part of speech and semantic category. *Victim(s)* does not collocate with any numbers or measurements, and *survivor(s)* has a single collocate: *800*. This only occurs in a single context giving information about the “U.S. Department of Health and Human Services confidential helpline for hurricane survivors: (800) 789-2647” at the bottom of articles, and is not meaningful to this analysis. Therefore, this has been discarded, and the basis of analysis in this section will be on the collocates of *resident(s)* and *evacuee(s)* alone – see Table 5-9 below for a full list.

Naming strategy	N category adjectival collocates of naming strategy
resident(s)	<b>60000</b> [LL=173.334, MI=4.886], <b>460000</b> [LL=172.983, MI=6.756], <b>nearby</b> [LL=165.708, MI=3.063], <b>100000</b> [LL=163.672, MI=3.875], <b>500000</b> [LL=121.18, MI=4.443], <b>200000</b> [LL=77.116, MI=3.516], <b>20000</b> [LL=63.293, MI=3.166]
victim(s)	-
evacuee(s)	<b>150000</b> [LL=189.684, MI=5.902], <b>300</b> [LL=181.716, MI=3.953], <b>estimated</b> [LL=143.107, MI=3.66], <b>10000</b> [LL=138.938, MI=4.124], <b>25000</b> [LL=137.048, MI=5.02], <b>2000</b> [LL=136.133, MI=4.065], <b>1000</b> [LL=133.634, MI=3.548], <b>5000</b> [LL=115.043, MI=4.31], <b>200000</b> [LL=112.641, MI=4.605], <b>20000</b> [LL=96.851, MI=4.256], <b>100000</b> [LL=95.179, MI=4.085], <b>200</b> [LL=92.311, MI=3.008], <b>4000</b> [LL=91.244, MI=4.34], <b>1500</b> [LL=86.117, MI=4.166], <b>30000</b> [LL=82.817, MI=4.335], <b>400</b> [LL=81.387, MI=3.475], <b>250</b> [LL=79.787, MI=3.947], <b>600</b> [LL=75.611, MI=3.689], <b>3000</b> [LL=73.784, MI=3.739]
survivor(s)	<b>800</b> [LL=425.447, MI=5.129]

Table 5-9: USAS N category adjectival collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*

There are two main patterns in these collocates – quantification (the majority pattern) and approximation (the minority pattern containing only two examples: *nearby* collocating with *resident(s)* and *estimated* collocating with *evacuee(s)*). In the name of expediency, I will dispatch with the minority pattern first before detailing the majority of cases.

The collocation of *resident(s)* with *nearby* neatly corresponds with the word's dictionary entry, and one of its most accessible senses. Measure of proximity sensibly fits with *resident(s)*, who are described as living in (or forcibly leaving from) specific areas throughout the corpus. There is no clear pattern in these 66 concordance lines, as the object of what is – or was – *nearby* is always changing (see, for instance, line 6 below).

6. Cannizaro said it was unclear whether the body, which was in advanced state of decomposition, was a patient from the nursing home or a *nearby resident who floated to the site*. (*New Orleans Times-Picayune*, 2 October 2005)

Figure 5-4: Sample concordance line of the collocation between *nearby* and *resident(s)*

The minority pattern collocation between *evacuees* and *estimated*, however, is more complex and measureable. Indeed, this links in to the majority pattern of quantification, as a count (understandably) always intervenes between the two collocates. When the number of *evacuees* must be *estimated* (and the estimates are quite large indeed, going into the tens of thousands and sometimes millions), they are the subjects of the *topos* of numbers in addition to the *topos* of threat. In concordance line 7 below, these *evacuees* threaten social order by leaving human waste and “mountainous heaps of trash” in their wake; in line 8, a machinery metaphor is employed, as *evacuees* “fueled” a “horrific traffic jam”. Lines 8 and 9 contain bursting/filling metaphors (again referencing the *topos* of threat by overfilling). In line 9, this overlaps with the sort of water metaphors demonstrated in line 10, where human *evacuees* begin “trickling back home” (see section 5.6 for more on water metaphors).

7. For the first week after the storm, the convention center served as an evacuation point for an *estimated 15,000 evacuees who left food, human waste, and mountainous heaps of trash*. (*Boston Globe*, 11 September 2005)
8. Houston was mostly empty, but just beyond the metro area, *the horrific traffic jam, fueled by an estimated 2.8 million evacuees*, was only beginning to taper off. (*Atlanta Journal-Constitution*, 24 September 2005)
9. And unbeknownst to FEMA, *a new circle of hell was opening downtown*, as the New Orleans convention center *filled with an estimated 25,000 evacuees*, many of them unable to get to the flooded area around the Superdome. (*Washington Post*, 11 September 2005)

10. Some of the *estimated 2.5 million evacuees* began trickling back home to Houston and surrounding Gulf Coast communities on Sunday. (*Washington Post*, 26 September 2005)

**Figure 5-5: Sample concordance lines of the collocation between *estimated* and *evacuee(s)***

Nearly identical patterns occur in the concordance lines arising from collocation between *resident(s)* and *evacuee(s)* with numbers (see Figure 5-4 below). In lines 11-12, lack of agency in large counts of PAKs can be seen – they are unable to move (“stuck”), or unable to choose to stay put (“driven out”). The *topos* of threat/strain is illustrated in lines 13 and 14, where *resident(s)* add “to already serious pressure” and *evacuees* “put ‘a considerable strain’ on the community” (while also being associated with an increase in crime). The filling/bursting metaphor also associated with the *topos* of strain is realized in these concordance lines as well; for an example, see concordance line 15 below, where *evacuees* “pack” shelters. In contrast to the estimation of *evacuees* “trickling” home, a known number of *residents* “poured into the Superdome” in another of many examples of the combination of the water metaphor/threat *topos* (line 16). These examples have prompted me to explore bursting/water metaphors further – see section 5.6 for more details.

11. For weeks last September, virtually all **460,000 residents** were stuck on someone’s sofa, in a hotel or on a cot in an arena somewhere. (*USA Today*, 29 August 2006)
12. Merlo’s determination to defy the abnormal state of everything here by acting as normal as possible stands out in a city that has seen nearly all of its **500,000 residents** driven out. (*Boston Globe*, 12 September 2005)
13. The **60,000** extra **residents** are adding to already serious pressure on critical highways, he said. (*New Orleans Times-Picayune*, 29 March 2006)
14. A survey by a Rice University sociologist found that three in four Houstonians believe the **150,000 evacuees** have put “a considerable strain” on the community, and two-thirds said the newcomers are responsible for “a major increase in violent crime.” (*New Orleans Times-Picayune*, 26 March 2006)
15. That’s down from the **30,000** or so **evacuees** who packed the three shelters last week. (*The San Francisco Chronicle*, 10 September 2005)
16. As water poured into the city, as many as **20,000** more **residents** poured into the Superdome. (*Washington Post*, 11 September 2005)

**Figure 5-6: Sample concordance lines of the collocation between figures and *evacuee(s)* and *resident(s)***

Use of ‘aggregation’ and ‘numerisation’ alongside the naming strategies of *residents* and *evacuees* “construct these people as passive, unified, motionless and mute people (or figures) who merely constitute the topic” rather than contributing to the discussion (KhosraviNik, 2010, p. 16). Indeed, they are shown to disrupt social order, creating damage literally (in disrupting traffic flows and leaving human waste in physical

structures) or threatening it metaphorically (by overstraining social systems and the physical boundaries of architecture). By contrast, *victims* and *survivors* are not construed in these terms, indicating a different understanding of their experience after the storm.

### 5.5.2 Nominal collocations

Collectively, the four naming strategies have 60 collocate lemmas with nominal headwords of statistical significance. The highest proportional USAS semantic categories of nominal collocates are S and W, which will be explored below.

#### 5.5.2.1 Category S: social actions, states and processes

At 20.41%, USAS Category S contains the largest proportion of nominal collocates of the four naming strategies. These show a diversity of preference, with *victim(s)* notably containing a far larger number of S category collocates than *resident(s)*, especially taking into consideration the lower overall frequency of this node. *Survivor(s)*, the lowest-frequency naming strategy under investigation, also has more nominal collocates from this category than *evacuee(s)*, which has none at all. This is not to say that association with this category of collocates necessarily leads to identical discourses; in fact, quite the opposite in this case.

Naming strategy	S category nominal collocates of naming strategy
resident(s)	<b>native</b> [LL=242.03, MI=3.25]
victim(s)	<b>aid</b> [LL=590.32, MI=3.84], <b>assistance</b> [LL=430.81, MI=3.33], <b>counseling</b> [LL=247.93, MI=4.33], <b>benefit</b> [LL=225.69, MI=3.05], <b>helping</b> [LL=73.72, MI=3.99], <b>advocate</b> [LL=54.86, MI=3.16]
evacuee(s)	-
survivor(s)	<b>wife</b> [LL=186.21, MI=3.87], <b>daughter</b> [LL=169.47, MI=4.00], <b>son</b> [LL=143.95, MI=3.69]

Table 5-10: USAS S category nominal collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*

Through USAS S category collocates, *victim(s)* are construed as being the beneficiaries of social support, whether financial (e.g. *aid*, *benefit*), psychological (*counseling*), legal (*advocate*), or more general (*assistance*, *helping*). In the financial sense, it is most commonly government agencies or organisations offering aid; in the psychological sense, it is civic groups such as the Family Service of Greater New Orleans. *Assistance* is rarely offered by any known agent; *helping* is the only ‘social’ noun deriving from (named, individual) human actions. Of the 21 instances of collocation between *victim* and *advocate*, 11 can be attributed to mentions of the “Senior Victim Advocate Program”. As in the previous (and subsequent) chapters, proper, non-human nouns are discounted.

The nine remaining cases do not meet the minimum frequency threshold of 20 and are therefore discarded.

*Resident(s)* only collocates with one S category noun, and this tends to occur in a very restricted context. In 76 of the 89 resulting concordance lines, *native* appears with *resident* in New Orleans Times-Picayune obituaries. The identity of “a *native* and a **resident** of New Orleans” appears to hold particular salience for this publication during this period. *Native* does seem to have a particular relationship to reporting of death; in the COCA corpus, the #9 collocate in the newspaper section (with the given spans) is *died*. Even beyond this, however, there is a sense of legacy or connection to the Gulf Coast imbued with the addition of ‘lifelong’, which modifies *resident* in 45 of the 89 *native* concordance lines – see line 17 for an illustrative example. The distinction of having lived in the area for one’s entire life (from inception to culmination) is a potent quality of identity and it appears proudly included in last statements on a person’s life. This identity remains static not just in death, but also when concessions must be made to the disruption of residence in the wake of Hurricane Katrina (as in line 18 below). Therefore, the most salient [nominal] social quality of a *resident* is that of longevity and nativity, or being one of the weathered ‘us’ in the frame of the *New Orleans Times-Picayune*.

17. A *native* and life-long **resident** of New Orleans, passed on to the eternal life at her home in Gentilly, in the aftermath of Hurricane Katrina. (*New Orleans Times-Picayune*, 20 January 2006)
18. Louise “BaBa” Scie Raymond Alvarez passed away at the residence of her son, Gilbert on Wednesday, August 23, 2006 at the age of 83. She was a *native* and lifelong **resident** of New Orleans, LA until she was displaced by Hurricane Katrina. (*New Orleans Times-Picayune*, 28 August 2006)

**Figure 5-7: Sample concordance lines of the collocation between *resident(s)* and *native***

The pattern within S category collocates of *survivor(s)* is readily recognized: the frequent co-occurrence of the items *wife*, *daughter*, and *son* suggest that *survivor(s)* are systematically framed by their place within the family. Closer inspection reveals that this pattern is actually entirely relatable to the one above: *survivor(s)* nearly always occur in list form within obituaries, once more predominantly (but not entirely) in the *New Orleans Times-Picayune*. Each of the 50 concordance lines of *survivor(s)* with *wife* and 46 lines of *survivor(s)* with *daughter(s)* occur in this context (e.g. line 19). Only one case in 44 of the *son* concordance lines was not from an obituary, and both the *son* and the *survivor* were still alive at the time of reporting (see line 20 below). These *survivor(s)* are, in fact, not PAKs, but their family members, iterated.

19. In addition to his son, **survivors** include his wife, Thelma M. Trapolin; three daughters, Jane T. Oaksmith of Seattle , Patricia T. Couret and Anne T. Britt; 10 grandchildren; and four great- grandchildren. (*New Orleans Times-Picayune*, 17 December 2005)
20. The son of one **survivor** said his mother escaped with little more than a scare. (*New York Daily News*, 3 October 2005)

**Figure 5-8: Sample concordance lines of the collocation between *survivor(s)* and *son*, *wife*, and *daughter***

Generally speaking, this indicates that S category nominal collocates appear alongside *survivor(s)* and *resident(s)* in the restricted context of obituaries, though the emphasis differs. *Survivor(s)* describe the remaining family members of deceased PAKs, whereas *resident(s)* (PAKs themselves) are represented in death as having lived their entire lives in a single place. Both of these celebrate a certain ‘in’ quality – ‘belonging’ to a place or a family. This is in contrast to the story told by S category nominal collocates of *victim(s)*, who (while still alive), are construed as having the familiar lack of agency and receipt of benefits and assistance granted by others. In these senses, *resident* runs counter to the definitions provided at the start, as *native resident* is primed for the context of obituary rather than a status of continued living in one area. *Victim* is someone who has suffered through a terrible event, but who has then been the recipient of aid in the aftermath. Here, the sense of *survivor* is not the dominant definition, but the minority (mentioned in only one of three dictionaries) of the person who survives longer than others.

In the following section, I deal with the next highest proportional category of collocates – W, where once more, *victim(s)* has an unexpectedly high number of collocates in comparison to the other naming strategies.

### **5.5.2.2 Category W: world and environment**

This section deals with USAS category W (world and environment), which contains 18.37% of nominal collocates overall. This is the second highest proportional category of nouns, and the last category to be discussed under this part of speech.

Collocates from category W are very nearly all concentrated in the *victim(s)* row in Table 15 below, with *resident(s)* lacking collocates in this part-of-speech and semantic field, and the only items under *evacuee(s)* and *survivor(s)* being a collocate shared with *victim(s)*: *hurricane*. While this is grammatically a noun, in nearly all cases, it functions as an adjective, modifying the naming strategy with the attribute of their circumstances. This is a type of specificity, whereby they are not just *victims*, *evacuees*, and *survivors*, but specifically *hurricane victims*, *hurricane evacuees*, and *hurricane survivors*, and thereby

differentiated from (for instance) victims of domestic violence, earthquake evacuees, and breast cancer survivors. To wit, there is little to add in the discussion of this pattern (as well as *storm*, *flood*, and *hurricanes*) that is not realized across all of the other features, and I move on now to discussion of the other assorted collocates of *victim(s)*: *disaster*, *tsunami*, and *earthquake*.

Naming strategy	W category nominal collocates of naming strategy
resident(s)	-
victim(s)	<b>hurricane</b> [LL=23285.87, MI=4.82], <b>disaster</b> [LL=1436.14, MI=3.50], <b>storm</b> [LL=736.58, MI=3.48], <b>tsunami</b> [LL=646.30, MI=5.40], <b>flood</b> [LL=445.212, MI=3.087], <b>hurricanes</b> [LL=295.32, MI=3.93], <b>earthquake</b> [LL=166.80, MI=3.62]
evacuee(s)	<b>hurricane</b> [LL=3101.34, MI=3.04]
survivor(s)	<b>hurricane</b> [LL=2498.40, MI=3.91]

Table 5-11: USAS W category nominal collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*

In the concordance lines of these remaining collocates, four patterns prevail. The first (and most common) is the call to action, as illustrated in line 21 below. Hurricane Katrina is likened to other disasters, and stories are retellings of methods in which ‘average people’, organizations, or charities have rallied to provide aid to *tsunami*, *earthquake*, and *disaster victims*, and continue to do so following Hurricane Katrina. Two subsequent patterns are branches of this. The first (illustrated in line 22 below) is a negative contrast, indicating that *victims* of other natural disasters have in fact had more access to aid than PAKs, which is considered a social problem. However, limited access to aid by Americans is not the only priority. The third most frequent pattern of this set concerns calls to action over the South Asian earthquake, which is reported as having attracted less than a tenth of the aid of either *tsunami victims* or Hurricane Katrina (whose victims are marked explicitly as the ‘in’ group using the term “their own” in line 23). These reports are contradictory, of course, but show a complex picture of what sort of civic responsibilities are expected of the readership or implied by the circumstances of the publishing periods. The last collocate, *disaster*, appears in conjunction with *victims* as the great equalizer; it is with this collocation that the experiences of PAKs and those who have experienced natural disasters at home and abroad are united under common needs and desires, for instance, access to information (see line 24).

21. Madison first baked cupcakes to raise money for the tsunami [victims], and when she heard about the hurricane, she said she wanted to help out,” said her mom, Julie Minear. (*Atlanta Journal-Constitution*, 15 September 2005)
22. We can send massive amounts of aid to tsunami victims, but we can't bail out the city of New Orleans. (*New York Daily News*, 2 September 2005)



23. Private donations by Americans totaled \$13.1 million for *earthquake victims*, according to The Chronicle of Philanthropy. In comparison, Americans doled out \$1.3 billion for the *tsunami victims* and roughly \$2 billion for their own after Hurricane Katrina. The response to the South Asian earthquake has raised a troubling question: In the face of such calamity, why is the world not doing more? (*New York Times*, 14 November 2005)
24. *Disaster victims*, including those who survived hurricane Katrina, have repeatedly cited information as among their most immediate concerns, often over material relief. (*Christian Science Monitor*, 21 November 2005)

**Figure 5-9: Sample concordance lines of *disaster, tsunami, earthquake*, collocating with *victim(s)***

Nominal collocations belonging to the USAS W category occur to mark the experiences of disaster *victims*, and to both unite and divide them. A concern for the distribution of aid and duty of civic care dominates these concordance lines, regardless of which group is currently being touted for increased assistance. This pattern is unique to the *victim(s)* naming strategy, and a developing semantic preference for aid/help is clearly emerging with each successive part of speech and semantic category.

I move now to verbal collocates to see whether these preferences continue hold true. Like nominal collocates (5.5.2.1), the verbal section contains a high proportion of S collocates, and the M category from the adjectival section (5.5.1.2) will also be revisited.

### **5.5.3 Verbal collocates**

This section on verbal collocates is made up of four subsections of USAS broad semantic categories containing proportions of greater than 10% of all verbal collocates. See section 5.5.3.1 for verbal collocates from category H, section 5.5.3.2 for category M, section 5.5.3.3 for category S, and finally, section 5.5.3.4 for category X.

#### **5.5.3.1 Category H: *architecture, housing and the home***

Immediately, I must acknowledge a methodological issue which has resulted in the erroneous inclusion of this category at all. The collocation of *victim(s)* with *shelter* is relatively rare, occurring only 20 times. Unfortunately, three of these occurrences are mis-tagged as verbs and are, in actuality, nouns (see concordance line 25 below for an example). In each of these instances, the noun is plural, indicating that the CLAWS7 tagger has difficulty disambiguating the plural common noun from the -s form of the lexical verb. With these 3 cases discounted, this collocation drops below the minimum frequency allowed, and is discarded. With this single collocation removed, this has the knock-on effect, however, of dropping the entire USAS H category below 10% representativeness of the verbal collocates as a whole.

25. Buses were converging there to shuttle flood **victims** to emergency *shelters* in Texas and Louisiana. (*Wall Street Journal*, 2 September 2005)

Figure 5-10: Sample concordance line of *shelters* (plural common noun) mis-tagged as a verb

Naming strategy	H category verbal collocates of naming strategy
resident(s)	-
victim(s)	<i>house</i> [LL=704.58, MI=3.99], <i>shelter</i> [LL=56.359, MI=3.32]
evacuee(s)	<i>house</i> [LL=1985.29, MI=5.13], <i>shelter</i> [LL=457.075, MI=5.47], <i>occupy</i> [LL=126.303, MI=4.19]
survivor(s)	<i>house</i> [LL=116.86, MI=3.53]

Table 5-12: USAS H category verbal collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*

It is, therefore, with the most cursory mentions that I will say that the collocates in category H exhibit a greater degree of homogeneity than previous categories. Three naming strategies share one collocate (*house*), and two share another (*shelter*), though *shelter* only collocates meaningfully with *victim(s)* as a noun. Concern about the availability, cost [to taxpayers], and distribution of *housing* for PAKs dominates these concordance lines. Only one collocate is unique to a naming strategy – *occupy* only collocates with *evacuee(s)* – and only one naming strategy (*resident*) has no collocates from this category at all. The occurrence of *occupy* with *evacuee(s)* draws upon the source domain of warfare, likening PAKs to a military (or other invading) force that *occupies* space not fairly allotted to them. As this category has not appropriately qualified for close analysis, I am unable to go into more depth here.

The unreliability of part of speech tagging is thus found to be a severe drawback to the method employed here, and this has been taken under consideration when revising the method for subsequent chapters. See section 5.7.2 for more discussion on this subject.

Going forward, I discuss three categories of verbal collocates that do maintain statistical significance, starting with movement.

### 5.5.3.2 Category M: movement, location, travel and transport

Given the content of the texts comprising the corpus – the aftermath of a natural disaster – it is unsurprising that verbs related to movement and transport would be an important (and statistically significant) factor. What is perhaps less expected is the imbalance of distribution of these collocates between the various naming strategies; while *resident(s)* and *evacuee(s)* both show a grammatical-semantic preference for this category, neither *victim(s)* and *survivor(s)* have a single M category verbal collocate. This is interesting, because though *resident(s)* includes in its own meaning an element of stasis and ‘staying

put' that might rightly be modified by verbs of motion and transport, *evacuee(s)* as a naming strategy contains within itself a description of movement. To find out more about why these naming strategies might prefer M category verbal collocates despite these points, I now inspect the collocates in turn.

Naming strategy	M category verbal collocates of naming strategy
resident(s)	<b>return</b> [LL=2802.19, MI=3.69], <b>evacuate</b> [LL=1207.47, MI=3.69], <b>flee</b> [LL=365.78, MI=3.70], <b>scatter</b> [LL=183.24, MI=3.53], <b>staye</b> [LL=89.41, MI=3.51], <b>transport</b> [LL=68.48, MI=3.09], <b>displace</b> [LL=17.29, MI=3.94]
victim(s)	-
evacuee(s)	<b>arrive</b> [LL=621.80, MI=3.97], <b>stay</b> [LL=414.90, MI=3.10], <b>settle</b> [LL=153.22, MI=3.47], <b>transport</b> [LL=142.46, MI=4.47], <b>flee</b> [LL=132.14, MI=3.47], <b>relocate</b> [LL=125.72, MI=3.30], <b>scatter</b> [LL=117.85, MI=3.83]
survivor(s)	-

Table 5-13: USAS M category verbal collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*

There are more collocates shared between *resident(s)* and *evacuee(s)* in this category than collocates that are unique to either individual naming strategy; both are construed as *fleeing*, *scattering*, and *staying*. (Note the tagging error in *resident(s)*, where the headword of the lemma *stay* has been incorrectly clipped to *staye* – another point that will be raised in the summary at the end of this chapter.) Two of these verbs stand out as particularly negative: *fleeing* and *scattering* call to mind a more beleaguered participant than storm *evacuees* or *residents*. Even as actors, they are associated with a number of verbs that signify lack of volition. For instance, they find themselves *scattered* across the country. As objects, they are *transported* by emergency workers, and lack agency in their own evacuation.

In concordance lines, PAKs are both the agents of their own *scattering* (as in lines 26 and 27 below) and the patients of the *scattering*, either by some unnamed agent (perhaps the hurricane itself, as in line 28 below) or government leaders (line 29). It is interesting to note that *evacuee(s)* have a higher proportion of agency of movement here in comparison to *resident(s)*, who are more often *scattered* by someone or something else. The *scatter* itself construes PAKs as an aimless, chaotic diaspora; top collocates in the COCA corpus include *randomly*, *widely*, and *debris*. In terms of management, nobody (from federal government officials to local city councils tasked with tracking or intake) has any idea how to handle (see lines 26, 27, 28, and 29 below) this *scatter* of humans. Line 30 has been included as another example of mistagging on the part of CLAWS. Though *scattered* is considered a verb here (likely due to its past participle ending), this

appears in a list of attributes including “confused, anxious”, and is indeed an adjective likely derived from the verb. Again, these errors will be discussed in the summary, and have contributed to a reconfiguration of the methods for the later chapters.

26. The initial lack of cooperation on the local, state and federal levels made it impossible to track them all, and by now **evacuees** have *scattered* across the nation and the world. (*Newsweek*, 21 August 2006)
27. Hundreds of thousands of **evacuees** are *scattered* around Louisiana and neighboring states in a sudden diaspora, and no one seems to have any idea what to do with them next. (*Washington Post*, 6 September 2005)
28. It was also a largely poor, black community, many of whose former **residents** have been *scattered* around the country - a fact that has led to some charges of racism when planners suggest the whole Lower Ninth might need to be razed and turned into a wetland or park. (*Christian Science Monitor*, 12 October 2005)
29. For weeks, Jackson has condemned government leaders for *scattering* New Orleans **residents** throughout the country after Hurricane Katrina and not having a plan to bring them back to help with the rebuilding effort. (*New Orleans Times-Picayune*, 23 October 2005)
30. More than half the city might remain in moldering ruins; its **residents** *scattered*, confused, anxious; its commercial activity moribund; its political leadership adrift -- and yet the people of greater New Orleans at least have managed to come together and save Katrina Ridge. (*Pittsburgh Post-Gazette*, 11 December 2005)

Figure 5-11: Sample concordance lines of the collocation between naming strategies and *scatter*

*Fleeing* is another verbal collocates held in common between *resident(s)* and *evacuee(s)*, and many now-familiar patterns arise in these concordance lines, as well. *Fleeing evacuees* strain relief systems in states where they have had to relocate (line 31 below). En masse, they are construed as water metaphors, as in a wave (line 32) or a mass that can clog a system (line 33) – an metaphor referring not just to water, but to dirty water backing up the smooth flow of a normal system. A new method of enumerating *evacuees fleeing* also presents itself in these concordance lines; in 34 below, they are referred to as “hordes” who stymie attempts to “get back to business”, and are contextualized as an overwhelming, invading, threatening mass, perhaps linked to the military *occupy* seen earlier. Finally, in lines 35-36, a different primed meaning of *flee* is referred to, as *residents* and *evacuees* are said to be *fleeing* from the law. They have taken Katrina as an opportunity to escape their “criminal records” (35) or to enact crime in their new locations (36). *Fleeing*, in the context of the Katrina corpus, is a highly negative action associated with threat to social order and criminality.

31. Georgia relief agencies, straining to serve thousands of storm **evacuees** who have *fled* here, took a step Friday toward working more closely together. (*Atlanta Journal-Constitution*, 3 September 2005)

32. When Hurricane Katrina broke the levees in New Orleans, it did more than create a wave of evacuees fleeing the city. (*Washington Post*, 22 December 2005)
33. "He drove with his wife, daughter and four friends to Mobile; with roads clogged with other residents fleeing, what is normally a three-hour drive took 12. (*New York Times*, 30 August 2005)
34. Efforts to get back to business are being stymied by the hordes of evacuees who fled to family members, hotels or emergency shelters. (*Wall Street Journal*, 6 September 2005)
35. I think some saw [Katrina] as an opportunity," Martin's bounty-hunting partner, Michael Wright, said of evacuees who fled New Orleans with criminal records. (*Chicago Sun-Times*, 15 August 2006)
36. It had been blocked by West Bank law enforcement agencies who viewed fleeing New Orleans residents as potentially dangerous looters. (*New Orleans Times-Picayune*, 8 November 2005)

**Figure 5-12: Sample concordance lines of the collocation between naming strategies and flee**

While *stay* is a collocate of both *resident(s)* and *evacuee(s)*, this is a good example of the differing prosodies that might arise upon closer inspection of concordance lines. In both cases, discussions relate to the dictionary definitions of these words; *residents* are described as *staying* in their houses during Hurricane Katrina, whereas *evacuees* *stay* in with friends, in shelters, etc. after Hurricane Katrina. Due to this difference, *staying* is far more negatively evaluated in the case of *residents*, who endangered their own lives and the lives of rescue workers by doing so. Compare line 37 below, where these *residents* experience a "sick feeling of panic" *staying* in their homes to line 38, where items are being prepared to give to *evacuees* *staying* northwest of New Orleans. The government is ultimately assigned responsibility for where both *residents* and *evacuees* are *staying*, however, regardless of whether this was in New Orleans during the storm (as in line 39) or afterwards (40), a symptom of risk society.

37. The sick feeling of panic began to rise up in the **residents** who had *stayed* in the projects during the hurricane. (*New York Times*, 12 September 2005)
38. The women are packing an 18-wheeler, set to leave Sunday, with hair care products, toys, clothes, suitcases, and other items to give to thousands of **evacuees** now *staying* in Lafayette, La., some 135 miles northwest of New Orleans. (*Boston Globe*, 9 September 2005)
39. The report also shows how poor communications prevented information about New Orleans from reaching the White House and the upper levels of DHS in time; how the failure to declare a "mandatory" evacuation of the city meant that many residents stayed in New Orleans; and how the absence of a state database of emergency shelters led to confusion. (*Washington Post*, 14 February 2006)

40. A government program that has been paying for tens of thousands of Hurricane Katrina evacuees to stay in hotel rooms around the nation while they look for other housing must be extended to Feb. 7, a federal judge ruled Monday. (*New Orleans Times-Picayune*, 13 December 2005)

Figure 5-13: Sample concordance lines of the collocation between naming strategies and *stay*

With all collocates shared between the two naming strategies touched upon, I will now wrap up discussion on those collocates unique to one naming strategy or the other. It is clear why *evacuate* is a unique collocate of *resident*: used in tandem with *evacuees*, this word would be repetitive. Previously, I have explored a variety of similar verbs meaning “to leave” which co-occur with *evacuees*. The relationship between *evacuate* and *residents* is very similar to these, insofar as the PAKs themselves only have agency in *evacuating* in slightly less than half of the resulting concordance lines. More often, they are evacuated by those with more social power, e.g. government officials or members of the military. Failures on the part of authority figures to effectively and successfully *evacuate residents* are seen as failures of the larger ‘system’ or even the state, demonstrative of risk society discourse. Unlike *flee*, *stay*, *return*, *arrive*, *displace*, and *relocate*, even when *residents* are the agents, they are the subjects in infinitive clauses that occur at the behest of others, for instance, the governor’s urging in line 41 and [lack of] forcing in line 42. Lack of sufficient institutional force in instigating this *evacuating* process on the part of *residents* is seen in line 43 to result in federal accusations of endangering lives. Failure to cooperate, or to somehow inhibit the agency of these *residents* on behalf of government officials is also described as criminal in line 44. Overall, the interplay in agency is extremely interesting in this collocation. Whether it is *residents* *evacuating* or *being evacuated*, it is always institutional bodies that assume agency or responsibility for the final outcome, and receive criticism for the failure of a completed task.

41. Authorities charged the husband and wife owners of a nursing home Tuesday with negligent homicide in the deaths of 34 elderly residents who died in the floodwaters from Hurricane Katrina. Louisiana Attorney General Charles C. Foti Jr. said the owners of St. Rita’s nursing home in Chalmette, just east of New Orleans, failed to *evacuate residents* despite repeated warnings. (*Atlanta Journal-Constitution*, 14 September 2005)
42. Louisiana’s governor urged coastal **residents** to *evacuate*. (*Wall Street Journal*, 23 September 2005)
43. The mayor of New Orleans and the governor of Louisiana, who had come to Capitol Hill yesterday to plead for more federal hurricane aid, faced Republican accusations that they had endangered lives by failing to force **residents** to *evacuate* days before the landfall of Hurricane Katrina. (*Boston Globe*, 15 December 2005)

44. When will Arthur Lawson, chief of the Gretna Police Department, be arrested along with the others who assisted him in trapping New Orleans residents trying to *evacuate* the city by foot -- the only means they had -- during a living hell that was not of their making? (*New Orleans Times-Picayune*, 24 September 2005)

Figure 5-14: Sample concordance lines of the collocation *resident(s)* and *evacuate*

Concordance lines in which *evacuees* are just *arriving* are as negative as any of the patterns witnessed before, e.g. they are enumerated, construed as threats, invasions and floods. However, once they *relocate* and *settle*, these negative prosodies die down, and they are accepted as part of the new status quo. *Residents* who are *displaced* by the storm are discussed in terms of duration and location of temporary housing, but the resolution of that status is not so neat as in the experience of *settled evacuees*. Analysis of the final collocate left in this category will end the section on a melancholy note. Though *resident(s)* collocates with *return* both strongly and frequently (847 hits in the corpus), this does not have the actual agency originally assumed. The concordance lines reveal a much more complicated relationship between the *residents* and their return – it is not one of agency and report, but of modality and prediction. In the weeks immediately after the storm, *residents* were advised not to *return* (line 45); in the months afterwards, they proved independently unwilling or unable to do so (lines 46 and 47). The population of New Orleans was extremely slow to recover, and has still not fully done so. The 2010 Census revealed that the population had dropped 27% in 10 years, and this is attributed to the storm (Plyer, 2011).

45. "It's still not yet safe for all of these **residents** to *return* to their homes, although we're making great progress," Guidry said. (*New Orleans Times-Picayune*, 17 September 2005)
46. In New Orleans, most **residents** had not *returned* after evacuating for Katrina, so the reflooded houses were almost all uninhabited. (*Atlanta Journal Constitution*, 25 September 2005)
47. Perhaps the biggest unanswered question is whether former **residents** will *return*. (*Christian Science Monitor*, 12 October 2005)

Figure 5-15: Sample concordance lines of the collocation between *resident(s)* and *return*

In this section, I found that neither *survivors* nor *victims* are construed in terms of their movement or transport, which was counterintuitive, given the nature of the corpus. Both *residents* and *evacuees* collocate with a large number of USAS M category verbs, and upon initial inspection, many of these seemed to indicate agency. However, closer analysis revealed that PAKs are the agents in less than half of the verbs in this category,

and even in sentences where they are affecting actions, often powerful, instructional bodies are ultimately steering or otherwise being credited for these processes.

It will be interesting to see whether this lack of unmitigated agency holds in the next section, verbs of social actions and processes.

### **5.5.3.3 Category S: social actions, states and processes**

USAS Category S has the highest proportion of verbal collocates overall, and once more, *victim(s)* shows a preference for this category, containing the greatest number of collocates from this semantic domain. Despite being the highest frequency naming strategy overall, *resident(s)* has the lowest number of social action verb collocates, reinforcing that the field of movement and transport is more salient to this identity. Due to the great number of collocates in this section and the high instance of overlap between naming strategies, I organize this section into various categories: collocates relating to rescue efforts (*strand, rescue*), collocates relating to social support (*encourage, welcome, honor*), collocates relating to social organization (*organize, rally*), and collocates relating to general assistance (*help, aid, benefit, assist*). Unlike other sections, these categories are of my own devising, rather than automatically generated. Though USAS could have grouped some collocates together, I found fine-tuning categories by hand to be more effective here in grouping collocates quickly and effectively.



Naming strategy	Collocates of rescue efforts	Collocates of social support	Collocates of social organisation	Collocates of general assistance
resident(s)	<b>strand</b> [LL=138.76, MI=3.88]	<b>encourage</b> [LL=354.80, MI=3.43]		
victim(s)	<b>rescue</b> [LL=137.72, MI=3.03]	<b>honor</b> [LL=70.20, MI=3.50]		<b>help</b> [LL=5109.03, MI=4.20], <b>aid</b> [LL=1706.14, MI=6.58], <b>benefit</b> [LL=1188.59, MI=5.10], <b>assist</b> [LL=838.94, MI=4.79]
evacuee(s)	<b>strand</b> [LL=73.92, MI=4.00]	<b>welcome</b> [LL=83.32, MI=3.15]		<b>assist</b> [LL=218.24, MI=3.59], <b>aid</b> [LL=97.16, MI=3.74]
survivor(s)	<b>rescue</b> [LL=200.51, MI=4.59]		<b>rally</b> [LL=116.93, MI=5.41], <b>organize</b> [LL=72.25, MI=3.95]	

Table 5-14: USAS S category verbal collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*

All four naming strategies collocate with words relating to success or failure of rescue efforts. Notably, *victim(s)* and *survivor(s)* are successfully *rescued*, whereas *resident(s)* and *evacuee(s)* are *stranded*, or more frequently associated with unsuccessful rescue attempts. This preference contributes to the negative prosody around residents and evacuees, as the state of being *stranded* and unable to be helped is compounded by emotional conditions of desperation as in lines 48 and 49 below, or inflated by the use of enumeration (line 49). Compare these to concordance lines of rescue below, wherein survivors (50) and victims (51) are the “focus” of federal attempts that “would move heaven and earth”, reinforcing both the vulnerability of these groups and the seeming obligation to relieve them of suffering.

48. Rescuers who ventured out to the hardest-hit areas reported **residents** stranded on rooftops and screaming for help. (*New York Daily News*, 30 August 2005)
49. The New Orleans Superdome, where thousands of hungry, desperate evacuees were stranded after the levees broke, is scheduled to be rehabilitated by Nov. 1, possibly in time for part of an NFL Saints season and the next Sugar Bowl. (*Pittsburgh Post-Gazette*, 29 December 2005)
50. Officials have said they were more focused on rescuing Katrina’s **survivors** immediately after the storm hit than in picking up bodies. (*New Orleans Times-Picayune*, 28 October 2005)

51. But he promised the United States would "move heaven and earth" to rescue **victims** of Hurricane Katrina, which he called "probably the worst catastrophe or set of catastrophes" in U.S. history. (*San Francisco Chronicle*, 4 September 2005)

**Figure 5-16: Sample concordance lines of the collocation between naming strategies and *rescue/strand***

The next category, which I have called 'collocates of social support' includes the collocation of *encourage* with *resident(s)*, *honor* with *victim(s)*, and *welcome* with *evacuee(s)*. None of the resulting concordance lines expose patterns not previously detailed here. In concordance lines of residents are encouraged to complete tasks – such as seeking information and returning home – by government officials (but are not shown actually completing them). Once more, the focus is on the institution rather than the agency of the PAKs. *Honor* collocates with *victims* due to the possible 'deceased' sense of this word, nearly always occurring as a result of a story on City Council dedicating a granite monument to *honor* Katrina's *victims*. Finally *evacuees* are welcomed in their new homes, notably in 29/32 cases these are without count, indicating that an unknown (or small, as in the other cases) number of *evacuees* can be welcomed, whereas the large numbers (or hordes) seen before are construed as threats.

As previously reviewed, *survivor* does share one verbal collocate (*rescue*) in common with *victims*, its two other collocates (*rally* and *organize*) are unique. It initially appeared that *rescue* would relate to *survivors* in the literal, physical sense (storm survivors) while verbs like *rally* and *organize* would create a prosody closer to the more metaphorical, emotional sense (*being* a survivor as a personality trait). However, both *rally* and *organize* occur in 21 New Orleans Times-Picayune texts duplicating an announcement of an upcoming event—"KATRINA SURVIVORS RALLY IN WASHINGTON, D.C. -- The Feb. 9 gathering will demand more help from lawmakers in the recovery and rebuilding of New Orleans and the Gulf Coast. The ACORN Katrina Survivors Association is organizing the caravan and events." In fact, *rally* has been mistagged as a verb. As a result, it can be said that *survivors* are afforded no extra agency, volition, or positivity—a surprising finding based on the dictionary definition of the word and differentiation from the anticipated outcome.

The last and most broad subcategory – what I call collocates of general assistance – includes *help*, *aid*, *benefit* and *assist*, in their collocational relationships with *victims* and *evacuees*. Once more, these verbs place PAKs in the passive position, as the beneficiaries of charity and aid. Over 95% of the time, PAKs receive the verbs rather than enact them.

This pattern has become typical of the verb categories overall. Pervasive lack of verbal agency reinforces the helpless/hopeless prosody built around each of the four naming strategies, and additionally removes them from social power by neglecting to portray them as interacting with, rather than being acted upon.

The final verb category of collocates, X, further reinforces these findings.

#### 5.5.3.4 Category X: psychological actions, states and processes

The final category of verbal collocates to be reviewed here is X: psychological actions, states and processes. Items from this category display a somewhat novel spread, with no common collocates and each being unique to a particular naming strategy. *Resident(s)* has the highest semantic preference for this group with three collocates, and *evacuee(s)* does not collocate significantly with any items at all. Close analysis of each collocate in turn reveals a continuation of denial of agency on the part of PAKs, and a focus instead on the actions of aid workers attempting to assist them.

Naming strategy	X category verbal collocates of naming strategy
resident(s)	<b>urge</b> [LL=1155.95, MI=4.57], <b>heed</b> [LL=119.80, MI=4.52], <b>resolve</b> [LL=88.34, MI=3.08]
victim(s)	<b>identify</b> [LL=196.81, MI=3.48]
evacuee(s)	-
survivor(s)	<b>search</b> [LL=264.98, MI=5.64]

Table 5-15: USAS X category verbal collocates of *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*

I begin with collocates of *resident(s)*: *urge*, *heed*, and *resolve*. In 255/260 concordance lines, they are *urged* to evacuate, clear debris, or be patient in returning home, as contrasted to only five concordance lines in which *residents* urge city officials to be more mindful of their needs, e.g. “to create a taxing district in the subdivision in order to raise money to buy the land and set it aside for use as a floodwater reservoir” (*New Orleans Times Picayune*, 20 October 2005). The remaining two unique collocates are a bit more complex as far as the relationship between agency and social power. In 33/35 concordance lines, *residents* are receiving help to *resolve* hurricane-related mortgage and insurance difficulties, again sharing agency with relief organisations and government workers, rather than being construed as *resolving* issues on their own. In the case of *heed*, *residents* are the ones who are responsible for this behavioural action. However, *heeding* is a deferential reaction to instructions provided by powerful institutional bodies, and therefore does not carry the same weight of agency. In only one of 27 concordance lines are *residents* not the ones doing the *heeding*; contrast lines 52 and 53 below (where the compliance of *residents* is being evaluated positively and

noncompliance is reported upon with deference) to line 54, where the *residents* themselves are (unusually) being *heeded*.

52. Officials estimate 80 percent of New Orleans' **residents** *heeded* the call to evacuate before Katrina roared ashore. (*Denver Post*, 30 August 2005)
53. **Residents** didn't *heed* warnings - they vowed to stay and fight rather than evacuate. (*New Orleans Times-Picayune*, 1 October 2005)
54. *Heeding* **residents'** concerns about construction projects near levees, Jefferson Parish planning officials added a provision to their proposed rules for high-rise housing that requires condominium towers to stand at least 50 feet from the bottom of any levee. (*New Orleans Times-Picayune*, 12 April 2006)

**Figure 5-17: Sample concordance lines of the collocation between *resident(s)* and *heed***

Lack of agency is ubiquitous in all other concordance lines resulting from collocations in this category. Of 45 concordance lines resulting from the collocation between *survivors* and *search*, in only 3 cases are survivors doing the searching, always for family members and friends, as illustrated in line 55 below. *Survivors* are far more commonly the objects of *search*, again placing more emphasis in this context upon the rescue teams (hailed as 'heroes' in line 56) than the PAKs themselves. Not a single one of the 65 concordance lines of *identifying* show the *victims* as agents of this action. Likely due to the dictionary definition of this word and relating to its sense as someone who is deceased, *identify* occurs with *victim(s)* only in the context of ascertaining the identities of bodies, as in line 57 below.

55. Its workers have created Web sites that allow survivors to search and connect with family and friends. (*Seattle Post-Intelligencer*, 8 September 2005)
56. Amid the unimaginable, the nation needed as many heroes as it could muster, so the rescue workers who pounded through the rubble *searching* for **survivors** day and night were hailed as well. (*The Boston Globe*, 6 September 2005)
57. A ChoicePoint-owned company that uses human DNA samples to determine identity is involved in the grim task of *identifying* **victims** from Hurricane Katrina. (*Atlanta Journal-Constitution*, 9 September 2005)

**Figure 5-18: Sample concordance lines of *identify* and *search***

As was the case in the S category analysis, collocation with verbs in the USAS X category compounds lack of agency on the part of *residents*, *victims*, and *survivors*, and deemphasizes the experience of PAKs in favour of emphasizing the agency of institutional bodies and their role in post-disaster management. Particularly throughout the verbal collocates categories, it has been increasingly difficult to pinpoint construal of

PAKs themselves beyond the scope of the actions of those around them upon them, as actions by them are very infrequently reported upon.

The next section is not strictly a part-of-speech category, but is related to some of the processes that PAKs did undertake, though incongruently. These have to do with metaphorical processes, namely those of water and volume metaphors.

#### **5.5.4 Water and volume metaphors**

Due to their proliferation throughout many subsections of this chapter, a separate section has been devoted here to discussion of water and volume metaphors. These are linked to both the pattern of enumeration of the negotiation/construal of threat. Not all of these were exposed through quantitative views of collocation tables; some were identified through search strings, combining once more quantitative and qualitative aspects of analysis.

Analysis of verbal and nominal collocates exposed a high number of water metaphors used to reference human actors. This argumentative strategy combines high counts with intimidating power and refers to elemental, uncontrollable forces leading to some sort of devastation or bursting. These metaphors are particularly employed in the case of *evacuees*, who *pour* into neighbouring states, 'makeshift shelters', and the Astrodome, or even *trickle* in, the local populations who struggle to *absorb* them.

As illustrated in the sample concordance lines below, 'waves' of evacuees are associated with shortage of resources and a certain degree of dread or dismay. When evacuees "pour" into cities and shelters, local officials and residents express an inability to care for outsiders, or at least a disinterest in shouldering a financial burden on their behalves.

58. **Evacuees poured** into Houston's Astrodome and other shelters, but nobody in Texas' largest city "knows for sure how many people are here or how long they intend to stay," says Frank Michel, a city spokesman. (*USA Today*, 8 September 2005)
59. Just weeks after Katrina **evacuees poured** across the state line, Georgia officials started to worry that FEMA would end up sticking the state and local governments with the task of housing evacuees, and with unpaid bills, according to e-mails and documents obtained through the state's Open Records Act. (*Atlanta Journal-Constitution*, 26 February 2006)
60. When Hurricane Katrina broke the levees in New Orleans, it did more than create a *wave* of **evacuees** fleeing the city. Democracy itself is now a disaster area. (*Washington Post*, 22 December 2005)

61. Each chopper brought more people. This second wave of **evacuees** was different from the group that showed up Sunday. Many of them had been rescued from a rooftop or plucked from the floodwaters. (*New Orleans Times-Picayune*, 28 August 2006)
62. He expects another wave of **evacuees** to hit Georgia shortly. These people have been staying in motels nearby and are running out of money. "Someone needs to channel people into appropriate housing," he said, "to save the emergency crisis beds for the next wave." (*Atlanta Journal-Constitution*, 2 September 2005)

**Figure 5-19: Sample concordance lines showing metaphorical use of *pour* and *wave***

Water metaphors extend beyond the initial "waves" of evacuees moving over and into surrounding areas; once relocated, they must then 'absorbed' by new locales – *absorb* collocates strongly with *evacuee(s)* (MI 3.79, LL 68.41). Absorption is often doubly marked with the vocabulary of volume; nearly every social actor below is modified by size or count ("most", "more", "a few", "25,00", "so many", "huge influx"). When viewed as a set of concordance lines, evacuee absorption is construed as a sort of contest, with communities claiming to have done their part in taking in "so many" PAKs, and using this to justify a sentiment of "no-vacancy" (line 66) or not "want[ing] you here anymore" (line 64). Absorption is a "strain" on local resources and economies, and communities are described as "struggling" to accommodate Americans who come from "pathetic and dangerous places" on "society's margins", finally asking where "the people nobody wants go" (line 68).

63. Outside the disaster zone, Georgia *absorbed* the most **evacuees** of any state other than Texas. It was as if one year's worth of new people showed up in one week. (*Atlanta Journal-Constitution*, 26 February 2006)
64. But Thomas almost touched off another controversy when he told the crowd that Houston -- which *absorbed* more hurricane **evacuees** than any other city in the country into housing, schools and jobs - - doesn't "want you here" anymore. (*New Orleans Times-Picayune*, 7 March 2006)
65. An epic migration not seen since the time of *The Grapes of Wrath* forced many cities and states to *absorb* **evacuees** workers into their economies. (*Chicago Sun-Times*, 21 December 2005)
66. Distant places such as Evansville might be able to *absorb* a few **evacuees**, but an overwhelmed Houston Astrodome is flashing the no-vacancy sign, sending 25,000 to Reunion Arena in Dallas, another 25,000 to a San Antonio warehouse that was once an Air Force base. (*USA Today*, 6 September 2005)
67. Baton Rouge Mayor Melvin "Kip" Holden told a Baton Rouge newspaper that the city would refuse to reopen an arena used as a major Katrina shelter. Baton Rouge is struggling to cope with the huge influx of New Orleans **evacuees** it *absorbed*. (*Wall Street Journal*, 6 May 2006)

68. Sections of New Orleans were pathetic and dangerous places even before the horrific images of squalor and lawlessness were beamed to the nation's television sets. About a quarter of New Orleans' 480,000 *residents* fall below the poverty line. Almost all are black. All must now live elsewhere. Even in good times, this population strains schools, social services and law enforcement agencies. Now some vexing questions are being asked. Which communities will *absorb* these **evacuees** for four, six, or eight months? Or even forever? What kinds of support --- and resistance --- will occur when communities are faced with accommodating those on society's margins? In short, where do the people nobody wants go? (*Atlanta Journal-Constitution*, 4 September 2005)

Figure 5-20: Sample concordance lines showing metaphorical use of *absorb*

The water metaphor holds exceptionally strong across months and publications; the narrative continues beyond “absorption” to “swelling”, evoking the containment *topos* of bursting or exploding.

As was the case with the *absorption* concordance lines above, the concept of *swelling* with *evacuees* is often magnified with tandem counts; swollen populations have “nearly triple[d]”, or gained “200,000 evacuees” in a “substantial impact”. Line 70 describes the number of *evacuees* inside the Dome as swelling to “dangerous levels”, as though the structure were on the verge of bursting from internal pressure. Where *evacuees* go, so too go extended water metaphors: “hotels are bursting” and “brimming with relief workers”, and local hotels and shelters are full (line 72).

69. *Swollen* from **evacuees** from Katrina, with a population nearly triple the 350,000 residents who lived there pre-Katrina, Baton Rouge has become a boomtown almost overnight. (*Wall Street Journal*, 17 September 2005)
70. Soon, floodwaters around the Dome rose, and the number of **evacuees** inside swelled to dangerous levels. (*New Orleans Times-Picayune*, 13 October 2005)
71. Though other communities suffered far worse damage than Norco, the community's population was *swollen* by **evacuees**. Many of them had lost everything. (*New Orleans Times-Picayune*, 11 August 2006)
72. The evacuees are making other kinds of substantial impact wherever they've landed: Baton Rouge, which had a population of 228,000, swelled by 200,000 evacuees, instantly surpassing New Orleans as Louisiana 's largest city. Hotels are bursting. Roads are clogged with traffic. (*USA Today*, 6 September 2005)

Figure 5-21: Sample concordance lines showing metaphorical use of *swell*

Above, an additional invocation of taint or filth dovetails with the water metaphor of mass force (e.g. “influx”) when humans are implied to be dirty water or detritus, “clogging” or “jamming” the roads with traffic as they disperse. A selection of extended concordance lines for these collocates reveals that the movement of *residents* away from

their homes and out of the path of the storm is often characterised as a sludge or unclean mass disrupting the normal flow of traffic. In line 76, the “influx of residents” will not only “clog the roads” but will also “hamper the cleanup” with their unwanted presence, making it “not safe to come in”.

73. Thousands of Houston **residents** *jammed* the already *clogged* roadways and created a Texas-sized traffic jam. They were fleeing a storm that never really hit Houston, and sadly, some died in hot Interstate 45 traffic jams. (*St. Petersburg Times*, 21 May 2006)
74. Coastal **residents** *jammed* one-way interstates and gas stations Saturday as they rushed to get out of the way of Hurricane Katrina, which threatened to gain even more strength and make a direct hit on the New Orleans area. (*Washington Post*, 28 August 2005)
75. Highways around Houston were *clogged* bumper to bumper with **evacuees** for up to 100 miles north of the city. The city’s two airports were *jammed* with **residents** seeking scarce seats on outbound flights. (*Boston Globe*, 23 September 2005)
76. Most hospitals have closed, and police communications are spotty. An influx of residents will *clog* the roads and hamper the cleanup, officials say. “It’s not safe to come in,” said Port Arthur Police Officer Janice Marshall. (*Pittsburgh Post-Gazette*, 27 September 2005)
77. “We need a ton of help. We could use National Guard units,” said one Biloxi officer who said he was not authorized to give his name. He said the city could use people to direct traffic on the streets increasingly *clogged* with **residents**, work crews and gawkers. (*Washington Post*, 31 August 2005)

Figure 5-22: Sample concordance lines showing metaphorical use of *swell*

Not all water metaphors used within the corpus are negative. Those that do not correspond to mass counts often construe *residents* and *evacuees* quite positively. Rather than being ‘swollen’ with numbers or inclined to defend against ‘floods’, when (seemingly ‘manageable’) small numbers of Katrina evacuees *trickle* or *stream* back, they are anticipated and even welcomed.

Note the return to normalcy signalled by the slow return of residents in lines 80, 82, and 83, where the trickles and streams are associated with positive elements such as future growth, safe homes, and “things...looking up”. Concordance lines 78 and 85 indicate an eagerness for a steadier stream, and while lines 79 and 83 also include references to clogging/flooding, they are centred on the mayor’s *invitation* (not permission) to return and “intact and dry homes”. Houston vows to treat the stream of *evacuees* with “great care” in line 81, and returning evacuees and newspaper writers alike praise the organisation and positive behaviour of the trickle in line 84.



78. As **residents stream** back into New Orleans and its suburbs, the question that rolls off every tongue -- right after "How'd you do in the storm?" -- is whether far-flung friends and neighbors also will return or whether the region has been permanently shrunk in the wash. (*New Orleans Times-Picayune*, 23 October 2005)
79. Earlier in the day, as **residents** began *streaming* in at the mayor's invitation, cars were backed up for two hours at an Interstate 10 checkpoint into the city. (*St. Petersburg Times*, 20 September 2005)
80. But Superintendent Gayle Sloan defended the temporary campus, pointing to expectations of future growth as Slidell **residents stream** back to the area. (*New Orleans Times-Picayune*, 21 January 2006)
81. Like other major American cities, Houston routinely has too few hospital beds for its own residents. But as thousands of **evacuees streamed** in last week, city leaders here vowed to treat them all with great care. The goal: to rapidly deploy doctors, nurses, paramedics and support workers with enough medical gear to treat every desperate storm *victim* while preserving the ever-strained hospital system for Houstonians. (*USA Today*, 8 September 2005)
82. As Waggaman **residents** began *trickling* back, cleaning up their homes, clearing debris, mowing their lawns and making repairs, things started looking up. (*New Orleans Times-Picayune*, 20 October 2005)
83. Upriver in St. Charles and St. John the Baptist parishes, the first wave of **evacuees** began *trickling* back to check out what one radio-station caller described as generally intact and dry homes. (*New Orleans Times-Picayune*, 1 September 2005)
84. "The return was excellent," said the tobacco company representative, 24. Some of the estimated 2.5 million **evacuees** began *trickling* back home to Houston and surrounding Gulf Coast communities on Sunday. Despite intense scrutiny by local officials and media to see how drivers would fare during the reentry, there were no headline-grabbing incidents. (*Washington Post*, 26 September 2005)
85. It's unclear when **evacuees** will begin *trickling* back in, or how long it will take for basic utilities to begin functioning again. But those still here yesterday seemed willing to tough it out. (*Boston Globe*, 6 September 2005)

Figure 5-23: Sample concordance lines showing metaphorical use of *stream* and *trickle*

Use of water metaphors in this context harkens back to "crime news stories and U.S. government reports during the nineteenth and early twentieth centuries [which] commonly used water metaphors to describe 'waves' of 'not-yet-white' people migrating to the United States to compete for employment, resources, and opportunities", signifying external threat "by 'aliens'" (Lacy & Haspel, 2011, p. 28). Therefore, metaphors of large amounts of water to describe PAKs "imply that they are a large alien force that would be competing against residents from other U.S. cities and states for resources" (Lacy & Haspel, 2011), an extreme case of othering.

## 5.6 Summary

### 5.6.1 Of findings

Adjectival collocates of counts and enumeration indicated that *resident(s)* and *evacuee(s)* were often construed as a threat or strain as a large and unmanageable mass. *Resident(s)* and *evacuee(s)* co-occur with nouns from the USAS S category in the restricted context of obituaries, framing them as belonging to a place or to a family, whereas [living] *evacuee(s)* are the beneficiaries of *assistance* and *aid*. With nominal W collocates, *residents* affected by Katrina were contrasted to *victims* of other natural disasters, a pattern not realized by other identities. It was, however, verbal collocates which were the most telling in distinguishing various naming strategies from one another. All are the objects of rescue efforts, but in the case of *residents* and *evacuees*, these fail, and with *survivors*, and *victims* they are successful. *Evacuees* and *victims* are the beneficiaries of *aid*, whereas *residents* receive *urging* to *heed* advice. PAKs enjoyed agency in less than half of the concordance lines arising from verbal collocations, and actions undertaken construed them as threats, invasions, and floods.

Ultimately, analysis of PAKs in this chapter proved particularly challenging; the attributes and actions of *residents*, *victims*, *evacuees*, and *survivors* are always corrupted (and indeed overwhelmed) by descriptions of rescue workers, government officials, and other, more powerful figures with greater agency. As soon becomes apparent, this is a feature of 'act of God' reporting in general.

### 5.6.2 Of method

I employed a method of downsampling results in this chapter that relied upon automated part-of-speech tagging using the UCREL CLAWS7 system, followed by automated semantic tagging by UCREL USAS and manual disambiguation/error correction. Collocates of four of the most frequent naming strategies for PAKs were categorized by part of speech and cross-tabulated by USAS broad semantic category. Only those categories representing more than 10% of the overall number of collocates of that part of speech were analysed qualitatively. This method had strengths and significant weaknesses.

Most helpfully, this method allowed for close analysis of a manageable number of collocates. These were objectively defined, rather than subjectively selected from a hundreds-long list of candidates. Grouping collocates by part of speech did make

comparing naming strategies somewhat easier, particularly in cases of attribution and transitivity. However, I believe that ultimately the weaknesses of this method have outweighed these strengths.

Tagging errors plagued a number of subsections of this chapter, most notably making the difference between inclusion and exclusion of an entire category that bordered on the minimum proportional threshold. Categorising by part of speech on the first pass places a very high degree of reliance on the machine for accuracy, and this failed throughout the chapter, particularly with adjectives ending in *-ed* and nouns ending in *-s* being mis-tagged as verbs, or vice versa. Separation of lemmas derived from the same word into different part of speech categories (e.g. *displaced* as an adjective and *displace* as a verb) did not ease analysis, and in fact made description more complex. Some automated categorization seemed arbitrary, for instance *thousand* as a noun, and *1,000* as an adjective. Categorising all items by semantic category in the first instance would have allowed more primary, dominant patterns to emerge, upon which part of speech analysis could have been applied. It is this course of action that I will take in AIDS/HIV chapters 7, 8, and 9.

In the next chapter, the method is revised for diachronic analysis of a single (contested) naming strategy—*refugee*.

## Chapter 6: Katrina's 'refugees'

### 6.1 Introduction

In the days and weeks following Hurricane Katrina, journalists and other public figures employed a wide variety of naming strategies to refer to American residents affected (and often displaced) by the storm. One in particular stirred an explicit controversy in the media:

In the days that followed the disaster on the Gulf Coast, once the scope of the ongoing tragedy became apparent, there arose a great hue and cry over using the word "refugee" to describe the thousands who fled the roiling waters kicked up by Hurricane Katrina. To some, because the vast majority of those who escaped were black, "refugee" took on a racist connotation. It's a term, they argued, that has been used to describe people from other countries -- Somalia, Rwanda, Vietnam, etc. -- where people fled to a different nation to escape oppressive governments or hardships imposed by poverty and hunger. So, the critics of "refugee" pointed out, to say the same of Americans who fled a natural catastrophe within their own country is to make them second-class citizens, and thus is racist. ("Having said that...", *News and Observer*, 17 September 2005)

As explored in previous chapters, terms such as *resident*, *evacuee*, *victim*, and *survivor* each carry unique connotations. The functional selection of any one over the other is a discursive choice constructing a certain meaning (Reisigl & Wodak, 2001; van Dijk, 2006; van Leeuwen, 2005). In most cases, this underlying meaning remains latent. However, one naming strategy—*refugee*—broke this pattern, turning cultural understanding of the word's connotations into a public discussion on the denotations of its use in the context of Hurricane Katrina.

Arrival at this naming strategy as a point of analysis came via a different route than those in Chapters 4 and 5; *refugee* has been selected specifically due to my own memory of its contested use in the weeks and months following Hurricane Katrina. Therefore, this chapter can be considered more in the vein of corpus-based analysis than the corpus-driven analyses that have come before. Specifically, the following research question will be addressed:

3. What evidence does corpus-based discourse analysis of the two corpora give of **changing or resistant discourses surrounding othered social groups over time**?
  - a. What kind of cultural changes in the broader socio-cultural context are indicated by these changes?

In the first week of September 2005, African-American civil rights activist and Baptist minister Rev. Jesse Jackson declared, “It is racist to call American citizens refugees” (“‘Refugee’ term stirs fight; Jesse Jackson says it’s racist – others say it fits dire situation”, *Chicago Sun-Times*, 7 September 2005). Similar outcries arose from politicians, journalists, and so-called ‘refugees’ themselves, as I will explore in this chapter.

Unlike the other naming strategies explored, uncontested use of *refugee(s)* dropped dramatically in frequency as it polarised journalists and introduced a meta-level, critical discussion of historical precedents, and socio-cultural semantics. Early on in corpus articles, perhaps partly “[b]ecause the concept of the ‘refugee’ summoned a sense of vulnerability that contradicted everything the United States—and its citizens—supposedly stood for, it was soon dropped in favor of some allegedly more suitable designations” (Masquelier, 2006, p. 736).

In this chapter, I will explore the print media’s use of the word in various contexts over the year following Hurricane Katrina. Discussion will begin with an examination of dictionary definitions of the term, which will be followed by frequency statistics, categorisation methods and schemes, analysis of the node word in various contexts, and finally, a diachronic overview of the differing patterns of usage and their interconnection will be introduced.

### **6.1.1 Dictionary definitions of term**

Before proceeding to analysis of *refugee* in the Katrina corpus, it would be helpful to establish definitions of the word from a broader social context. These are included not as a benchmark of the ‘correct’ way of using a term, but to identify the way its use is designated by standard-makers, by contrast to use in the press. Three dictionary definitions are offered below, mirroring those set out in the previous chapter. It must be noted that these definitions were collected *before* Hurricane Katrina, and there is some evidence that the event influenced lexicographic choices in the years that followed—an argument which will be reflected later in this chapter. Here we can see that of the three

dictionaries, both Longman and the New Oxford stipulate that a *refugee* must leave his or her country, whereas Random House simply suggests that this a strong possibility.

	Longman Dictionary of Contemporary English (2003)	Random House Webster's College Dictionary (2000)	The New Oxford Dictionary of English (1998)
<b>Refugee</b>	A person who has been forced to <u>leave their country</u> in order to <u>escape war, persecution or natural disaster</u>	A person who flees for refuge or safety, <i>esp. to a foreign country</i> , as <u>in time of political upheaval</u> .	Someone who has been forced to <u>leave their country</u> , especially during a war, or for <u>political or religious reasons</u> .

Table 6-1: A selection of three pre-Katrina dictionary definitions of *refugee*

Owing to its domain of legal discourse, definitions and understandings of what, exactly, a 'refugee' is, can also be drawn from political discourse. The United Nations Convention Relating to the Status of Refugees defined conditions under which refugee status would be applied in a document adopted in 1951, entered into force in 1954, and amended in 1967 (Office of the United Nations High Commissioner for Human Rights, 1967). In this document, a *refugee*:

...owing to 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 as a result of such events, is unable or, owing to such fear, is unwilling to return to it.

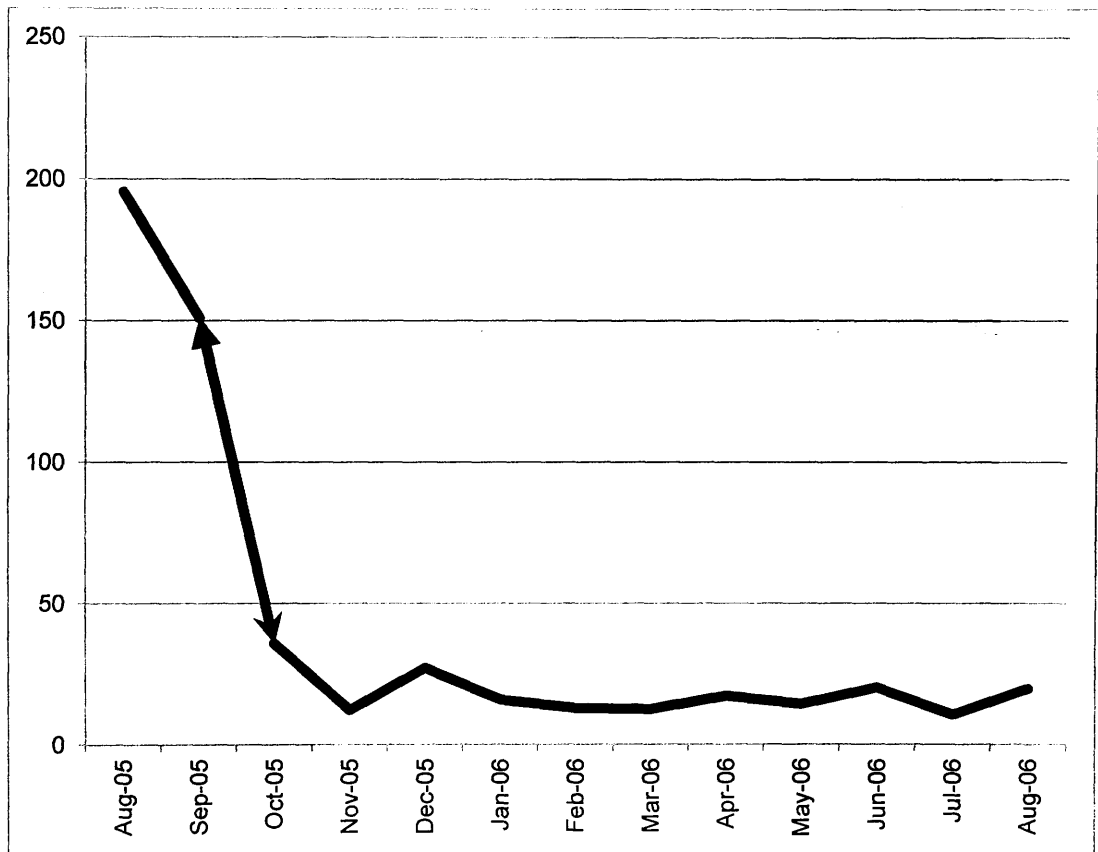
The United Nations definition clearly stipulates both cause (fear of persecution) and location (outside of one's own country) of refugee displacement. These two variables, and their appearance in arguments both for and against application of the word *refugee* to people affected by Katrina (PAKs), played a large part in the discourse of Hurricane Katrina, and will be explored shortly.

## 6.2 Method

Though this chapter will be of a more qualitative nature than the two preceding it, this thesis is fundamentally corpus-based, and as such, details on categorisation schemes of concordance lines (and frequencies) will be outlined before proceeding to finer analysis.

A simple search for both the plural and singular form of *refugee* returns 2,017 matches in the corpus, in 1,047 texts from 20 publications. The plural is far more frequent than the singular form, comprising 1,546 (76.65%) of the total frequency, and only

publications represented by six or fewer texts overall show no usage whatsoever. The unusual diachronic distribution pattern of this naming strategy was one factor recommending it for further investigation. Over 73% of all instances of *refugee(s)* occur in the initial 37 days of data, with a markedly steep drop in usage from October 2005 onwards (see the period marked by arrows in Figure 6-1 below).



**Figure 6-1: Overall frequency of *refugee(s)* in the Katrina corpus by month, expressed in frequency per million words.**

Each instance of the word returned by this simple search, along with its immediate context, was manually categorised into a broad rhetorical category, as described below. Salient features of concordance lines belonging to each category will be explored successively throughout the remainder of this chapter.

### **6.2.1 Broad categorization of concordance lines**

Categorisation was completed in three successive rounds. For the initial round of coding, I categorised each instance of *refugee(s)* as either referring to a) people affected by Hurricane Katrina or b) UNRELATED others, e.g. Bosnian refugees. In the first case, I had initially intended to further categorise each hit into one of two broad subcategories:

UNCONTESTED and CONTESTED (descriptions of criteria will soon follow). However, as the first round progressed, it became clear that several other subcategories were needed. These—DEFINITIONS, PARALLELISMS, CAMPS, and ANIMALS—were added *ad hoc*, and points at which each of these were implemented were noted separately. The second round consisted of reviewing the batch of first-round categorisations that had preceded completion of the final categorisation scheme, and modifying codes as necessary. The third and final round was a blind re-categorisation of the entire set to ensure an enhanced level of categorisation consistency. There were 16 cases of inconsistency occurring between PARALLELISMS and UNRELATED, which were resolved through extended analysis of the concordance lines. The number of hits in each category is as follows:

Category	Subcategory	Frequency	Number of texts	% Overall	% Katrina
Katrina	UNCONTESTED	1,258	748	62.37%	73.22%
	CONTESTED	238	101	11.80%	13.85%
	DEFINITIONS	42	23	2.08%	2.44%
	PARALLELISMS	100	65	4.96%	5.82%
	CAMPS	64	53	3.17%	3.73%
	ANIMALS	16	14	0.79%	0.93%
Non-Katrina	UNRELATED	239	143	11.85%	
	AGENCIES	60	36	2.97%	

Table 6-2: Frequency of *refugee* in various categories of semantic sense

### 6.2.2 Description of categories

- Category: Katrina
  - Subcategory UNCONTESTED was assigned to concordance lines where *refugee(s)* was used to describe American residents displaced by Hurricane Katrina. These concordance lines do not feature scare quotes, negation, or allusion to dictionary definitions, showing implicit approval of the term in its application to the social group. This subcategory is the most frequent by a vast margin, representing over 67% of all hits for *refugee(s)*, and over 78% of the hits in the Katrina category.
    - *Example: “Refugees from along the Gulf Coast, fleeing deadly Hurricane Katrina, said their search for affordable hotels that had vacancies led them all the way to Gwinnett.” (Atlanta Journal-Constitution, 31 August 2005)*



- Subcategory CONTESTED encompasses concordance lines where *refugee(s)* was used to describe American residents displaced by Hurricane Katrina, but differs from the previous subcategory due to the presence of scare quotes (see example), negation (see example), and argumentation strategies. Concordance lines in this subcategory show explicit disapproval of the term, contesting use with a variety of strategies that will be explored in this chapter. The CONTESTED subcategory is the third most frequent overall, and the second most frequent of the Katrina category.
  - *Example:* “I say something about ‘**refugees.**’ Andre is quick to correct me. ‘No, don’t say that, we’re not refugees at all. That means something else. We’re ‘displaced citizens.’” (*Atlanta Journal-Constitution*, 11 September 2005)
  
- Subcategory PARALLELISMS lays somewhere between the Katrina and non-Katrina categories. In these concordance lines, *refugee(s)* may not be used in direct reference to those displaced by Hurricane Katrina, but appears as an allusion, metaphor, or simile likening other historical or international refugees to the Katrina displaced, comparing or contrasting attributes, behaviours, and experiences through intertextual referencing. Frequencies of comparison to various social groups, and analysis of whether the proposed relationship construes the Katrina social actors in relatively positive or negative light will be analysed later in this chapter.
  - *Example:* “For the last hundreds in the arena, the day was particularly trying. Families sat on bundles of possessions like war refugees.” (*New York Times*, 21 September 2005)
  
- Subcategory DEFINITIONS often collocates with *define*, appearing in the phraseology, “refugee is defined as” or some similar iteration. The frequency of this subcategory is a symptom of discussion on contested terminology. Whether definitions are invoked more frequently to support or refute the application of *refugee(s)* to American residents displaced within America after Hurricane Katrina will be explored shortly.

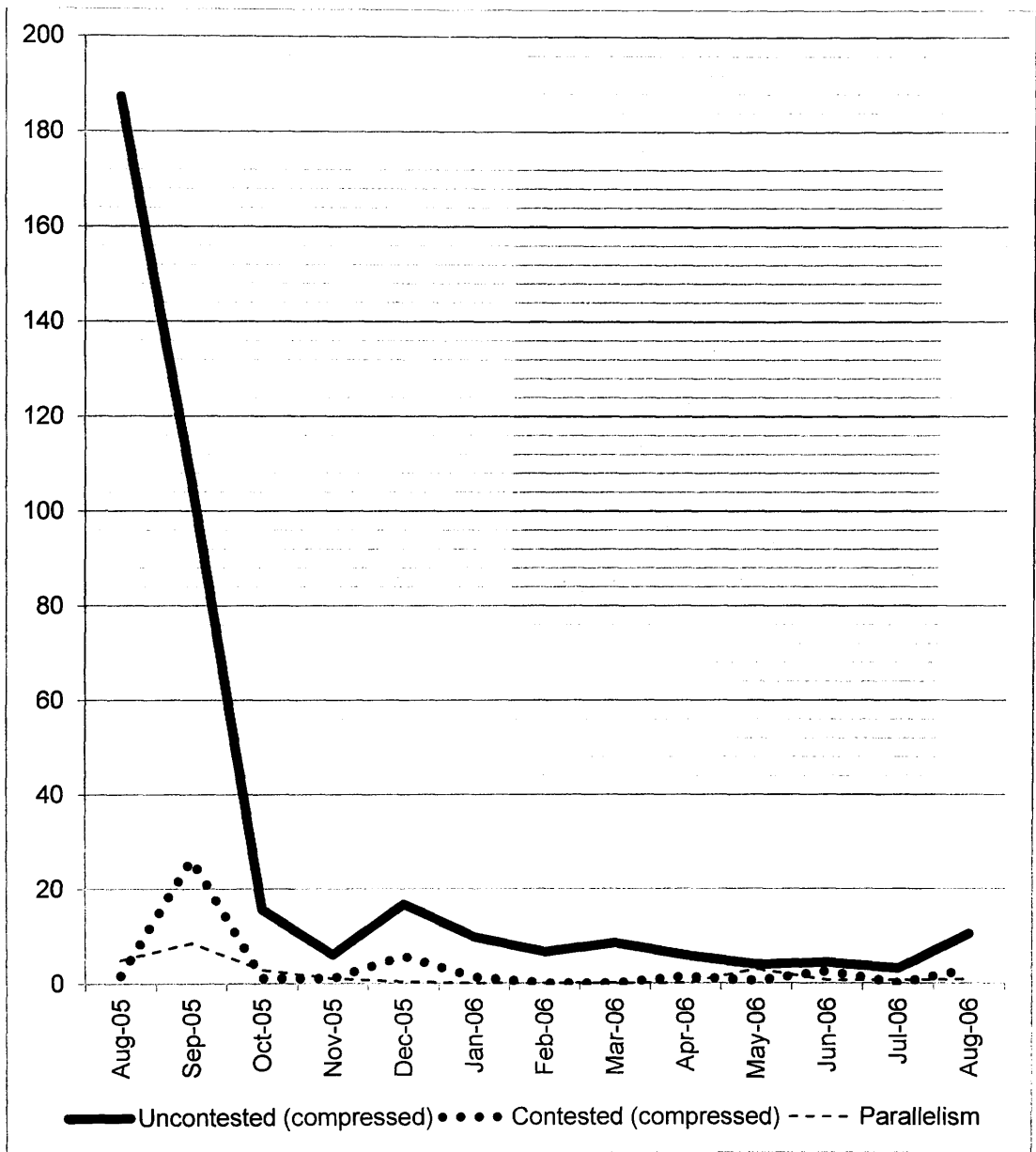
- *Example: “Under international law, **refugees are defined as** people who cross national borders to flee persecution.” (*New York Times*, 11 September 2005)*
  - Subcategory CAMPS refers to areas occupied by groups of PAKs as ‘refugee camps’ or semantically similar. While these do not explicitly refer to the PAKs as refugees, their conditions and collective counts as likened to those of refugees. Presence of this phraseology in discourses surrounding PAKs surely impacted overall construal, and indicates a certain level of UNCONTESTED use.
    - *Example: “In Snellville, a two-person household has been transformed into a **refugee staging area**.” (*Atlanta Journal-Constitution*, 2 September 2005)*
  - Subcategory ANIMALS is a very infrequent subcategory with only 12 concordance lines. These are rare cases where reporters use *refugee(s)* in reference to pets or other animals affected by Hurricane Katrina. Due to its indirect relationship to human participants, and low frequency (with resultant incompatibility with any level of generalisation), this subcategory will be discussed only very briefly outside of the methodological overview.
    - *Example: “Since Hurricane Katrina devastated the South, four-legged **refugees** have been arriving in Butler County by the truckload.” (*Pittsburgh Post-Gazette*, 6 November 2005)*
- Category: Non-Katrina
  - Subcategory UNRELATED includes uses of *refugee(s)* not relating to American residents displaced or otherwise affected by Hurricane Katrina. Lexical indicators were most often found directly to the right (“Bosnian refugees”) or left (“refugee of Ethiopia”) of the node. In roughly a fifth of cases, extended concordance lines were viewed to determine the nationality of the refugee(s).

- *Example:* “Orchard House, which was the setting for Alcotts’ “Little Women,” housed one of the first adult education centers in America, sheltered **refugee** Native Americans during King Philip’s War and was home to a Concord Minuteman.” (*Boston Globe*, 17 November 2005)
  - Subcategory AGENCIES was often indicated by capitalisation of the node and the words closest to it, indicating proper names of associations. As this is a study of the representation of humans, discussion of groups such as the *Massachusetts Immigrants and Refugee Advocacy Coalition* have been deselected. However, it is interesting to note the frequency with which government agencies traditionally working in [foreign] refugee relief were associated with Katrina’s displaced.
    - *Example:* “Calling on **refugee** agencies to help Americans displaced by a hurricane in their own country hasn’t been an easy idea to sell in Washington.” (*Wall Street Journal*, 6 October 2005)

### **6.2.3 Frequency of categories**

One advantage of manual categorisation of the search results is discovery of infrequent categories. Here, the PARALELLISMS, DEFINITIONS, and CAMPS, ANIMALS categories combined represent just over 10% of the overall frequency of *refugee(s)*. In studies of nodes occurring in higher frequencies, many corpus linguists might have relied upon collocation profiles (e.g. with *defined* or with negation) to categorise concordance lines. Spot-checking results would have been unlikely to bring these subcategories to the fore, particularly in very rare cases such as the ANIMALS concordance lines.

For the purpose of simplicity, categories have been conflated and depicted in frequency over time. Here, Contested (compressed) includes CONTESTED and DEFINITIONS; Uncontested (compressed) includes UNCONTESTED and CAMPS; Parallelism consists of this subcategory only.



**Figure 6-2: Frequency of compressed semantic categories of sense of *refugee(s)* over time**

The extremely high frequencies of the UNCONTESTED category in the earliest portion of the date range make the later, less frequent trends more difficult to analyse. A clipped version of the lower portion of the graph appears below. In observing the patterns occurring at this level, particularly between September and October 2005, the effect of the CONTESTED discourses upon the UNCONTESTED discourses becomes very clear; a spike in the former in September leads to a trough in the latter in October.

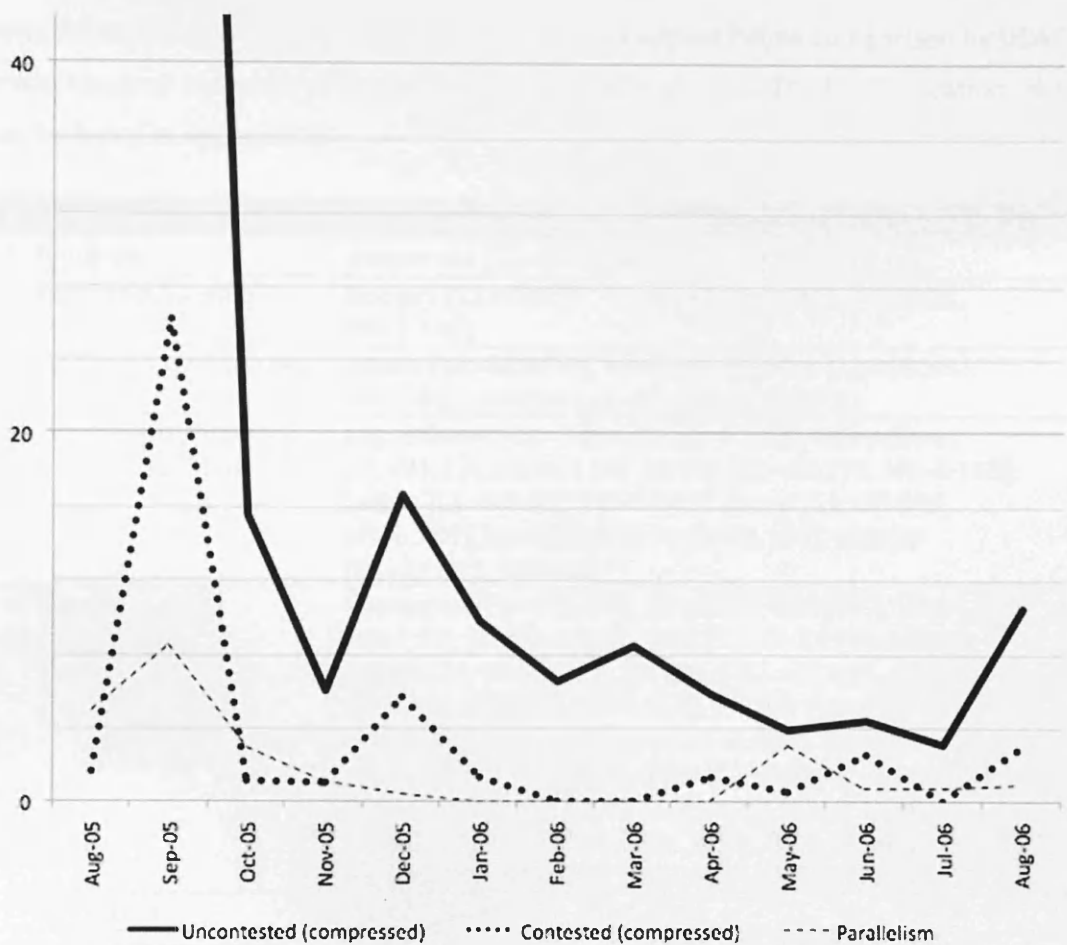


Figure 6-3: Frequency of compressed semantic categories of sense of *refugee(s)* over time (clipped)

This pattern suggests that diachronic semantic change might have been a major factor in the chronology of word use. In the next sections, I will be exploring through qualitative analysis whether the less frequent CONTESTED discourse managed to undermine and overrule the UNCONTESTED discourse.

### 6.3 Analysis of categories

The following sections contain analyses of the various patterns of use found in each of the semantic sense categories of *refugee(s)*.

#### 6.3.1 Category: KATRINA, Subcategory: UNCONTESTED

To establish an overview of the construal of this social group generally, collocation data was generated for the 1,258 instances of *refugee(s)* categorised in the UNCONTESTED category. Collocates are part-of-speech tag lemmas, within a +/-3 span,  $MI \geq 3$  and  $LL \geq 10.83$ , with minimum frequency decreased to 5 to account for the drastically smaller

node occurrence. As in previous chapters, collocates appear below categorised by USAS broad category and ordered by decreasing log likelihood value. The full collocation table can be found in Appendix G.

USAS broad category	Significant collocate lemma
E: Emotion	<b>desperate</b> [LL=37.753, MI=4.797]
F: Food and farming	<b>hungry</b> [LL=59.281, MI=6.166], <b>meal</b> [LL=23.651, MI=3.778]
H: Architecture, housing, and the home	<b>house</b> [LL=222.738, MI=5.13], <b>shelter</b> [LL=50.582, MI=5.46], <b>rooftop</b> [LL=30.543, MI=5.071]
M: Movement, location, travel and transport	<b>superdome</b> [LL=139.921, MI=5.133], <b>astrodome</b> [LL=91.239, MI=6.114], <b>arrive</b> [LL=80.272, MI=4.115], <b>influx</b> [LL=67.443, MI=6.281], <b>camp</b> [LL=36.855, MI=6.739], <b>bus</b> [LL=28.184, MI=3.125], <b>scatter</b> [LL=25.611, MI=4.461]
N: Numbers and measurement	<b>thousand</b> [LL=415.309, MI=5.39], <b>25000</b> [LL=164.466, MI=7.65], <b>hundred</b> [LL=65.631, MI=3.978], <b>10000</b> [LL=63.22, MI=5.551], <b>busload</b> [LL=57.901, MI=8.357], <b>9000</b> [LL=57.557, MI=7.344], <b>23000</b> [LL=47.378, MI=8.235], <b>fill</b> [LL=42.475, MI=3.504], <b>20000</b> [LL=40.799, MI=5.615], <b>1000</b> [LL=38.096, MI=4.429], <b>pack</b> [LL=24.043, MI=4.263], <b>3000</b> [LL=23.245, MI=4.739], <b>300</b> [LL=22.264, MI=4.032], <b>dozen</b> [LL=21.599, MI=3.241], <b>200</b> [LL=18, MI=3.483]
P: Education	<b>enrol</b> [LL=18.799, MI=4.066]
S: Social actions, states and processes	<b>huddle</b> [LL=53.544, MI=6.93], <b>stranded</b> [LL=40.491, MI=5.582], <b>crowd</b> [LL=31.698, MI=3.886], <b>gather</b> [LL=29.683, MI=4.032], <b>rescue</b> [LL=29.32, MI=3.684], <b>crisis</b> [LL=26.725, MI=3.455], <b>strand</b> [LL=25.915, MI=5.138], <b>invite</b> [LL=12.252, MI=3.04]
W: World and environment	<b>houston</b> [LL=75.193, MI=3.801]
Z: Grammatical bin	<b>katrina</b> [LL=396.06, MI=3.106]

Table 6-3: Collocates of *refugee(s)* in the UNCONTESTED category

Viewing collocates in semantic categories, it becomes evident that UNCONTESTED discourses around *refugee(s)* are primarily focussed on the concepts of: 1) Numbers and measurement [N], 2) Social actions, states and processes [S], and 3) Movement, location, travel and transport [M]. None of these are unexpected, nor radically different from collocates of the other naming strategies observed in Chapter 5. The attributes appear to have been simplified further than those of the other naming strategies, however. Mental and physical health does not feature as statistically significantly connected to *refugee(s)*, and the Money and Commerce category is completely absent from this group. In the UNCONTESTED category, *refugee(s)* are distilled to a basic form: a huge and transient

group in need of social support. However, this is not to say that *refugee(s)* has no unique collocates of its own.

The analysis of this naming strategy has uncovered the first instance of naming strategy collocation with *crisis* (underlined in the table above). The bigram *refugee crisis* occurs quite infrequently—only five times in the UNCONTESTED set, reflected in line 1-5 below—but does highlight the acceptance of this nomenclature by the majority of articles immediately following Hurricane Katrina. Here, displacement of PAKs is a *crisis* to be reckoned with.

1. Fox News Channel's Shephard Smith retained objectivity as he described the lethal mix of water, gas and human waste flooding New Orleans in 'a **refugee crisis** the United States has never known.' (*Denver Post*, 5 September 2005)
2. But Hurricane Katrina delivered America its own **refugee crisis**, arguably the worst since Sherman's army burned its way across the South. (*New York Times*, 11 September 2005)
3. At least 18 U.S. states are dealing with the Katrina **refugee crisis**, either by providing emergency shelter or by offering longer-term housing and education. (*San Francisco Chronicle*, 4 September 2005)
4. With a shattered New Orleans all but emptied out, an unprecedented **refugee crisis** unfolded across the country Sunday, as governors and emergency officials struggled to feed, shelter and educate more than a half-million people dispossessed by Hurricane Katrina. (*St. Louis Post-Dispatch*, 5 September 2005)
5. THE DOW JONES Industrial Average gained 50 points last week as Katrina crushed the Gulf Coast, drowned New Orleans, killed thousands, created an American **refugee crisis**, made a pro-business president look incompetent and sent gasoline prices into orbit. (*St. Louis Post-Dispatch*, 8 September 2005)

Figure 6-4: Sample concordance lines of *refugee crisis*

### 6.3.1.1 Verbal collocates

Viewing collocates by part-of-speech tag allows for further investigation into processes enacted by and upon *refugee(s)*. A concise transitivity analysis (Halliday, 1985) of each of the concordance lines associated with verbal collocates reveals a roughly equal number of *refugee(s)* occurring in the agent/actor position and those in the patient/beneficiary position (a small number of collocations were not true collocates of *refugee(s)* in context and have been excluded from transitivity analysis). However, in viewing verbal collocates within USAS semantic categories, it becomes clear that *refugee(s)* are 'doers' in a limited number of domains, namely: M (Movement, location, travel and transport) and N (Numbers and measurement). Again, the most salient features of this social group are the manner of their dispersion and their growing numbers. One collocate from category M—*evacuate*—is contrary to this trend, with *refugees* construed as 'being

evacuated', in contrast to the agency implied in the alternate naming strategy, *evacuee*. Not unexpected disproportions in category S also indicate a deeper inequality of relations between PAKs and non-PAKs; while *refugees gather* and *huddle*, they are being invited, rescued, or further disempowered with the "stigma of charity" (Fothergill, 2003, p. 7).

Tagged lemma	Collocate frequency	In no. of texts	MI	LL	USAS sem-tag	Refugee as agent/actor	Refugee as patient/beneficiary
<i>shelter</i>	9	9	5.46	50.58	H1	0	9
<i>house</i>	43	41	5.13	222.74	H4	0	42
<i>arrive</i>	21	18	4.12	80.27	M1	18	0
<i>scatter</i>	6	6	4.46	25.61	M1	4	1
<i>camp</i>	5	5	6.74	36.86	M7	5	0
<i>pack</i>	6	6	4.26	24.04	N3.4	5	0
<i>fill</i>	14	12	3.50	42.48	N5	14	0
<i>enrol</i>	5	5	4.07	18.80	P1	1	4
<i>gather</i>	8	8	4.03	29.68	S1	8	0
<i>invite</i>	5	5	3.04	12.25	S1	0	5
<i>huddle</i>	7	7	6.93	53.54	S5	7	0
<i>strand</i>	5	5	5.14	25.92	S5	0	5
<i>rescue</i>	9	6	3.68	29.32	S8	0	2
<b>Total</b>						62	68

Table 6-4: Frequency of verbal collocates of UNCONTESTED *refugee(s)* by USAS semantic category and transitivity

### 6.3.1.2 Instances of metaphor in verbal collocates

The appearance of verb lemmas *fill*, and *pack* in the table of collocates above indicates the presence of constraint and space metaphors, similar to those explored in the previous chapter. In reviewing the concordance lines, a variety of other familiar metaphors began to appear. While searching for metaphorical constructions in a corpus is always somewhat problematic, searching for water metaphors in a corpus of newspaper articles describing a natural disaster involving catastrophic amounts of water is particularly difficult to do without a large level of qualitative analysis. For instance, a lemma search for *pour* returns 1,412 matches in the corpus, but humans are the actors *pouring* into or out of physical locations in only 146 of these.



Therefore, it was necessary to use a combined approach. Firstly, lexical indicators of metaphors previously discovered in Chapter 5 (*stream* and *flood*) became lemma node words searched within the UNCONTESTED subcorpus, in addition to related words identified in concordance lines: *absorb*, *swell*, *pour*. Each concordance line resulting from these was analysed for the presence of *refugees*. Nine times, a collective group of *refugees* is metaphorically described as a *stream*, and seven times, as a *flood*. Five times *refugees* are *absorbed* into cities other than their own. *Swell* can allude to both water metaphors and to the *topos* of numbers; in eight concordance lines, school enrolments, state populations, and the Astrodome are depicted as being *swollen* with *refugees*. After revisiting indicators from Chapter 5, each concordance line was read for the presence of any new or unique metaphors. This rereading uncovered a final water metaphor in the form of *saturation*, another combination of water metaphor and the *topos* of numbers. Four times, *refugees* are described as having *saturated* the apartment market, hotels and motels, and the city of Shreveport. Sample concordance lines of water metaphors appear below. As illustrated in sample concordance lines 7 and 8 below, water metaphors are also frequently accompanied by additional markers of risk, including size (*massive*) and threat (*overwhelmed*).

6. The Superdome swelled with thousands of storm **refugees** from all parts of the city. (*New York Daily News*, 31 August 2005)
7. In Baton Rouge, refugees lined up for cash and gasoline and quickly saturated the apartment market... All were part of a massive, chaotic and unprecedented stream of American **refugees** fleeing the ravaged Gulf Coast area. (*Wall Street Journal*, 1 September 2005)
8. Storm **refugees** overwhelmed the state of Louisiana and poured into cities from coast to coast, crowding sports arenas, convention centers, schools, churches and the homes of friends, relatives and even strangers. (*Pittsburgh Post-Gazette*, 4 September 2005)
9. The **best** thing Entergy could do for New Orleans, Mr. Hemlick said, was to restore service as quickly as possible to cities like Jackson and Baton Rouge that have been absorbing **refugees**. (*New York Times*, 6 September 2005)

**Figure 6-5: Sample concordance lines illustrating water metaphors associated with *refugee(s)***

Only two attributive adjectives appear in the collocate table for UNCONTESTED instances of *refugee(s)*: *hungry* and *desperate*. Positive attributes such as those observed in collocation tables of *citizen*, and *survivor* are notably absent here. Markers of income and ethnicity are likewise non-existent; some might infer that (low) income and (minority) ethnicity are incorporated in the word *refugee* when used in such a context, but with such a small sample of attributive adjectives to base analysis on, this is not open for further exploration.

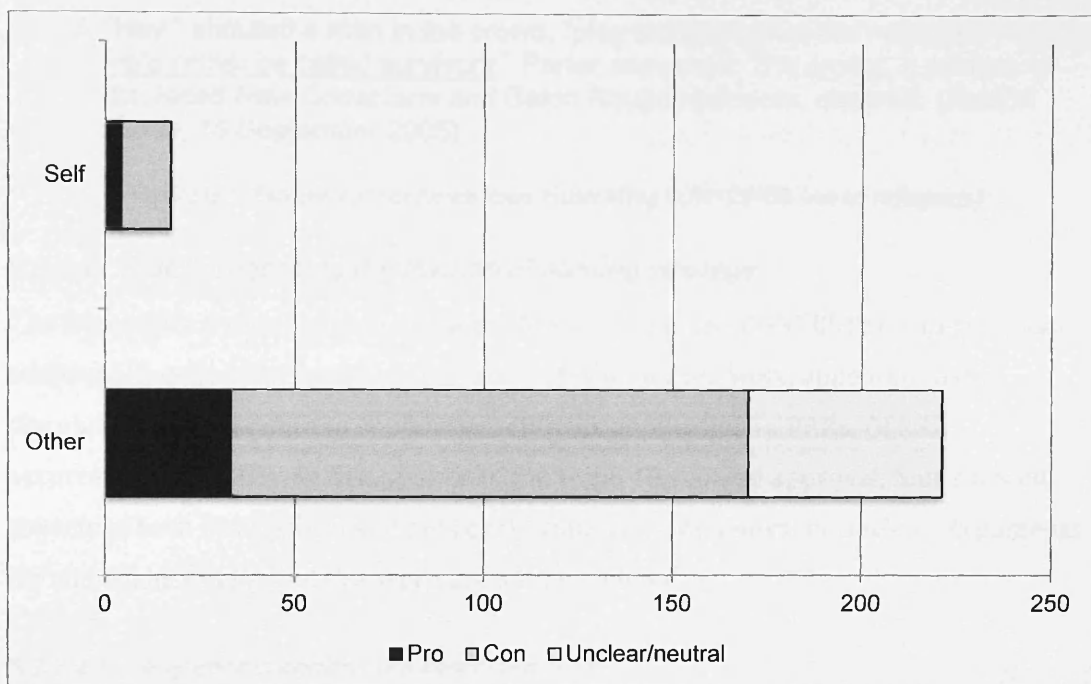
Overall, the application of the word *refugee*—which is frequently defined to indicate flight from one’s home country to escape persecution—to American residents/citizens in America results in the group being effectively “stripped of the specificity of culture, place and history” (Malkki, 1995, p. 12). Transitivity analysis reveals that *refugees* lack agency and are additionally constructed as relatively powerless, effecting action almost exclusively in regards to (their own) movement. Water and volume metaphors further ‘other’ this social group, defining and collectivizing them in terms of threat to city/state membership boundaries and established social order. These arguments, and others, formed the basis of a public debate on the semantics of what it means to be the ‘other’ in America, further explored below.

### **6.3.2 Category: KATRINA, Subcategory: CONTESTED**

The CONTESTED category—wherein the suitability of the term “refugee” in reference to PAKs is disputed—accounts for 236 instances (out of the overall 2,017 frequency of *refugee(s)*) in 100 texts from 17 (of a possible 24) publications. This is a minority category, with an approximate frequency of 6.42/million words. Analysis of this category will cover: 1) the direction of the naming strategy (as a marker of self- or other-identification); 2) frequency of this sense of *refugee(s)* in various publications; 3) common arguments in favour of continuing or ceasing to use this naming strategy for PAKs; and 4) any alternative naming strategies offered as more agreeable and/or appropriate.

#### **6.3.2.1 Direction of naming strategy**

Of 236 total hits for *refugee(s)* in the CONTESTED subcategory, the label is being applied to someone else (outwardly) 219 times, and to refer to oneself (inwardly) in only 17 cases. Only four of these show positive approval or alignment with this identity (two of which come from the same quotation in the same article)—labelled PRO; the other 13 explicitly reject group membership—labelled CON. By and large, Katrina evacuees did not identify themselves as *refugee(s)*, but were rather identified by others as belonging to this social group. Figure 6-6 below illustrates the proportion of use by self and others.



**Figure 6-6: Overall frequency of CONTESTED *refugee(s)* application by direction (self-reference versus other-reference) and appraisal (appropriateness versus inappropriateness of application)**

For those who gave reasons for rejecting the nomenclature, one explained that he was not a refugee because he had not fled from political capture, seven stated that they were not refugees because they are Americans, and one more said that in addition to being American, they are law-abiding taxpayers. Two concordance lines clarify that the speaker self-identifies as an *evacuee* (see line 12 below) or a *survivor* (line 13), both more positive naming strategies, and in the case of the latter—with its association with overcoming adversity—certainly more empowering.

10. Comedian and Louisiana native Paul Mooney was among several stars who perceived racism in how the media had addressed the natural disaster: “The news referred to the white people as survivors and the black people as refugees. I don’t like the idea of being a refugee.” (*USA Today*, 12 September 2005)
11. The African-American leaders in Chicago said they were also disturbed by some commentators’ description of the displaced men and women camping in stadiums and convention centers. “We are not refugees. This is our home. We built this country. America profited on the backs of black people,” fumed Ald. Freddrenna Lyle (6<sup>th</sup>). (*Chicago Sun-Times*, 4 September 2005)
12. Although Preciado and her children are clients of the same relief agency that is aiding her Cuban cousins, she bristles at the word ‘refugee.’ “We’re evacuees,” she said adamantly. “I’ve been here 37 years. I was a refugee when I came from Cuba.” (*Atlanta Journal-Constitution*, 5 October 2005)

13. “Hey,” shouted a man in the crowd, “play something for the **refugees**.” “I think we’d rather be called survivors,” Porter answered. The crowd, a mixture of displaced New Orleanians and Baton Rouge residents, cheered. (*Boston Globe*, 15 September 2005)

Figure 6-7: Sample concordance lines illustrating **CONTESTED** use of *refugee(s)*

### 6.3.2.2 *Stance regarding application of naming strategy*

The immediate context of each instance of *refugee(s)* in the **CONTESTED** category was additionally coded for stance, or indication of whether the word appeared in the discourse of contestation to support or refute its application to PAKs. Of 237 occurrences, 143 showed disapproval of the term, 40 showed approval, four showed aspects of both disapproval and approval, while 47 were neutral or unclear. Arguments for and against the use of this word are detailed below.

#### 6.3.2.2.1 *Arguments against the contested term*

Appearing in 43 of the total 100 articles, the most frequent justification for cessation of *refugee* use in Katrina reporting was an insistence that ‘refugee’ indicates a crossing of international borders and a dislocation to a country other than one’s own. A tandem identity of ‘law-abiding tax payer’ (see line 16 below) often accompanies this primary depiction of PAKs as American citizens first (see line 14 below), and hurricane survivors second. The state of being a homeowner—or of being only *temporarily* homeless—appears in seven articles (see line 15 below), with an additional two specifying that PAKs have jobs, and therefore are not *refugee(s)*. This identity, the employed, tax-paying American citizen, is the most frequently occurring.

14. It is the idea of storm victims as outsiders, as others, as foreigners that is driving some people crazy. If those “**refugees**” are not quite Americans, not quite our own, it becomes acceptable for the nation to do less than its best by them. (*Boston Globe*, 5 September 2005)
15. Although the people of New Orleans and Biloxi, Mississippi, and Mobile, Alabama, and other places may have lost their homes, they are not refugees,” Hamilton declared passionately from the high pulpit of the ornate old church. (*Boston Globe*, 9 September 2005)
16. All those folk who are going to be over there aren’t hoodlums. Most of them had jobs. They’re taxpayers, not refugees. (*Atlanta Journal-Constitution*, 15 September 2005)

Figure 6-8: Sample concordance lines illustrating arguments **against** the use of *refugee(s)*

Another frequent argument against application reiterates Rev. Al Sharpton’s declaration that “It is racist to call American citizens refugees” (‘GIMME SHELTER’, *Boston Globe*, 11 September 2005). What I found most interesting in this argument was that while some

journalists and public figures agreed that the term was racist, they did not agree on exactly *why* this was the case. Considering the texts contained CONTESTED category concordance lines as a subcorpus, the word *racist* appears 18 times in 13 of the 100 total texts. Of these, the word is used in discussions of applicability of the word *refugee* 17 times, with a single additional concordance line stating, “Some left-wing bloggers asserted that concern with looting was inherently racist” (‘A HURRICANE OF BLAME’, *Boston Globe*, 12 September 2005). Of the 17 concordance lines relating racism to *refugee(s)*, there are nine attributions (eight direct quotation and one indirect quotation) to Rev. Jackson.

However, a certain grounds for rejection of the term emerges as even more frequent in the corpus. A hierarchy of struggle emerges, in which *refugee* is seen as an affront not to individuals of a certain ethnicity, but to patriotism in general. The word ‘*refugees*’ is ‘tainted’ (see concordance line 17 below) with a prosody of a foreign other. Despite their current misfortune, PAKs are still set apart in a hierarchy relegating non-Americans firmly to the bottom. *Refugees* are associated with “problems ‘over there’” (line 18), including persecution or war (line 19). Concordance line 20 hints at racial undertones, likening the physical appearance of (African-American) PAKs to Sudanese *refugees*. For recent generations of Americans who have never known large-scale enemy invasion, employment of a naming strategy that denotes statelessness runs counter to the American culture of patriotism, by which being *any* kind of American is superior to being *any* kind of foreigner. “In other words, their U.S. citizenship—and what it entailed in terms of power, civil rights, and access to resources—precluded their ability to become refugees, no matter how dire their circumstances” (Masquelier, 2006, p. 736).

17. But others heard **refugees** as an un-American, blame-the-victim word, and their interpretation has stuck; many evacuees now say the word sounds like an attempt to keep them at arm’s length, out of the mainstream... And so many news outlets, whether or not they agree that **refugee is tainted**, have banned or restricted the use of the word. (*Boston Globe*, 11 September 2005)
18. I am also stunned to hear people in our own land referred to as “**refugees**.” That is a word associated with problems ‘over there,’ isn’t it? (*Atlanta Journal-Constitution*, 5 September 2005)
19. “A **refugee** is a prisoner of war or something,” said Paulette Jolla, a New Orleans [sic] who left before the storm hit. (*Washington Post*, 6 September 2005)
20. The real reason “**refugees**” doesn’t sit well with many people, particularly African-Americans, though, seem to be that “**refugees**” seems to equal “victims.” They are the people they see on television, suffering in places like the camps of Sudan. (*Christian Science Monitor*, 21 September 2005)

Figure 6-9: Sample concordance lines illustrating ‘un-American’ arguments against *refugee(s)*

Rejection of one naming strategy does not occur solely without proposed substitution with another. In 37 instances, journalists extend their arguments against using *refugee(s)* by offering alternative naming strategies. Some of the most frequent (and empowering) naming strategies from the previous chapter top this list: *evacuees*, *survivors*, and *displaced citizens* are the most preferred, followed by the previously discovered (and relatively disempowered) alternative *victims*. It is interesting to note that *displaced citizens* and *American citizens* overtly mark nationality; the appearance of *taxpayers* contributes to another discourse of financial contribution.

Proposed alternative	Frequency
evacuees	7
survivors	7
displaced citizens	5
american citizens	4
homeless	4
victims	4
guests	3
displaced	2
taxpayers	1

**Table 6-5: Frequency of alternative naming strategies for *refugee(s)* proposed in the CONTESTED set**

A sample of the concordance lines leading into alternative naming strategies is shown below. Lines 21 and 22 illustrate typical structure of the CONTESTED category; membership in the category can be established through presence of negation (line 21) or quotation (line 22). The second sentence in line 21 is an imperative call to action for readers. Following the colon in line 22, the journalist attributes justification to an outside (powerful) party without specifically assigning quotations—this is the reasoning of ‘some’ members of Congress. While line 21 shows an explicit renaming, line 22 offers an alternative in usage. Line 23 is an interesting view into the journalistic process of selection. Though *refugees*, *displaced citizens*, and *evacuees* are rejected, one of the most empowering naming strategies—*survivor*—is offered alongside one of the least empowering—*victim*.

21. These are not refugees. Let’s stop calling them **refugees**. They are American citizens. (*The Boston Globe*, 5 September 2005)
22. Another reason cited for not using the word arises from what some members of Congress point out: The word “**refugees**” suggests that the displaced victims, many of whom are poor or people of color, are being treated like second-class citizens or non-Americans in their own country. (*Seattle Post-Intelligencer*, 8 September 2005)

23. 'The United Nations has a term for people uprooted by natural disasters or unprotected by national authorities: "internally displaced people." "I.D.P.'s remain in their own country," Cohen said. I don't go for the bureaucratic initialism, but also resist applying **refugee** to people who live in the U.S. homeless, though currently accurate, implies permanent rootlessness. Displaced citizens does not cover the many victims who are not citizens, and evacuees is a highbrow concoction. My choices: Katrina survivors overall and, specifically for the inundated of New Orleans, flood victims. (*The New York Times*, 18 September 2005)

Figure 6-10: Sample concordance lines illustrating alternative naming strategies provided in place of *refugee(s)*

### 6.3.2.2 Arguments for the contested term

Precise justifications for arguments to continue the use of *refugee(s)* as an acceptable naming strategy for PAKs are much more difficult to uncover and quantify than those arguing against the same. Only 40 concordance lines seemed to support the continued use of this naming strategy, and arguments found in these can be categorised and enumerated as follows: to be called a 'refugee in America' is not pejorative, as America is a nation of proud refugees (5 cases), the term is not racist or insulting in the opinion of the journalist (6), that the term is the most descriptive (3) or in popular use (2). Other publications simply stood by their use of the word, stating that it would be applied where appropriate (7), and that so-called 'refugees' in their experience did not object to the term (2). One journalist rejects the debate outright, stating that "the debate over the word 'refugee' is nonsense but I'm not going to enter into it; I was watching a PBS documentary on the First World War the other night, and the Frenchies who fled Paris to Normandy to escape the Germans were 'refugees' without any apparent protest from Jesse Jackson" ("Aldermen forget where money comes from: us", *Chicago Sun-Times*, 14 September 2005).

### 6.3.3 Category: KATRINA, Subcategory: DEFINITIONS

This section is a small departure from sections 6.3.1 and 6.3.2, which dealt with the use of *refugee(s)* in the context of direct reference to humans (whether contested or not). This discussion is focussed instead on a semantic sense that is more metalinguistic in nature, that of incorporating and discussing the definition of *refugee* in determining its applicability or inapplicability to PAKs.

#### 6.3.3.1 Method

Each instance of *refugee(s)* within the DEFINITIONS subcategory has been manually coded at two levels:

For type of definition, including:

1. *Dictionary definition*: Where the definition is attributed to a specific, named dictionary.
2. *Political definition*: Where the definition is attributed to (largely international) political institutions or conventions.
3. *Miscellaneous definition*: Where the definition is attributed to a person, or where the source is not given.

In cases where *refugee(s)* occurred in discussion about the heightened frequency of the word, most commonly in lists of 'buzz words' from 2005, these were categorised *meta-discussion* and excluded.

In the cases of 1A-1C, cases were additionally coded for construal of the word *refugee* as being either applicable or not applicable to PAKs (or showing aspects of both); three cases were unclear.

Applicability	Category Frequency	Reference Type	Reference Frequency
Inapplicable	19	Dictionary definition	4
		Political definition	5
		Miscellaneous definition	10
Applicable	9	Dictionary definition	5
		Political definition	0
		Miscellaneous definition	4
Aspects of applicability and inapplicability	4	Dictionary definition	2
		Political definition	2
		Miscellaneous definition	0
Unclear	3	Dictionary definition	2
		Political definition	0
		Miscellaneous definition	1
<i>Meta-discussion</i>	10		

**Table 6-6: Frequency of *refugee(s)* in the DICTIONARY category by reference type and statement on applicability of the term**

The frequency of various definitions can be seen in Table 6-6 above. Most frequently, dictionary definitions are employed to support an argument against use of the term *refugee* to refer to PAKs, and so this is where we start in section 6.3.3.2.

### **6.3.3.2 Arguments for inapplicability**

In the majority of cases (19 out of 37), definitions are invoked to argue against the application of the term *refugee(s)* to PAKs. Justifications for considering this term



inapplicable are twofold: Firstly, the reasons for displacement do not include escape from invasion or oppression (2 instances), religious, racial, ethnic, or political persecution (11 instances, e.g. lines 24-25 below), violence, civil wars, or human rights violations (3 instances); secondly, location of displacement is not across foreign borders (13 instances, e.g. lines 24-26 below) or specifically from a [foreign] Third World country (1 instance, line 27 below). Though 16 instances offer reasons for displacement as a distinguishing feature setting PAKs apart from *refugees* compared to only 14 instances citing location of displacement, the latter is more salient in another facet. In two concordance lines, alternatives to *refugees* are given—the *Washington Post* prefers ‘internally displaced persons’ while the *Chicago Sun-Times* favours ‘displaced Americans’—and both explicitly mark the location of PAKs, not their reasons for displacement. While it could be assumed that reasons are known (PAKs were all fleeing a natural disaster), it is interesting to note that from four articles in the DEFINITIONS subcategory offering alternative naming strategies, only one concordance line from the *Seattle Post-Intelligencer* (which offers aspects of both inapplicability and applicability) proffers a label that marks for reasons for displacement, stating that PAKs are in fact ‘**environmental refugees**’ (see line 28 below).

24. Critics of the word “**refugees**” in the hurricane’s context can find backing in the Oxford English Dictionary, considered by most wordsmiths to be authoritative. Its first definition of “**refugee**” says: “One who, owing to religious persecution or political troubles, seeks refuge in a foreign country.” None of the other definitions would fit those who fled from Hurricane Katrina. (*St. Louis Post-Dispatch*, 7 September 2005)
25. He had a point; although a **refugee** can be defined as “a person who seeks refuge,” it has carried the connotation since 1685 of “one who seeks refuge or asylum in a foreign country to escape religious or political persecution.” (*New York Times*, 18 September 2005)
26. Roberta Cohen, a senior fellow at the Brookings Institution, defined **refugees** as “people who cross borders to escape violence, civil wars and human rights violations -- because they do not have the protection of their own country. A natural disaster is not considered a cause of refugee status.” (*Pittsburgh Post-Gazette*, 18 September 2005)
27. While a **refugee** simply may be someone seeking refuge from disaster, it also is frequently associated with someone who flees a Third World country to find safety and comfort elsewhere. (*San Francisco Chronicle*, 12 September 2005)
28. IN MODERN POLITICAL terms, the hundreds of thousands of people left homeless by Hurricane Katrina are considered “environmental refugees” -- people forced out of their communities not because of a tyrannical government or violent civil war but because of a natural disaster. (*Wall Street Journal*, 5 September 2005)

**Figure 6-11: Sample concordance lines of DEFINITIONS being incorporated to argue against applicability of *refugee(s)* in referring to PAKs**

A variety of sources are cited for the definitions given in this category, with political definitions from authorities such as the United Nations and international laws more prominent here than in the other categories of applicability. Both of the alternative naming strategies in the INAPPLICABLE category are contained within the POLITICAL subcategory, indicating that those inclined to rule out *refugee(s)* on the basis of political definitions are more likely to proffer alternatives explicitly stating location of displacement or reasons for displacement (see line 29 below for an example) to distinguish PAKs from the cultural understanding of (or ideology around) 'real' *refugees*.

29. Under international law, **refugees are defined as people who cross national borders to flee persecution**. Ms. Cohen of the Brookings Institution said the evacuees from the Gulf Coast fit neatly into a newer category: "internally displaced persons." (*New York Times*, 11 September 2005)

**Figure 6-12: Sample concordance of an alternative naming strategy for *refugee(s)* offered following the political definition subcategory**

Sources under MISCELLANEOUS include two people: Roberta Cohen, a senior fellow at the Brookings Institution is quoted on two occasions as saying that "A natural disaster is not considered a cause of refugee status" (cited in line 26), while American Heritage executive editor Joe Pickett is quoted four times in two articles from different publications: "Traditionally, Pickett said, 'refugee' is used to describe somebody fleeing from political persecution or war" (*St. Louis Post-Dispatch*, 7 September 2005). The second mention of refugee in his statements even shows the possibility of change, as "Even so, he said his staff would revisit the word 'refugee' to see whether the definition needed a nuance about leaving one's native nation" (*St. Louis Post-Dispatch*, 7 September 2005) indicating the critical nature of the DEFINITIONS category, as a type of companion to the CONTESTED category.

The final MISCELLANEOUS definitions from this category include three unattributed statements iterating justifications given above: one states that *refugee* is "frequently associated with someone who flees a Third World country to find safety and comfort elsewhere" (line 27 above) while two unattributed quotations claim that *refugee* "has carried the connotation since 1685 of 'one who seeks refuge or asylum in a foreign country to escape religious or political persecution'" (line 25) – a definition I was unable to find associated with any published work, despite its appearance in quotation marks. Dictionary definitions from the Oxford English Dictionary (2 instances) and Webster's (2 instances) are also employed to argue applicability, but they appear more commonly in the APPLICABLE category, which we move to next.

### 6.3.3.3 Arguments for applicability

Overall, use of DEFINITIONS to argue that *refugee(s)* is an applicable naming strategy for PAKs was much less frequent than arguments against application of this term, though this mirrors the much rarer pattern of applicability over all. Employment of definitions to argue for the applicability of *refugee* to PAKs was second most frequent in the DEFINITIONS subcategory, appearing 9 times in 37 cases. Location of displacement features in justifications for applicability only twice. One article in The Boston Globe argues, “Some dictionaries say ‘refugees’ must cross borders, but that just means they’re abridged too far”. Another in the Atlanta Times Journal-Constitution cites Webster’s New World College Dictionary, Fourth Edition, as defining *refugee* as “a person who flees from home or country to seek refuge elsewhere, as in a time of war or of political or religious persecution”. This example is particularly interesting, as a (unspecified) version of Webster’s is cited in the INAPPLICABLE category, with “from home or country” replaced by “to another country”. Further, the APPLICABLE example cited above *does* include an extended example of reasons for displacement contrary to PAKs’ reasons, but does not problematize this. Reasons for displacement are offered as justifications in eight of the nine total concordance lines, and these include the example above, as well as definitions of *refugee(s)* from very specific, e.g. “people who have been removed or have removed themselves from the Gulf Coast in search of shelter, safety and food” (1 instance) or “people seeking refuge from natural disasters” (1 instance), to quite vague, e.g. “one who flees in search of refuge, as in times of war” (2 instances) or more simply, “one who seeks refuge” (3 instances). It is worth mentioning that three of these ignore exemplification of reasons for displacement not applicable to PAKs whilst making the case for applicability.

Sources for definitions given in this category differ substantially from the INAPPLICABLE category above; definitions from dictionaries (5 instances) and miscellaneous sources (4 instances) comprise this category, while political definitions are never used to argue *for* this label, only against it. All dictionary sources named—Webster’s, Oxford, and the American Heritage Dictionary—have also been used to argue for INAPPLICABILITY above, demonstrating a certain degree of ‘spin’ in news media. Further, MISCELLANEOUS attributions include two unnamed sources (2 instances) and direct quotes from news columnist William Safire (1 instance) and news editor Leland Senn (1 instance). In comparison to MISCELLANEOUS sources in the INAPPLICABLE category (a dictionary editor and a research fellow at a non-profit public policy

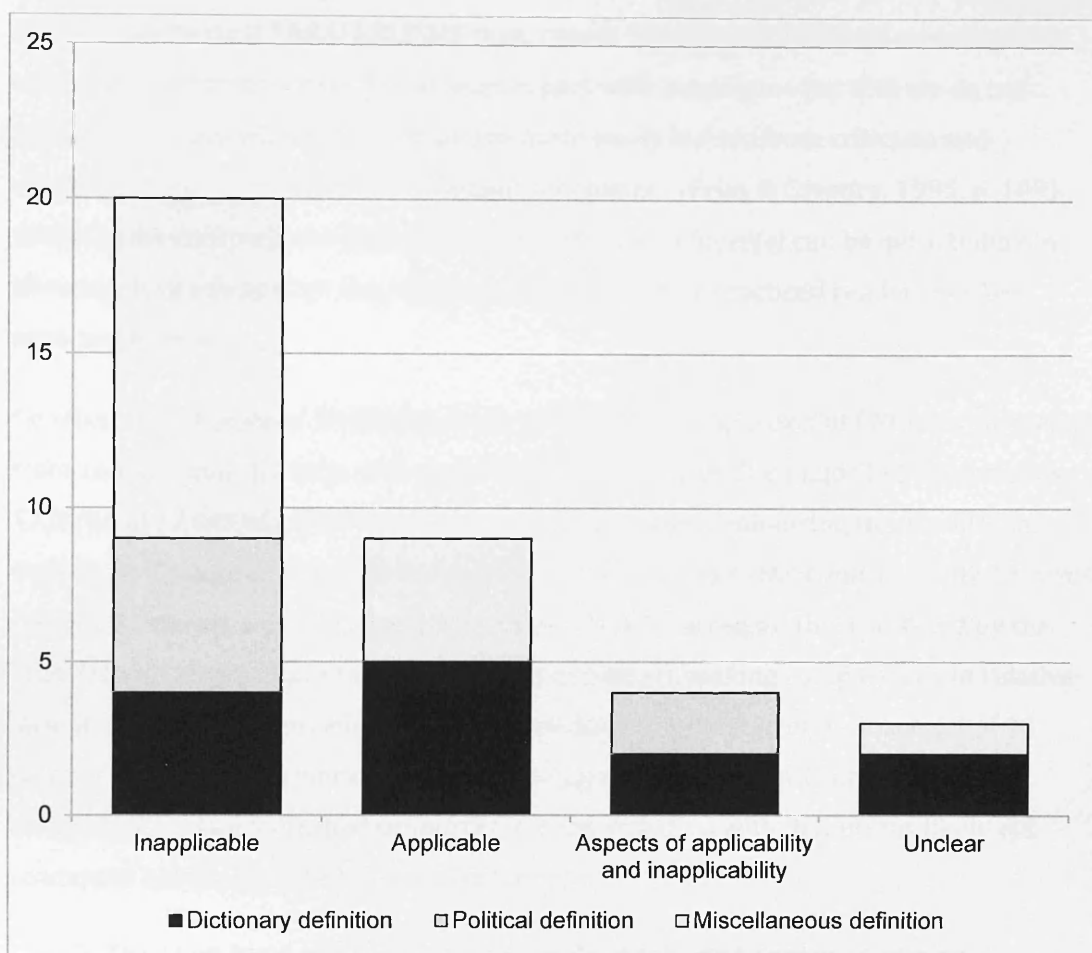
organisation), incorporation of two 'insider' media sources to reinforce institutional discourse supporting applicability of *refugee(s)* to PAKs is noticeably skewed.

#### **6.3.3.4 Aspects of both applicability and inapplicability**

Four additional cases convey reasons both for and against applicability of *refugee(s)* to PAKs, again foregrounding the debate regarding contested use. Of these, three cite domestic location of displacement as a feature of inapplicability, before expanding that *refugees'* reasons for displacement can be extended beyond escape from persecution or violence to include escape from danger or natural disaster. The fourth article employs a rhetorical style attributing a counter-argument to a reader, who "insisted that The Chronicle adhere to the 1951 U.N. Refugee Convention, which declares that a refugee is someone who escapes from their country to avoid persecution" before proceeding to the editorial argument, that "That may be fine when you deal with international affairs, but I'm not sure the newspaper should be bound by U.N. decree. I'd let the United Nations sit on the paper's style council when it puts The Chronicle on the U.N." ("Refugee" kicks up a storm', *San Francisco Chronicle*, 12 September 2005). This effectively dismisses incorporation of political definitions supporting inapplicability of *refugee(s)*, reclaiming the rights to employment of naming strategies on the terms of journalists alone.

#### **6.3.3.5 Definitions as supports of discourses**

As demonstrated in the sections above and visualised in Figure 6-13 below, most often, DEFINITIONS were included to argue that *refugee(s)* was an inappropriate naming strategy to apply to PAKs.



**Figure 6-13: Frequency of *refugee(s)* by stance category and definition source**

However, particular sources drawn upon differed depending on the preferred meaning, with political definitions being called upon more frequently in the INAPPLICABLE category, and dictionary definitions more frequently in the APPLICABLE category; miscellaneous definitions in the INAPPLICABLE category were more often from credentialed outside sources, whereas those in the APPLICABLE category were more likely to be sources from inside the publishing industry itself.

#### **6.3.4 Subcategory: PARALLELISMS**

The next subcategory refers to PAKs even more obliquely; rather than discussing this group in the scope of metalinguistic definition, PARALLELISMS are employed to liken a variety of attributes or experiences of PAKs to other (more 'traditional') *refugee(s)*. This introduces a strong element of intertextuality, which relies upon readers' knowledge of a vast and nuanced range of previous stories, events, and cultural touchstones, as well as the ability to draw connections between these and the current discussion. The analysis

of these intertextual PARALLELISMS then, covers “primary sociological meanings that we *make*—and make most often at least in part with language—but that we do not explicitly say, and which thus remain the more easily hidden from criticism and concerted social opposition or movement for change” (Fries & Gregory, 1995, p. 109). Studying the comparisons drawn between PAKs and *refugee(s)* can be quite telling in allowing us to extrapolate the relations that educated or practiced readers can be expected to infer.

To reiterate, 63 cases of PARALLELISMS, or linguistic comparison of PAKs to *refugee(s)* from another time or place, were found within the corpus. The majority (48 instances, 57.83%) of PARALLELISMS link PAKs to foreign *refugees*, enhancing the ‘traditional’ definition of *refugee(s)* despite the presence of arguments to the contrary. Only 17 cases (20.48%) reference domestic *refugee(s)*, and 10 of these regard those affected by the ‘Dust Bowl’<sup>17</sup> in the 1930s (see lines 30 and 31 below), making comparisons to relative lack of welcome, income, education, skill, and health in these groups. In a total of 18 cases, it is unclear whether the *refugee(s)* being compared to PAKs are foreign or domestic, though established proportions in combination with definitions likely led readers to believe that these were also foreigners.

30. The Dust Bowl **refugees** also faced the anger and contempt of many Californians, who watched as their state filled up with poor, often uneducated, unskilled and unhealthy **refugees**. (*Wall Street Journal*, 7 September 2005)
31. These are the very entitlement programs most needed by the sudden underclass of hundreds of thousands of hurricane **refugees** cast adrift like Dustbowl Okies. (*New York Times*, 3 September 2005)

Figure 6-14: Sample concordance lines of PARALLELISMS between PAKs and Dust Bowl victims

The appearance of PARALLEL references is interesting intertextually, but analysis of frequency does not expose the *meanings* created by inclusion. All concordance lines belonging to this category have additionally been coded as comparing or contrasting PAKs to *refugee(s)* (either domestic or foreign). A total of 57 concordance lines compare PAKs to *refugee(s)*, with only 23 contrasting the qualities of these two social groups; the stance of 14 cases were unclear. In PARALLELS to domestic refugees, 12 comparisons are accompanied by 3 contrasts. Most comparisons are drawn to large-scale diaspora (7), relation to other natural disasters (7), and American citizenship (4), with less

---

<sup>17</sup> A man-made environmental disaster taking place in the 1930s, where a period of droughts combined with longstanding poor farming techniques led to severe dust storms that devastated large areas of the American and Canadian prairie lands. See *Dust Bowl: The Southern Plains in the 1930s* by Donald Worster (2004) for more information.

frequent parallels being lower social class (3), experience of negative reactions to relocation from locals (1), and conception of *refugee(s)* as having undesirable qualities, specifically being unskilled and unhealthy (1). Domestic *refugee(s)* are contrasted to PAKs by virtue of differing ethnicity (1), more conservative political and religious outlooks (1), and receiving lesser support than PAKs (1).

While this category is relatively uncommon in comparison to the UNCONTESTED and CONTESTED categories of *refugee(s)*, its frequency is approximately 42% of that of the UNRELATED *refugee(s)*, meaning that a relationship between PAKs and other (so-called 'real') *refugees* is not an infrequent association. Even when the word is applied to non-American, non-PAKs, it can be (and frequently was) related to the discourse of Katrina. Association with established (foreign, often poor, minority) 'others' through constructions of PARALLELISM is another level of Social Sequestering, by which PAKs are identified through the frame of an accepted (threatening, unwanted) other. Another collocate—*camps*—has a similar effect.

### **6.3.5 Subcategory: CAMPS**

The CAMPS subcategory consists of 62 instances of *refugee(s)* occurring in 2- or 3-grams describing locations designated for temporary PAK resettlement or assistance. A frequency table of these phrases occurs in Table 6-7 below. The purpose of this layer of analysis, and indeed the thesis as a whole, is to explore the discourses surrounding (human) social actors; as such, an in-depth analysis of locational noun phrases formed by 2- and 3-grams such as *refugee shelters* and *refugee ghettos* would not be centrally relevant to the study.

Secondary component of bi- and tri-grams	Frequency
camp/s	26
center/s	14
shelter/s	11
assistance center	2
city	1
emergency center	1
encampment	1
ghettoes	1
reception center	1
relief center	1
relocation center	1
site	1
staging area	1

**Table 6-7: Frequency of secondary components of bi- and tri-grams following *refugee* in the CAMPS subcategory**

However, the appearance of ideologically loaded phrases such as *refugee camps* used in reference to temporary accommodation/assistance sites of PAKs does contribute to the UNCONTESTED discourse, drawing explicit connections between PAKs and the experiences of others in *refugee camps*.

### 6.3.5.1 Exploring the discourse of the *refugee camp/city*

Analysis of concordance lines from this subcategory reinforce the squalid conditions alluded to with the use of these terms, including “smelly, messy, dark and dank” (line 32), fetid, without running water (line 33), or containing human waste or dead bodies (line 34). Alternative naming strategies, such as *people* (line 32) and *victims* (line 33) sometimes populate the *refugee* locations, reaffirming that while *refugee* does not serve as the primary referential tactic in this subcategory, it does still contribute (negatively) to the semantic prosody of PAKs.

32. There are four levels of hell inside the **refugee city** of the Superdome, home to about 15,000 people since Sunday. On the artificial-turf field and in the lower-level seats where Montrel sat sweltering with her family, a form of civilization had taken hold—smelly, messy, dark and dank, but with structure. (*Washington Post*, 1 September 2005)
33. Hotels that withstood the storm have become **refugee camps**, with victims packed into fetid rooms with no lights, water or flushing toilets. (*New York Times*, 2 September 2005)
34. The center, from what officials can see, sustained little damage from the high winds. But since then, it has been a **refugee center** and a rather desperate one at that. Reports from inside the center describe dead bodies and overflowing toilets. (*New York Times*, 3 September 2005)

**Figure 6-15: Sample concordance lines belonging to the CAMPS subcategory**



Discursively placing PAKs into *refugee camp* conditions further distances their experiences from that of the typical newspaper reader and American, and construes PAKs as a locust-like population, transforming, with their forced resettlement, cities and structures into simulations of the Third World.

## **6.4 Summary**

In this chapter, a contested term was analysed in terms of its distribution over time and across publications, and categorised according to pragmatic uses. This investigation has led me to some recommendations regarding the feasibility of the utilised method, and to some conclusions regarding the construction of PAKs as *refugees* in Hurricane Katrina reporting. Both are detailed below.

### **6.4.1 Of findings**

Using the term *refugee(s)* as a naming strategy for PAKs was problematized within the articles included in the corpus. Despite a relatively high frequency of UNCONTESTED *refugee(s)* and a relatively low occurrence of the CONTESTED debate appearing within the corpus, the minority discourse ultimately prevailed.

Initial rejection by public figures, newspaper readers, and even some journalists changed the discourse around (and, I argue, the common meaning of) the word *refugee* in American English. A critical meta-discussion on the meaning of *refugee(s)* in America—and to Americans—took place between September and October 2005, and resulted in a dramatic decrease in the term's use. As we can see in a selection of post-Katrina dictionary entries in Figure 6-16, two out of three now stipulate that *refugees* leave their "homes" rather than their "country", which is the less frequently endorsed definition. The creation of an open debate in a popular forum such as the media is a heartening 'grass roots' outcome for the critical discourse analyst, who aims to affect change through exposure of power dynamics hidden in words.

	Longman Dictionary of Contemporary English Online (accessed 2012)	Collins Cobuild Dictionary Online (accessed 2012)	Oxford English Dictionary Online (accessed 2012)
<b>Refugee</b>	Someone who has been forced to leave <u>their country</u> , especially during <u>a war, or for political or religious reasons</u> .	Refugees are people who have been forced to leave <u>their homes or their country</u> , either because there is <u>a war</u> or because of <u>their political or religious beliefs</u> .	A person who has been forced to leave <u>his or her home</u> and seek refuge elsewhere, esp. in <u>a foreign country</u> , from <u>war, religious persecution, political troubles, the effects of a natural disaster</u> , etc.

Figure 6-16: A selection of three post-Katrina dictionary definitions of *refugee*

Alienating, othering discourses can be seen particularly through analysis of UNCONTESTED uses of the word. This was characterised by a great number of collocates of quantification (or the *topos* of numbers), and the movement and resettlement of *refugees* was often described in the context of water metaphors—both patterns which represent amorphous threat, dehumanizing *refugees* and “constructing them as an out-of-control, agentless, unwanted natural disaster” (Baker et al. 2008: 287). Given the context of reportage—a literal natural disaster characterised by floods—employment of metaphors exploiting the water-as-threat source domain seems to be in particularly poor taste. UNCONTESTED appearances of *refugee* were further typified by lack of agency and defined almost solely in terms of their number and movement.

Use of *refugee(s)* was deemed racist in the CONTESTED subcategory, though deeper analysis of its use in the UNCONTESTED and PARALLELISM subcategories indicates that, like *people* in Chapter 4, the racial aspect was most explicitly cited, but seems not to be as central as issues of classism and nationalism. More frequent than arguments of discrimination were those indicating that association with *refugee* crises or ‘problems over there’ were demeaning to Americans, who, despite being temporarily homeless, unemployed, or destitute, are still constructed as somehow superior to any foreign ‘other’ recalled with this naming strategy. Despite having been failed by their government (a feature of the risk society), PAKs did not appreciate being likened to social actors previously associated with moral panics. In other words, “the discomfort that so many people in the United States reportedly felt at hearing (or reading about) fellow U.S. citizens being called ‘refugees’ was revealing of their self-image. An image of power, prosperity, and self sufficiency had been proudly projected onto national and international scenes as testimony to the vitality of the ‘American dream’; all of this was

now under threat, thanks to Katrina” (Masquelier, 2006, p. 736). Despite the breakdown of social support structures available to PAKs, and the reality that they would have likely received more streamlined, ready assistance as UN-defined *refugees*, PAKs and journalists alike largely rejected the nomenclature, though there was no clear consensus reached on what was most appropriate to use in its place.

Even this discourse is not as straightforward and unified as the social context may suggest; as a relatively new nation composed of immigrants of various descriptions (including UN-designated *refugees*), the argument that this nomenclature was offensive to PAKs was, in itself, offensive to former *refugees* who had escaped problems ‘over there’, proudly adopted citizenship, and yet found their identities part of collateral damage; ‘real’ *refugees* are still further down in the hierarchy of ‘otherness’.

#### **6.4.2 Of method**

While the previous chapter focussed on high-frequency naming strategies uncovered in corpus-driven analysis, this chapter has explored a relatively lower-frequency, contested naming strategy with corpus-based analysis. *Refugee* is not in the top 100 most frequent common nouns, plural or unmarked for number in the Katrina corpus, but this is likely due to its extreme diachronic frequency change. Personal familiarity with both the underlying social context and the event itself indicated the unique treatment of this word. Other researchers with a similar advantage may likewise benefit from oscillation between corpus-driven and corpus-based analysis of naming strategies.

However, if contested discourses surrounding another event in another corpus event were not previously known to the researcher, or were not as explicit as those of *refugee(s)*, some methodological steps can be used to expose these:

1. **Frequency lists** will help researchers to identify all naming strategies, in order of their frequencies. In part-of-speech tagged corpora, this step can be streamlined by restricting frequency lists to (plural) nouns. Frequency lists of untagged corpora must be manually analysed, with researchers taking care to identify ‘people’ nouns, which become candidates for the next step.
2. **Distribution** patterns for high frequency, uncontested naming strategies tend to be generally similar over time. Steep drops or spikes in distribution data of one naming strategy in comparison to the others indicate a sudden change in preference, possibly attributed to a debate (see Figure 6-1, for instance). Cross-

referencing diachronic distribution with publication distribution will ensure that data has not been skewed to reflect one publication's reporting frequency or digitisation efforts.

3. If **collocations** of a naming strategy showing unusual distribution patterns include indicators of debate—negation, quotations, reporting verbs, etc.—this is quite likely a candidate for an isolated analysis. Review of a randomly thinned number of sorted **concordance** lines may then reveal contested discourses in these cases.

With these recommendations, I now move on to the second half of the analysis chapters of this thesis, which deal with the HIV/AIDS corpus. As this corpus is much larger, it has been necessary to combine all lessons learned and heavily refine the method to reflect a more accurate way of taking into account the part of speech, semantic category, and proportional salience of collocation patterns.

## **Chapter 7: The attributes and actions of people with AIDS/HIV**

### **7.1 Introduction**

This chapter marks the start of the second case study of the thesis, which is based on the representation of people with AIDS (PWAs) in a corpus of American news articles. The analysis contained in this and the following two chapters is designed to mirror the corresponding units in the Katrina case study. More specifically, Chapter 7 is a study of the most frequent naming strategy and an examination of the semantic domains associated with the social actors in the corpus; Chapter 8 contrasts a series of other high-frequency nomination strategies to discover covert meanings; and Chapter 9 explores one static variable in a diachronic investigation.

While the research questions and types of analysis remain broadly consistent across the two case studies, the method has necessarily evolved. Readers are reminded that the AIDS/HIV corpus is approximately 440% the size of the Katrina corpus, containing 161,144,924 words compared to Katrina's 36,736,679. While working with a corpus over four times the size of that used in the last case study has had the positive consequence of containing a greater variety of linguistic features and demonstrative examples of data to draw upon during analysis, a negative consequence has also occurred. Search (node) terms are *far* more numerous in the AIDS/HIV corpus, which has the knock-on effect of generating an unmanageably copious list of significant collocates, contained within too many concordance lines to manually code for variation. The collocates of the less frequent naming strategies in the smaller Katrina corpus were sufficiently categorised automatically by part-of-speech tags, and then manually semantically sorted. But, in dealing with hundreds of collocates per naming strategy (as is the case in the AIDS/HIV corpus, in the chapters ahead) it becomes necessary to develop a finer strategy of computer-assisted categorisation. Such a method is detailed below.

### **7.2 Methods**

In line with the methods sections in previous chapters, the section below will detail the process by which the search term was selected before providing information on frequency and distribution of the node over time. I will then outline the enhanced method of categorising collocates, which will be employed in all further analysis (Chapters 7, 8, and 9, all based on the AIDS/HIV corpus).

### 7.2.1 Identification of the search term

The method of identifying the search terms that serve as the basis for this chapter and the next one closely mirrors the methods employed respectively in Chapters 4 and 5 in the Hurricane Katrina subsection. To wit, a frequency list of POS-tagged lemmas was generated from the complete corpus and inspected for appearance of naming strategies for social actors.

Of the 1,000 most frequent lemmas in the corpus, the CLAWS tagger identified 483 as common nouns, either plural or unmarked for number. Manual inspection of each of these lemmata led to the identification of 101 'human' nouns with frequencies over 17,000. The full list of high-frequency human nouns will be revisited in Chapter 8; the initial exploration in construal of social actors in the AIDS/HIV corpus is focussed on the most frequent human noun: *people*.

*People* is the 38<sup>th</sup> most frequent lemma in the corpus, and the most frequent plural common noun or common noun unmarked for number, appearing 331,858 times in 102,772 texts (out of a total of 161,144,924 words in 166,576 texts). That this was likewise the most frequent 'human' noun in the Katrina corpus—therefore providing the basis for a precise comparison study—at first seemed extremely felicitous<sup>18</sup>. The lemma has fairly stable frequency over time, appearing 1711.84–2300.15 times per million words in each year 1981–2009. However, an initial compilation of all collocates of the word *people* in the AIDS/HIV corpus hinted that this word could not be analysed in as straightforward a manner as previously accomplished, given the vast differences between the corpora.

The AIDS/HIV corpus—which spans 28 years—is designed for diachronic study, as opposed to the 'snapshot' one-year Katrina corpus. Though each text in the AIDS/HIV corpus contains at least one mention of the epidemic, there is no guarantee that this is also the main topic (or what the text is predominantly 'about'). Various methods could have been used to enhance 'aboutness' in the corpus, including: 1) Restricting texts to those containing search terms in headlines (the stated 'theme' of newspaper texts), or 2) Increasing the minimum frequency of the original search terms (e.g. *AIDS* OR *HIV* must occur at least three times in a given text for it to qualify for inclusion). However, headlines appear in a wide variety of formats in *Factiva*, a feature that does not allow for

---

<sup>18</sup> *People* is also the most frequent noun in a subcorpus of COCA made up of newspaper texts, 1990-2009, with a frequency of 142,490.

reliable mining. Increasing the minimum frequency of search terms would be a superior strategy, but would eliminate a swath of texts that are short and/or contain many cohesion features (e.g. anaphoric references)—potentially a good number of texts indeed, given the genre! It is for these reasons that both options were rejected at the collection stage.

To narrow the results to just those *people* under analysis (PWAs) in the large resulting corpus, I instead utilise a very specific search term: “people with (AIDS|A.I.D.S.|HIV|H.I.V.|GRID|G.R.I.D.)”. While this term may immediately seem overly explicit, the size of the corpus does compensate: it returns 13,236 matches in 9,666 texts, more than enough to draw ‘baseline’ conclusions from. The majority of solutions arising from the ‘PWA’ search is, in fact, *people with AIDS*; see Table 1 below for more details. A search for “(PWA|PWAs)”, by contrast, only returns 140 results, and has therefore not been included as part of the final search term. Note that “people with GRID” and “people with G.R.I.D.” return no results. As such, this search string will be referred to going forward as *PWA/H*.

No.	Search result	No. of occurrences	Percent
1	people with AIDS	10755	81.26%
2	people with HIV	1938	14.64%
3	people with H.I.V.	543	4.1%

**Table 7-1: Frequency of all solutions arising from a search for “people with (AIDS|A.I.D.S.|HIV|H.I.V.|GRID|G.R.I.D.)” in the AIDS/HIV corpus, expressed by number of occurrences and percentage of all occurrences**

One drawback of using this search term rather than exploring the prosody surrounding *people* on its own is that *people with AIDS* or *people with HIV* are more neutral, ‘politically correct’ terms, and do indicate a certain consideration of the social group, as the state of having the disease is just one attribute assigned to a complete person, in contrast to naming strategies encountered in Chapter 8 (e.g. *AIDS victim*, *AIDS patient*), in which the social actor is essentialised by their illness. While this is acknowledged as a limitation, a view of the distribution of solutions for the search term over the decades (Table 7-2) and years (Figure 7-1) does show decreasing frequency over time. This aligns with decreasing ‘interest’ in AIDS/HIV, rather than an uptake in ‘politically correct’ terminology, such as PWAs, as indicated by the parallel decline of frequency of this naming strategy shown in Table 7-2.

Decade	Words in category	"people with (AIDS A.I.D.S. HIV H.I.V.  GRID G.R.I.D.)"		"(PWA PWAs)"	
		Hits in decade	Frequency/million words in decade	Hits in decade	Frequency/million words in decade
1980s	35,286,877	3644	103.27	69	1.9
1990s	77,681,146	7572	97.48	67	0.86
2000s	48,176,901	1996	41.43	4	0.08

Table 7-2: Frequency of solutions arising from search for "people with (AIDS|A.I.D.S.|HIV|H.I.V.|GRID|G.R.I.D.)" and "(PWA|PWAs)" in the AIDS/HIV corpus, expressed by decade of occurrence

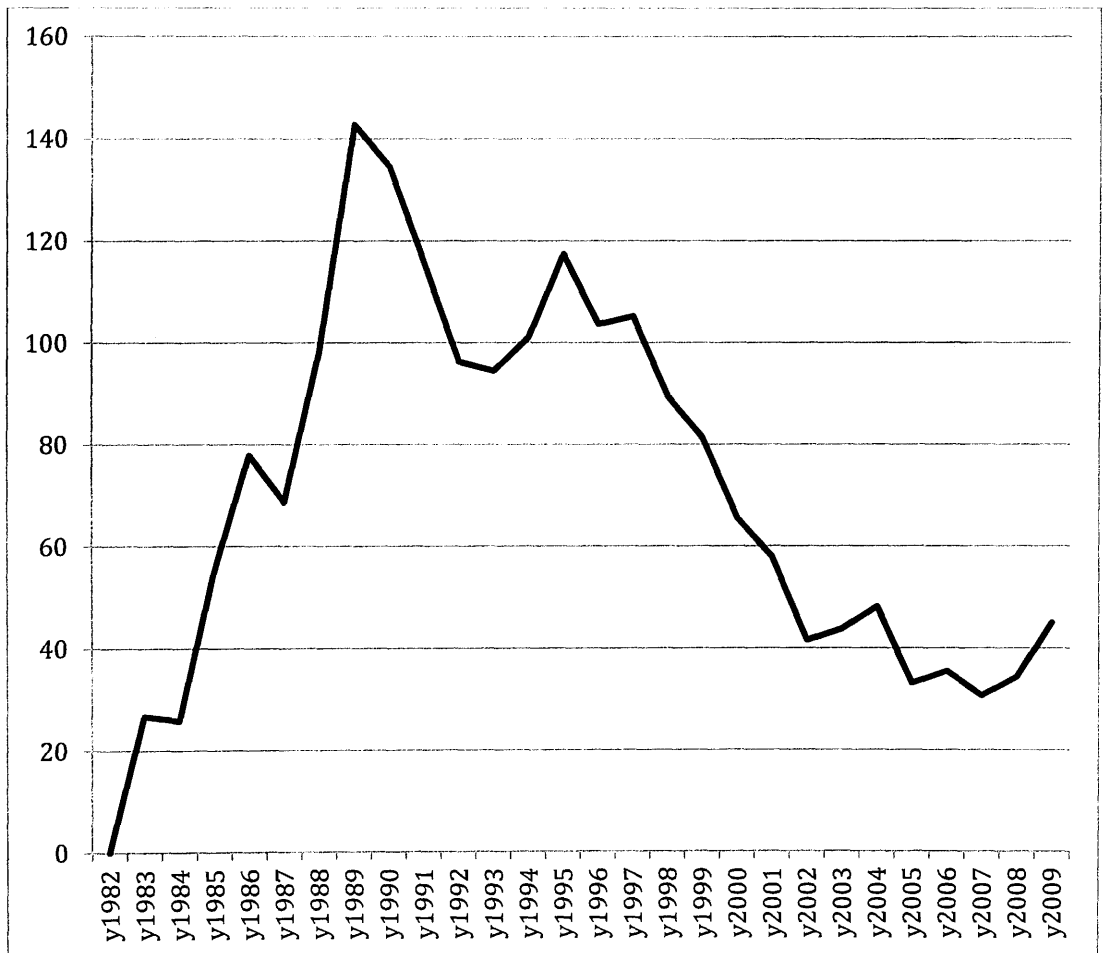


Figure 7-1: Frequency of solutions arising from a search for "people with (AIDS|A.I.D.S.|HIV|H.I.V.|GRID|G.R.I.D.)" in the AIDS/HIV corpus, expressed by frequency/million words in category (year)

Having considered the collocations and concordance lines associated with *people* in the AIDS/HIV corpus, I decided that this high-frequency term was too general to lead to highly conclusive results when considered within a corpus of this size<sup>19</sup>. While some

<sup>19</sup> For instance: B category collocates *healthy, retarded, overweight, and inoculate*; I category collocates *self-employed*; S category collocates *colored* (as part of National Association for the



choices at the corpus collection stage could have heightened the AIDS/HIV ‘aboutness’ of the corpus, too many texts would have been sacrificed from the sample. Acknowledging the limitations above, I consider the sheer frequency of results, in combination with the declining frequency pattern, as assurance enough of this search string’s viability as a basis of analysing ‘baseline’ sentiments associated with PWAs.

### **7.2.2 Categorisation of collocates, and selection of ‘salient’ domains**

In the given collocation spans (+/-1) and cut-off points of frequency (5) and statistical significance ( $MI \geq 3$ ,  $LL \geq 10.83$ ), *people with AIDS/HIV/H.I.V.* collocates with 38 lemmata. Each of these were run through the USAS semantic tagger, which assigns each word loaded into it a string of candidate tags, ranked in order of likelihood. Researchers do have the option of relying upon the machine-generated list of tags, which a program like WMatrix can then use to calculate items such as key semantic domains on the basis of the first candidate tag. However, to explore the feasibility and efficacy of this method, each collocate was manually checked against its tag, and re-tagged to a more appropriate one, if needed. In disambiguating semantic tags of polysemous words, each concordance line was checked, and the semantic tag associated with the sense used in the majority of cases was selected.

Automated tagging was largely quite successful here. Of the 38 lemma collocates of *PWA/H*, the broad semantic category of the first tag was deemed the most appropriate in 32 cases (84.21%)—see Appendix H for a full collocate table including all candidate tags. Researchers could, therefore, rely upon a completely automated use of semantic tagging with quite a high instance of accuracy. However, I have chosen here to manually disambiguate the remaining 15% of collocates where the first tag was not the most suitable. There are two categories of these cases.

In two instances (5.26%), a candidate tag that was automatically assigned but does not appear in the list as the first (or most likely) was selected as the most appropriate. For instance, *counsel* was initially given the semantic tags G2.1/S2mf S8+/Q2.2, with G2.1/S2mf (the most likely) being ‘Crime, law and order: Law and order’ with ‘People’. However, *counsel* in this (American) context does not mean legal *counsel* but emotional/psychological *counsel*, so this has been changed to S1.1.1 (Social actions, states, and processes) rather than G2.1 (Law and order).

---

Advancement of Colored People) and *iraqi*; and X category collocates *talented, knowledgeable, like-minded* are not related to PWAs, but to other *people* in the large corpus.

In the remaining five cases (13.16%), a tag not allocated by the USAS tagger was assigned to the collocate. This could happen in the case of unmatched items, e.g. *quarantining*, which was tagged Z99 'Unmatched' but was changed to A1.8, 'Inclusion/Exclusion' (matching *quarantine*, correctly automatically tagged A1.8). Using an automated semantic tagger means that some items might be successfully assigned a tag, but this tag will be inappropriate given the context. For instance, *serve* was initially given the semantic tag K5.1, for 'Sports'. However, *serve* in this context does not mean to *serve* a ball, but to *serve* a group of people by offering financial, legal, administrative, or emotional support, so this has been changed to an S category (S1.1.1: Social actions, states, and processes) rather than a K one.

### **7.2.3 Manual disambiguation of B category collocates**

Two special cases of manual disambiguation should be flagged here: *help* and *treat*. One of these falls into each of the categories above. The first candidate tag for *help* was B2 (Health and Disease), followed by S8+ (Helping/hindering). This collocate was initially categorised in the USAS B category: The body and the individual, on the basis of genre type, but upon analysing the collocate, it became clear that *help* was not necessarily (or even dominantly) given in medical cases, but rather spanning a range of social practices (see section 7.3.4 for discussion of agency and types of *help*). This has therefore been re-categorised as S8+, the second candidate tag. The other special case, *treat*, was inappropriately tagged E4.2+ A9+ (Happy/sad: Contentment, Getting and giving; possession). In this corpus, *treat* is polysemous, meaning both to offer (medical) treatment, and to behave toward others in a certain way (usually compassionately). If a cross-categorisation scheme were being used, this could therefore be both a B and an S category collocate, but as the choice to *treat* or refuse to *treat* a PWA/H makes up a large number of these concordance lines and is a social process, this has been categorised under S.

This leaves the USAS B category with an insignificant percentile appearance of collocates, which is surprising given the search term and the corpus used. It would seem logical that a search for PWA/H in a collection of texts each containing at least one reference to an illness will often be surrounded by further discussion of B category themes, e.g. health and disease. However, only one collocate (*infection*) lies strictly within the medical field, and is automatically and manually categorised in the B category. For this reason, *infection* was actually a somewhat 'expected' collocate, but its collocational position should have been an indicator of another pattern; it occurs in the

R1 position in 150/151 concordance lines, referring to “people with HIV infection”, an over-lexicalised, tautological term (Baker, Gabrielatos, Khosravini, McEnery, & Wodak, 2008). Only once (in the L1 position) is this describing *PWA/H* with other infections related to AIDS/HIV; in others, it is used to foreground the positive status—or the state of being infected—of the social actor being described. This is not isolated to one decade, but is used more in the 1990s (1.26/million words) than the 1980s (0.96/million words), with a decrease in the 2000s (0.39/million words).

It is my feeling that an automated semantic tagging accuracy rate of approximately 85% could be adequate for future researchers, particularly when the number of collocates (or key words, etc.) grow into the thousands. However, given the room for detail in this thesis, and my interest in testing the method for accuracy, moving forward, all semantic tags have been manually reviewed for accuracy and disambiguation.

### **7.3 Analysis**

The categorised collocates of *PWA/H* appear in Table 3 below, expressed in raw frequency of collocates in a given USAS semantic category and the percentage of overall collocates that these comprise. The highest proportion preference is in boldface. Those semantic categories not containing collocates for *PWA/H* are shaded out. The names of semantic categories represented by >10% of the collocates are in all capitals—these will be analysed below.

	Collocates in category	Percentage of overall collocates
<b>A GENERAL AND ABSTRACT TERMS</b>	<b>5</b>	<b>13.16%</b>
B the body and the individual	1	2.63%
C arts and crafts	0	0.00%
E Emotion	1	2.63%
F food and farming	0	0.00%
G government and public	0	0.00%
H architecture, housing and the home	1	2.63%
<b>I MONEY AND COMMERCE IN INDUSTRY</b>	<b>4</b>	<b>10.53%</b>
K entertainment, sports and games	0	0.00%
L life and living things	0	0.00%
M movement, location, travel and transport	2	5.26%
<b>N NUMBERS AND MEASUREMENT</b>	<b>11</b>	<b>28.95%</b>
O substances, materials, objects & equipment	0	0.00%
P Education	0	0.00%
Q language and communication	0	0.00%
<b>S SOCIAL ACTIONS, STATES AND PROCESSES</b>	<b>9</b>	<b>23.68%</b>
T Time	0	0.00%
W world and environment	0	0.00%
X psychological actions, states & processes	0	0.00%
Y science and technology	0	0.00%
<b>Z NAMES AND GRAMMAR</b>	<b>4</b>	<b>10.53%</b>
Total	38	100.00%

Table 7-3: Collocates of *PWA/H*, categorised into *USAS* semantic categories, and expressed as a percentage of the overall number of collocates

Considering only those semantic categories containing the highest number of collocates (here, over 10%) pinpoints the semantic fields of most salience to a given node or naming strategy. Using this method, researchers can effectively downsample their results while ensuring that the most representative patterns are not being omitted. This is a proposed improvement of subjective downsampling (wherein the researcher selects semantic fields that they find most 'interesting' personally to explore further), or application of statistically-motivated downsampling (whereby the researcher implements rather arbitrary statistical or rank order cut-offs, e.g. 'top 100 collocates'), which are not as objective or data-driven.

The semantic categories containing the highest proportion of collocates are: N: Numbers and measurement (11 collocates, 28.95% of overall); S: Social actions, states and

processes (9, 23.68%); A: General and abstract terms (5, 13.16%); I: Money and commerce in industry (4, 10.53%); and Z: Names and grammar (4, 10.53%).

### 7.3.1 A: General and abstract terms

In a general corpus of English, it could be guessed that A category (General and abstract terms) and Z category (Names and grammar) have the greatest potential to represent high frequencies of collocates. In the AIDS/HIV corpus, *PWA/H* collocate with USAS A category collocates only to create a negative prosody of ostracism (medical and social).

#### A category collocates of people with AIDS

<b>quarantining</b> [LL=111.339, MI=8.101], <b>exclude</b> [LL=70.493, MI=4.788], <b>quarantine</b> [LL=69.89, MI=5.29], <b>stigmatize</b> [LL=36.79, MI=5.857], <b>isolate</b> [LL=13.038, MI=3.195]
---

Table 7-4: USAS A category collocates of *PWA/H*, listed in order of descending log likelihood value

Two of the collocates shown above—*quarantining* and *quarantine*—are clearly related; the former lemma is the gerund (noun) form, and the latter is the verb form. Additionally, these collocate with low enough frequencies (12 and 13, respectively) to allow for tandem analysis in this section.

In the AIDS/HIV corpus, support of *quarantining* is linked to single individuals, or not attributed clearly, as in sample line 1, below. This is due to the fact that there was never any federal recommendation supporting quarantine (see line 2), which resulted in newspapers' necessarily distancing themselves from these calls. *Quarantining* is negatively evaluated as a method of "segregating and degrading" *PWA/H* (line 1), and individual supporters of *quarantining* practices are associated with the "lunatic fringe" (line 3) and [anti-gay] extremist groups (line 4). This practice is likened to "concentration camps" in line 4, the first WWII allusion I spotted in the corpus, but certainly not the last—see Chapter 9 for comparisons drawn between gay rights opposition in 1980s America and Nazi Germany. Calls for *quarantining PWA/H* are construed as, themselves, a public health risk, due to the associated decrease in likelihood of people seeking testing or treatment (and risking quarantine), and some instances—notably, by a radio personality and not by one of the many politicians/religious leaders named—were addressed directly by the Human Rights Commission (line 5).

1. "When AIDS first hit, we were all terrified it would be used to heighten discrimination," said McFeeley, who nine months ago moved from directing the Boston-based New England Human Rights Campaign Fund to directing the national HRCF. "Initially, there was talk about *quarantining people with AIDS, further segregating and degrading them,*" added McFeeley. (*Boston Globe*, 2 June 1990)
2. Federal health officials have never recommended *quarantining people with AIDS*. (*New Orleans Times-Picayune*, 17 June 1992)
3. One of the nine respondents with a favorable opinion, Ralph Chambers of Little Ferry, N.J., 32 years old, said he liked some of Mr. LaRouche's stands, such as his views on *quarantining people with AIDS*, or acquired immune deficiency syndrome. He also welcomed Mr. LaRouche's style, saying "It's good to have the lunatic fringe from both sides to keep the parties in balance." (*New York Times*, 18 April 1986)
4. When reports surfaced Sunday that Fong had donated \$50,000 to the Rev. Lou Sheldon and his Traditional Values Coalition, political insiders were stunned that Fong could support a group so opposed to gay rights that it advocated *quarantining people with AIDS*. "Mr. Sheldon's group would put people like me in concentration camps," said Ron Hill, a San Francisco health commissioner with AIDS. (*San Francisco Chronicle*, 27 October 1998)
5. That the Human Rights Commission "become involved" in public complaints against a KSFO radio personality who reportedly called for *quarantining people with HIV* or AIDS and told his audience that HIV can be transmitted by coughing. (*San Francisco Chronicle*, 24 January 1995)

Figure 7-2: Instances of *quarantining* collocating with *PWA/H*.

Similar patterns arise in close analysis of *quarantine*, the verb form of the collocate above. Medical officials speak out against calls for quarantine, calling it "inappropriate" in the British context (line 6), while in California, opponents of a proposition to maintain a state-wide list of HIV-positive individuals criticise the bill for its potential to ease quarantine and increase other forms of discrimination (line 7). *Quarantining PWA/H* is again associated with a 'fringe' social actor, politician Lyndon LaRouche, who in line 8 below is described as wanting to *quarantine PWA/H*, "colonize Mars", and "outlaw the death penalty" (a list of goals that are implied to be preposterous, which also tells us something about normed American views of capital punishment). Additionally, *quarantining* is linked with a former dark point in history, the detention of "girls with venereal disease in WWI" (line 9). Despite an initial working hypothesis that *quarantining* and *quarantine* both collocated with *PWA/H* due to press reiteration supporting calls to do so occurring early in the corpus, it is now clear that these always appear heavily distanced from the newspaper's own ideological standpoint, either through association with 'others', historical injustices, or negatively evaluated potential consequences of enactment.

6. "From a medical standpoint," he said, "It makes no sense. There is no legitimate reason to *quarantine people with AIDS* from the general public because the disease is not casually contagious." The British regulations are totally "inappropriate," he said, and the suggestion of a parallel with the black plague smacked of a "panic reaction" and public "hysteria." (*San Francisco Chronicle*, 23 March 1985)
7. However, that analysis and opponents of Proposition 69 say the initiative could make it easier to *quarantine people with AIDS* and to eliminate them from jobs as food workers and in schools. (*San Francisco Chronicle*, 23 May 1988)
8. [LaRouche] also wants to colonize Mars, outlaw the death penalty and *quarantine people with AIDS*. (*New Orleans Times-Picayune*, 18 October 1992)
9. "Flipping through this book," he said, "I saw a sentence about a proposal to *quarantine people with AIDS*, which she compared to the detention of girls who had venereal disease in World War I. I was thunderstruck." (*Boston Globe*, 29 January 2007)

Figure 7-3: Instances of *quarantine* collocating with *PWA/H*

The use of *isolate*—a word quite similar to *quarantine* in its enactment—collocates only five times with *PWA/H*, but its use is quite varied in context. In the late 1980s, it is used once (line 10) in reporting the results of a poll in which most respondents do not support *PWA/H* being *isolated* from society. Indictments of attempted isolation do appear twice, once offensively and once defensively; in line 11, a political action chair opposes a bill through accusations of attempted isolation of *PWA/H*, and in line 12, a housing representative defends a new development by stating that it is not meant to isolate *PWA/H*. In this small sample, it seems that in the 1990s, the word retained an emotional charge, and was negatively appraised. However, not all uses fall under this formula. In 2007, Republican politician Mike Huckabee defended an earlier 1992 appeal to *isolate PWA/H* by stating that this was not a call for a quarantine (line 13 below), marking the two as somehow dissimilar, though the boundaries between *isolating* and *quarantining* an individual are not made clear. By 2009 (line 14), a civil servant in Vietnam is seen to use the first demand for isolation in a direct reported speech act, framed within a new risk society panic: the outbreak of swine flu. As examples of this collocation are rare, it would not be prudent to draw firm conclusions on the basis of the concordance lines analysed, though I will observe that *isolating PWA/H*—though not without its supporters—seems to be perceived as an unpopular, Draconian response to the epidemic, but is set apart somehow from *quarantining* them, as evidenced by its appearance and its contrastive use in line 13.

10. Two-thirds [of poll respondents] (65 percent) say they would not balk at working alongside someone who has AIDS, and 71 percent do not want people with AIDS isolated from society. (*San Francisco Chronicle*, 23 November 1987)
11. “At the heart of those bills is a mean-spirited attempt to identify and isolate people with AIDS,” said Jim Lansdowne, the political action chair for the Harvey Milk Lesbian and Gay Democratic Club. (*San Francisco Chronicle*, 22 August 1987)
12. Day said the housing is not being developed to isolate people with AIDS. (*New Orleans Times-Picayune*, 3 November 1996)
13. [Mike Huckabee] began the day defending his record on “Fox News Sunday,” where he argued that when he called in 1992 for taking steps to isolate people with AIDS, he was not advocating a quarantine. (*New York Times*, 10 December 2007)
14. “I don’t know why we don’t isolate people with AIDS,” said a civil servant in Ho Chi Minh City, about 20 miles southeast of the village. “Even with swine flu we isolate people, and this disease is much more dangerous.” (*New York Times*, 14 October 2009)

Figure 7-4: Instances of *isolate* collocating with PWA/H

A fourth collocate in the USAS A category also exists on the cline of codified marginalization. Unlike *isolate*, *exclude* is not linked to the medical prosody of *quarantine*, and seems to be situated further onto the social side of the scale. The things that PWA/H are *excluded* from differ from those they are *isolated* from, though in some cases, the agents are quite similar. PWA/H are *excluded* in immigration policies (6 times), legal definitions of handicapped/disabled in bills and laws protecting these social groups (6 times), cover from health insurers (twice), and [not] from serving in the Peace Corps (once). Unlike *isolating* PWA/H, in all cases, *excluding* them is problematized. It is interesting to note that the most frequent agent—the government—is also the most frequent agent in the 1990s examples of *isolate*, which was also viewed very critically.

The final collocate—*stigmatize*—is interesting, despite its low frequency, due to its observable change over time. All instances of this collocation are reflected in concordance lines 15-20 below. In the earliest cases from the 1990s, it is laws (15), policies (16), and bills (17) that *stigmatize* PWA/H, and these are all opposed through direct and indirect quotation from various (named) expert sources. It appears that in the 1990s, *stigma* was something that could be restricted at the state/national governmental-institutional level. However, by the 2000s, it is not only legal documents that *stigmatize* PWA/H, but other institutions and other people. In line 18, it is the [Chinese] state-run media that is accused of *stigmatizing* PWA/H, and in lines 19 and 20, society at large is the source of *stigma*, which in turn is blamed for the spread of the epidemic. In these larger concordance lines, social stigmatization of PWA/H is ‘given’



information, and not problematized or called into question. Solutions are not as easily offered in the 1990s. I suggest that this is due to the later (2000s) reports detailing information regarding *PWA/H* being *stigmatized* in foreign contexts rather than domestic ones, which could indicate implicit American criticism of other countries' now-'archaic' treatment of *PWA/H*.

15. "This law is bad policy," said Murphy, whose organization offers legal advice to people with AIDS." It stigmatizes people with HIV. No other laws single out people with a disease. (*New Orleans Times-Picayune*, 22 November 1992)
16. The union disagrees with a city of Portland policy issued in June that forbids dispatchers from broadcasting that a person has an ailment such as HIV, the virus that causes AIDS, or hepatitis. [...] The police chief said officers should always take precautions to protect themselves, but he doesn't want to stigmatize people with AIDS. (*Seattle Post-Intelligencer*, 18 October 1995)
17. Mr. Gottfried, Assemblywoman Deborah J. Glick of Manhattan and other opponents, including groups that help AIDS patients, argued that ordering H.I.V. tests for defendants is an invasion of privacy that does not help crime victims and serves only to further stigmatize people with AIDS. (*New York Times*, 12 January 1996)
18. Li Xiguang, a prominent journalism scholar at Beijing's Qinghua University who has criticized the state-run media for stigmatizing people with AIDS, noted that Wen's visit took place less than a month after former president Bill Clinton visited China and was shown on television hugging a young man with HIV. (*Washington Post*, 2 December 2003)
19. Authorities in Russia have no information on the mode of transmission for at least half of the new cases, preventing clear analysis of the disease's trajectory through a society that continues to stigmatize people with HIV, the study found. (*Washington Post*, 12 January 2005)
20. "We know that stigmatizing people with HIV, or who are presumed to have HIV, is one of the root causes of the pandemic. So why would we issue statements that might exacerbate that?" [Maurice I. Middleburg, acting president of EngenderHealth, a 62-year-old public health charity working in 16 countries] said. (*Washington Post*, 18 May 2005)

Figure 7-5: Instances of *stigmatize* collocating with *PWA/H*

Close, qualitative analysis of collocates in the USAS A semantic category has been enlightening. One of the advantages of combining corpus linguistics with critical discourse analysis is that quantitative patterns appearing to tell one story can be found to be telling quite another when analysed qualitatively. A cursory view at the collocates from this category indicates that *PWA/H* are associated with a negative semantic prosody of social and medical segregation. However, analysis of the individual concordance lines has shown that, in fact, segregation of *PWA/H* is itself very negatively evaluated, and often associated with 'others' (e.g. political extremists), unpopular policies from powerful institutions (e.g. immigration laws), and even heightened risk (e.g. contributing to spread of the disease).

### 7.3.2 I: Money and commerce in industry

The next broad semantic category containing more than 10% of the collocates of *PWA/H* is I: Money and commerce in industry. This category also featured in the Katrina analysis (see Chapter 4.3.5) where markers of low income and African-American ethnicity often coordinated, as did markers of middle- to high-income and Caucasian ethnicity. But unlike the Katrina corpus, not even a minority of the I category collocates represent *PWA/H* on the upper end of the income scale; all four (*low-income*, *poor*, *indigent*, and *needy*) associate *PWA/H* with poverty, meaning that collocates from this category consistently orient *PWA/H* in an additional out-group, or show deviancy doubling.

There are two arguments that may explain this disproportionate distribution. The first is that to be 'poor' is a marked condition in newspaper texts. Habitual newspaper readers generally have higher levels of education and income than the national averages (The Pew Research Center, 2012), and the relatively low occurrence of intermediate- and high-income markers could signal that these are the 'norm' or default reader position. The other explanation is that lack of income is of particular salience to *PWA/H* in the corpus; this can only be established through closer analysis.

#### I category collocates of *people with AIDS*

**low-income** [LL=364.942, MI=6.116], **poor** [LL=121.677, MI=3.103], **indigent** [LL=92.118, MI=6.172], **needy** [LL=28.688, MI=3.947]

Table 7-5: USAS I category collocates of *PWA/H*, listed in order of descending log likelihood value

The collocates above appear in a wide range of frequencies; *low-income* is the most frequent (with 56 instances of collocation), followed closely by *poor* (48 instances), whereas *indigent* (14 instances) and *needy* (only 8 instances) are far less frequent, relatively. In addition to the absence of middle- and upper-income collocates, I wondered whether the two corpora differed in relation to the ordering (or strength of association) of the various collocates. With the semantic selection scheme in place, it is now possible to do a much stricter comparison between *low-income* collocates of *people* [with AIDS] in the Katrina corpus versus the AIDS/HIV corpus. I revisited the Katrina corpus data and applied the WMatrix/USAS tagging system developed between working on the two case studies. The Katrina corpus collocates contained a much greater variety of income markers, including upper-income and middle-income, which was not witnessed in the AIDS/HIV corpus. In Table 7-6, items not collocating with the search term(s) with statistical significance in a particular corpus are hashed out.

Collocate lemma	AIDS/HIV corpus		Katrina corpus	
	MI value	Rank	MI value	Rank
lower-income			5.084	1
indigent	6.172	1		
low-income	6.116	2	4.595	2
needy	3.947	3	4.416	4
poor	3.103	4	4.48	3
disadvantaged			3.485	5
impoverished			3.465	6

**Table 7-6: Comparison between ‘lower-income’ collocates of PWA (AIDS/HIV corpus) and *people* (Katrina corpus), expressed by Mutual Information value and relative rank within the lower-income collocate set**

While the Katrina corpus shows a greater variety of low-income collocates of *people* than the AIDS/HIV corpus has of PWAs, the preference is largely the same. As shown in Table 7 above, *low-income*, *needy*, and *poor* are collocates shared between the two corpora, with similar preference displayed with MI values (despite a small inverse in the positions of *needy* and *poor* between the two). Though it is technically categorised under the H category, I would also like to note that *homeless*—a related but not identical social group—collocates with PWA/H (AIDS/HIV) and *people* (Katrina corpus) with an identical MI value (4.421). Markers of employment—e.g. *unemployed*, *self-employed*—occurring in the Katrina corpus are not present here. Notably, the two corpora do not share their highest MI-ranked collocates; *indigent* (highest in AIDS/HIV) does not feature with statistical significance in the Katrina corpus, and *lower-income* (highest in the Katrina corpus) is not significantly used to modify PWA/H in the AIDS/HIV corpus. To find out whether this is an effect of the ‘sense’ of these words, we move now to analysis of the concordance lines.

Organising collocates by automated semantic categorisation, and selecting the highest frequency categories (or those of highest ‘salience’) have been extremely helpful methodological strategies in this instance: with only 126 concordance lines associated with collocations in the I category, I am easily able to read extended context in every case, and draw comparisons between the related meanings arising. In doing so, it becomes clear that low-income PWA/H are represented in the frame of wanting, needing, ‘deserving’ (a concept revisited when discussing S category collocates in this chapter, below) or receiving aid of various descriptions, often overlapping or interconnected within the same context: governmental/legislative (60 concordance lines), healthcare (54), residential (39), and miscellaneous (18). It is of note that the theme of governmental/legislative action is more common than that of provision of healthcare or residential services (though these do overlap with some frequency),

indicating that *low-income PWA/H* are mainly the responsibility of (or at the mercy of) state and federal government.

The concept of (low-) income was explored at great length in Chapter 4, and will be revisited in both Chapters 8 and 9. For longer discussions on the consequences of collocation with this semantic domain, please refer to sections 4.3.5.1 and 5.5.1.

### 7.3.3 N: Numbers and measurement

The USAS semantic category representing the greatest percentage of collocates (by a margin of approximately 10%) is N: Numbers and measurement, under which 11 or 28.95% of all collocates of *PWA/H* are categorized. This largely echoes the semantic preference of earlier naming strategies in the Katrina corpus, where one of the major *topoi* used to negatively depict PAKs was the *topos* of numbers (accompanied by flood and [bursting] container metaphors).

N category collocates of <i>people with AIDS</i>	
<b>many</b>	[LL=809.593, MI=3.425], <b>235470</b> [LL=79.201, MI=12.15], <b>5000</b> [LL=40.34, MI=4.298], <b>600</b> [LL=36.253, MI=3.966], <b>10000</b> [LL=30.076, MI=3.765], <b>1000</b> [LL=27.39, MI=3.069], <b>200000</b> [LL=18.247, MI=3.986], <b>40000</b> [LL=17.322, MI=3.886], <b>2000</b> [LL=15.651, MI=3.166], <b>250</b> [LL=14.584, MI=3.416], <b>3000</b> [LL=13.722, MI=3.32]

Table 7-7: USAS N category collocates of *PWA/H*, listed in order of descending log likelihood value

In the AIDS/HIV corpus, the majority of the USAS N category collocates are cardinal numbers, which give various estimations of the numbers of infections rates, estimated sizes of risk groups, etc. Enumeration of ‘cases’ of AIDS/HIV rather than discussing the human lives represented by the figures is a dehumanizing and alienating discursive strategy (Gabrielatos et al., 2008). Having already discussed this pattern in the PAK chapter, however, I refer readers back to sections 4.3.7 and 5.5.2 for a review of the implications of association with the *topos* of numbers, and move on to another section with more unique collocates in the HIV/AIDS corpus.

### 7.3.4 S: Social actions, states and processes

In this section I explore the S collocates of *PWA/H*. In another echo of the Katrina corpus results (see Chapter 5.5.3.3), collocates from the USAS S category convey the stigma of charity (Fothergill, 2003); *PWA/H* are construed as being the objects of others’ efforts to *help, treat, assist, protect, and counsel*. Two collocates do exist on the cline a bit further from total disempowerment—*serve* indicates that *PWA/H* are being aided in a manner that places them on equal (if not higher) footing as those serving them. Likewise, *enable*

signposts an ultimate model of enhanced self-sufficiency being strived toward through charity with *PWA/H*. *Deserve* keys into existing social understandings of basic human rights and access to those, which will be explored shortly below, but it is of note that what is *deserved* cannot be gained on one's own, but must be given by others. Therefore, it is only one collocate in the SOCIAL ACTIONS, STATES, AND PROCESSES semantic category that shows empowerment of the *PWA/H*, and this is *coalition*. In this section, I explore the various ways (presence or lack of) power is negotiated through collocation with collocates of Social actions, states, and processes.

**S category collocates of people with AIDS**

**coalition** [LL=1987.09, MI=7.121], **help** [LL=1585.338, MI=4.251], **treat** [LL=799.269, MI=4.795], **assist** [LL=372.346, MI=5.672], **serve** [LL=276.663, MI=3.59], **protect** [LL=139.474, MI=3.557], **deserve** [LL=37.266, MI=3.563], **enable** [LL=36.592, MI=3.994], **counsel** [LL=25.922, MI=3.396]

**Table 7-8: USAS S category collocates of *PWA/H*, listed in order of descending log likelihood value**

Starting from the collocate with both the highest LL and MI values, *coalition* stands apart in terms of statistical significance and social meaning. This is clearly the item most loaded with social empowerment, but unfortunately it is not the product of a social process, but rather a named collectivized social actor (van Leeuwen, 2005) in its own right. In all cases, *coalition* occurs as part of the proper noun *People With AIDS Coalition*, which undertakes processes such as “electing”, “running”, etc. While this pattern is interesting, the focus of this study is not entities realised through collectivization in proper nouns, but the attributes of- and actions done by- or to [groups of] individual humans as common nouns. Therefore, this collocate must be set aside in this study, but is very well-suited to be taken up again in a later, further analysis.

In the following section, I refer to manual analysis performed on each concordance line produced by the collocation of *help*, *treat*, *assist*, *serve*, *protect*, *deserve*, *enable*, or *counsel* with *PWA/H*. Each concordance line was annotated for three layers of features:

1. Agency of *PWA/H* – Is the *PWA/H* the social actor the active participant of the collocating process?
  - a. Yes: Occurred in 1/777 cases.
  - b. No: Occurred in 776/777 cases.
2. If the *PWA/H* is not the agent (99.87% of the time), who is the agent? Agents were grouped into categories (illustrative concordance lines appear below):
  - a. Organisation/Business (32.30%): Charitable organisations, foundations, religious groups, private businesses, corporations, etc.

- i. *Sample concordance line:* "We expect to be serving 200 to 250 people by the end of the year,' said Matthew Hamilton, executive director of Open Hands Chicago, an agency that helps people with AIDS." (*Chicago Sun-Times*, 30 December 1988)
- b. Individual person (24.84%): A single human social actor, not collectivized.
  - i. *Sample concordance line:* "SHE CAN'T PLACE exactly where the passion comes from, but 23-year-old Emily Pawul always had a desire to *help people with AIDS.*" (*Seattle Post-Intelligencer*, 19 January 2008)
- c. Government/Legislation/Program (17.63%): City, State, or Federal departments, legal policies and laws, government programs, etc.
  - i. *Sample concordance line:* "And budget-slashed programs must be rebuilt, including public health programs that *help people with AIDS, substance abuse, and teen pregnancy.*" (*Boston Globe*, 7 May 2005)
- d. Medical Treatment/Centre (12.23%): Drugs, treatment regimes, health clinics, etc.
  - i. *Sample concordance line:* "Many people with HIV never develop AIDS, and new medicines are *helping people with AIDS* live longer and better lives." (*St. Louis Post-Dispatch*, 16 November 1997)
- e. Financial assistance (9.78%): Monetary donations, fundraising efforts, etc., but *not* government benefits (which are categorised under b.)
  - i. *Sample concordance line:* "Such accelerated death benefits could *help people with AIDS, cancer and other severe illnesses.*" (*The New York Times*, 29 March 1996)
- f. Events (0.90%): Organised gatherings intended to raise funds, awareness, or other forms of civic support.
  - i. *Sample concordance line:* "The sixth annual Greater Boston Clergy Walk to *help people with AIDS* will take place today in Arlington and Lexington." (*Boston Globe*, 11 November 1991)
- g. Food (0.47%): This category contains only one case found in the concordance line below.

- i. *Sample concordance line*: "In many cases, nourishing meals enable people with AIDS to fight the recurring bouts of illness at home..." (*The Boston Globe*, 23 October 1992).
      - h. In 1.80% of concordance lines, the agent was Unclear.
        - i. "The topics ranged from questions about health care (Streisand favors a single-payer system such as the one in place in Canada); how Judaism has affected her work; and what she thinks should be done to help people with AIDS." (*Boston Globe*, 4 February 1995)
3. Finally, are applicable actions being assigned a primary agent, other than the PWA? (In 7/393 total *help* concordance lines)
  - a. In how many of these cases is the PWA being *helped/assisted/enabled* by actions attributed to the primary agent?
    - i. This is almost always realised through the use of a by-phrase, e.g.: "The photographs are part of 'A Tasteful Affair Four,' which benefits Food Outreach, an agency that helps people with AIDS by providing meals." (*St. Louis Post-Dispatch*, 24 April 1994)
    - ii. In one case, a prepositional phrase gives more information about the type of help being provided: "Jackson has changed from 'a strung-out junkie living in the Tenderloin' to a volunteer at St. Mary's Hospital who helps people with AIDS into public assistance programs and support groups." (*San Francisco Chronicle*, 29 November 1990)
  - b. By contrast, in how many cases is the PWA/H being *helped/assisted/enabled* to undertake further actions of their own agency? (In 88/393 total *help* concordance lines)
    - i. Most frequently, the PWA is the secondary agent in the subordinating clause, with futurity marked by the bare form of the verb, e.g.: "There was an ample supply of antiretroviral drugs, which can help people with AIDS stay healthy for years." (*Newsweek*, 1 October 2007)
    - ii. In eight concordance lines, the verb attributed to PWA/H as the potential agent appears in full infinitive form in a non-finite secondary clause, making this grammatically more dependent on the first clause, enhancing futurity and potentiality. For example:

“Lenny Larson of Seattle has *helped* **people with AIDS to cope with the deadly disease**, providing support through the Northwest AIDS Foundation, the Chicken Soup Brigade and the Seattle AIDS Support Group, three organizations he co-founded.”  
(*Seattle Post-Intelligencer*, 3 April 2002)

From the very first step of concordance annotation, it is clear that the same lack of agency characterising the naming strategy throughout this chapter’s analysis continues to dominate the discourse of *PWA/H*. In nearly every concordance line featuring collocates of statistical significance from the S category (coincidentally all verbs, save for the proper noun discussed and excluded above), *PWA/H* are the passive recipients of social actions. Behind the USAS N category of Numbers and measurement (whose high incidence of statistical significance has been discussed above), the S category is the most highly populated by collocates for *PWA/H*. This indicates, then, that a dominant discourse surrounding this group is one of passivity—or powerlessness.

In Step 2 of concordance analysis, the agents acting upon *PWA/H* were calculated. In Table 9 below, these agents are shown with: a) the percentage of all instances of the collocation between *PWA/H* and the verb in which this actor is the agent; and b) the raw frequency of times they undertake the collocate verb actions (shown in brackets). The largest percentile preference of agent for each verb is indicated in boldface; preference of each verb collocate is italicised. Items in both bold and italics have a mutual preference.

The most common agent is an Organisation/Business (frequency: 251, agency in all S category concordance lines under analysis: 32.30%), followed by Individuals (193, 24.84%) and Government/Legislation/ Program (137, 17.63%). Large collective agents (Organisation/Business and Government etc.) account for approximately 50% of all USAS S verb collocate concordance lines, whereas Individuals are named agents less with less than half of that frequency. It appears that social support—specifically by way of serving/assisting/protecting/enabling—is not conveyed in the HIV/AIDS corpus texts as something that can be given by the average civilian, but instead, as something that must be distributed through institutionalised means. The connotations associated with various agents differ, of course, by their frequency of association with the various collocates under analysis, just as the semantic meanings of these processes themselves are constructed by their frequent relationship with a certain agent.



Agent	help	treat	serve	assist	protect	enable	counsel	Total
Organisation /Business	151 (19.43%)	16 (2.06%)	46 (5.92%)	36 (4.63%)	0 (0.00%)	1 (0.47%)	1 (0.47%)	251 (32.30%)
People	90 (11.58%)	84 (10.81%)	3 (0.39%)	11 (1.42%)	0 (0.00%)	0 (0.00%)	5 (0.64%)	193 (24.84%)
Government/Legislation/ Programme	56 (7.21%)	4 (0.51%)	18 (2.32%)	7 (0.90%)	45 (5.79%)	6 (0.77%)	1 (0.13%)	137 (17.63%)
Medical Treatment/Centre	44 (5.66%)	32 (4.12%)	16 (2.06%)	0 (0.00%)	0 (0.00%)	2 (0.26%)	1 (0.13%)	95 (12.23%)
Financial Assistance	37 (4.76%)	24 (3.09%)	5 (0.64%)	9 (1.16%)	0 (0.00%)	0 (0.00%)	1 (0.13%)	76 (9.78%)
Events	7 (0.90%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	7 (0.90%)
Food	3 (0.39%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1 (0.47%)	0 (0.00%)	4 (0.51%)
Unclear	5 (0.64%)	9 (1.16%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	14 (1.80%)
Overall frequency of verbal collocates	393 (50.58%)	169 (21.75%)	88 (11.33%)	63 (8.11%)	45 (5.79%)	10 (1.29%)	9 (1.16%)	777 (100.00%)

Table 7-9: Contextual agents of verbal collocates help, treat, serve, assist, protect, enable, and counsel, expressed by number of instances of concordance lines of a given agent and percentage of 777 overall concordance lines of these verb collocates

The verb lemma *help*—the most frequent collocate overall—strongly prefers the Organisation/Business category of agent and this preference is mutual, with this agent also occurring most often with this action. Individuals and Government/Legislation/Programme are also frequent agents of *helping* of which *PWA/H* are the beneficiaries, which makes the sort of ‘help’ indicated quite variable. The exact nature of this *help* is indicated in only 89 out of 393 concordance lines (22.65%), perhaps indicating that the reader is able or intended to infer the type of aid based on the agent providing it, or to comprehend it after reading the extended context of the entire article.

An interesting pattern does appear in the 95 concordance lines where the type of *help* being supplied is expounded upon. In 87 cases, the agent of the main clause is *helping PWA/H to do* something in the following subordinate infinitive clause, marking them as the agent in the possible secondary action. By contrast, in only six cases, the agent of the main clause is *helping PWA/H by* performing a secondary action in the subordinate clause (providing circumstances or range), thereby maintaining agency. Two concordance lines provide both.

Agent	Total freq. of agency with <i>help</i>	<i>help + by</i>	<i>help + secondary clause</i>
Organisation /Business	158	2 (1.27%)	33 (20.89%)
People	90	2 (2.22%)	10 (11.11%)
Government/Legislation/Programme	51	2 (3.92%)	17 (33.33%)
Medical Treatment/Centre	44	2 (4.55%)	22 (50.00%)
Financial Assistance	36	0 (0.00%)	3 (8.33%)
Events	7	0 (0.00%)	1 (14.29%)
Food	3	0 (0.00%)	1 (33.33%)
Unclear	5	0 (0.00%)	1 (20.00%)
Total	393	8 (0.90%)	89 (11.33%)

**Table 7-10: Common agents *helping PWA/H*, with frequency of *by*-phrases and secondary clauses**

In Table 7-10 above, the frequency of concordance lines containing secondary information about the nature of *help* (whether this is on the part of the main agent, or *PWA/H*) is shown. As we can see, help from agents in the Medical Treatment/Centre category is the most empowering. With the aid of treatment, *PWA/H* are construed as being enabled, among other things, to “sort through a maze of medical protocol”, “fend off infection”, “regain muscle mass”, and “live longer and healthier lives”. The Government/Legislation/Programme agent empowers *PWA/H* to “navigate through city agencies” and “gain access to new treatments and therapies” but also to *help* others,

giving them civic responsibility to *help* them “from transmitting the disease to others”. Agency is also often distributed when the main agent is from the Organisation /Business category. With assistance from this agent, *PWA/H* can “cope with problems associated with terminal illness”, “to come out and not commit suicide”, “to cook and get about town” and “care for their pets”. Individuals are the only agents who have a proportion of help + BY concordance lines (2, 2.22%) even a fifth of the help + TO concordance lines (10, 11.11%). This strengthens the argument that details on the range or type of help is only given when this cannot be inferred from the identity of the agent.

The extremely high frequency of *help* indicates that this is the preferred S category verb in the corpus, and in viewing its distribution across agents in Table 9 above, it seems that this is also typical of verbal preference in newspaper discourse about AIDS/HIV. Agents *helping PWA/H* are generally large organisations or governmental agencies, and this is true across nearly all of the other verb collocates. To learn more about this collocate and to test the typicality of its use, it would be helpful to compare it to a collocate showing a different pattern. In viewing the preferences of the verbs (in bold) and of the agents (italicised) listed above, it becomes clear that the most frequent collocate *help* and the second-most frequent collocate *treat* appear in quite dissimilar agentive contexts, indicating a divergent semantic ‘sense’ being used. For instance, Agencies/Organisations/Centres are more likely to *help PWA/H*, whereas Individuals are more likely to *treat* them. Below, more nuanced comparisons between the uses of these collocates in the various categories are made.

While the Agencies/Organisations/Centres category is the most frequent overall, the actual agents contained within this category are quite different depending on whether *help* or *treat* has been used. It is most frequently volunteer groups, not-for-profits or charitable organisations that *help PWA/H* in various manners, e.g. to find residence, lobby for their rights, create networks of carers within their respective cities (see concordance lines 21 and 22 below for examples). Perhaps predictably, named agents in this category who *treat PWA/H* are more likely to be medical centres, hospice groups, and organisations of practitioners (lines 23 and 24). This is not a particularly interesting finding unless considered in light of the percentages of overall concordance lines this category represents. *PWA/H* are *helped* by Agencies/Organisations/Centres in close to half of all concordance lines for this collocation, but only *treated* by collective groups of practitioners, etc. approximately 10% of the time. This is somewhat contrary to my

expectation, that *helping* would be a more individual affair, with *treatment* being provided by large entities.

21. Forty nonprofit organizations helping people with AIDS throughout the region were awarded grants totaling \$263,600 yesterday by the Northwest AIDS Foundation. (*Seattle Post-Intelligencer*, 13 December 1995)
22. "We expect to be serving 200 to 250 people by the end of the year," said Matthew Hamilton, executive director of Open Hands Chicago, an agency that helps people with AIDS. (*Chicago Sun-Times*, 30 December 1988)
23. The project supports small groups who treat people with AIDS and do alternative research. Such groups include the Society for Tibetan Medicine, God's Love: We Deliver, the AIDS Education Program at Harvey Milk High School, the Lesbian and Gay Community Services Center, the Hispanic AIDS Forum, the AIDS Education Program at the Association of Hispanic Arts, and the Community Health Project. (*New York Times*, 16 December 1987)
24. A clinic to treat people with AIDS and the HIV-virus will open Wednesday in the Warren County Health Dept. (*USA Today*, 27 July 1992)

**Figure 7-6: Instances of the Agencies/Organisations/Centres agent in concordance lines where *help* and *treat* collocate with PWA/H**

This pattern is continued when looking at the next most frequent category of agent, the Individual, though more contrasts do appear in the concordance lines. Individuals almost always choose to *help* PWA/H, citing religion or a sense of charitable responsibility for their actions (see line 25 below). I only find one case where an individual does not choose to *help* a PWA/H (line 26), and this is presented in the form of a personal attack (and possible allegation of homophobia) on the agent allegedly denying help, not as a direct quotation from the agent himself. Here, it appears that to *help* PWA/H is charitable and admirable, whereas refusing or neglecting to help them is a character flaw. By contrast, the Individual agent is the most frequent in the *treat* concordance lines, but is also much more frequently negated. As often as individual doctors and practitioners state that they are willing to *treat* PWA/H, they are depicted as feeling somehow inept in doing so (line 27), or in continually ("still") denying the task outright (line 28). When Individual practitioners refuse to help PWA/H, this is not seen as a moral issue, but rather as an objection to fulfilling all areas of one's position. *Treating* PWA/H is expected and demanded (despite practitioners' reservations), but *helping* them is only undertaken by ethical, well-appraised Individuals.

25. As one of the volunteers at Bread For The Soul, Motley is committed to *helping people with AIDS* attain some spiritual comfort before death. (*Washington Post*, 15 April 1995)
26. Mr. Pomeroy doesn't like the idea of *helping people with AIDS*, said Brian Pardo, president of Life Partners Inc. He views it, I think, as a gay problem. (*New Orleans Times-Picayune*, 20 August 1992)

27. Family doctors who may feel ill-equipped to *treat* **people with HIV**, the AIDS virus, got help from the government Thursday. (*USA Today*, 21 January 1994)
28. "A lot of dentists still don't want to *treat* **people with HIV**," said Becky Harmon, an HIV case manager with The Partnership Project. (*Seattle Post-Intelligencer*, 23 November 1998)

**Figure 7-7: Instances of the Individual agent in concordance lines where *help* and *treat* collocate with PWA/H**

The next most frequent agent is categorised as Drugs/Treatment. This category is somewhat problematic in its unusually high use of the infinitive, indicating that the drugs are not always necessarily performing the action themselves, but being used as part of a larger process. However, as the treatments are foregrounded in these concordance lines over the practitioners, it is best to keep them segregated from the results above. In this category, *help* and *treat* are quite synonymous, though *help* explicitly refers to the potential of Drugs/Treatments to improve quality of life (lines 29 and 30), a feature which is not prominent in the *treat* concordance lines, which indicates that while *treat* is a medical process, *help* is this medical process combined with a more social remediation, particularly in the second half of the corpus, ranging from the late 1990s through the 2000s.

29. Activists are hoping for more good news in the near future, as combination drug therapies using protease inhibitors continue to show promise in *helping* **people with AIDS** live longer and healthier lives. (*San Francisco Chronicle*, 21 June 1997)
30. In a videotaped message to participants in a ceremony organized by the Elizabeth Glaser Pediatric AIDS Foundation, Clinton said the battle against AIDS has shown some signs of success, with powerful new treatments *helping* **people with AIDS** live longer and healthier lives. (*New Orleans Times-Picayune*, 2 December 1997)
31. By studying children born with compromised immune systems, researchers are learning how gene therapy may be used to *treat* **people with AIDS**, a California scientist said Tuesday. (*Denver Post*, 24 September 1997)
32. The drug --- also known as Viread and currently used to *treat* **people with HIV** --- is being assessed for eventual use in preventing HIV infection in people at high risk. (*Atlanta Journal-Constitution*, 17 March 2005)

**Figure 7-8: Instances of the Drugs/Treatments agent in concordance lines where *help* and *treat* collocate with PWA/H**

Like the Drugs/Treatment category, the Financial Assistance category was somewhat problematic to categorise. In many cases, *help* or *treat* appeared in the gerund form, meaning that no agent was necessarily attached. However, in cases of collocation where the *helping/treating* was related to financial measures or bursaries as opposed to organisations, governments, or individuals, this did seem the most appropriate. Also like the Drugs/Treatment category, the Financial Assistance category showed a real

divergence in concordance lines belonging to *help* versus those to *treat*. Financial Assistance was nearly always a positive force in *helping* PWA/H; grant awards (line 33) or fundraising attempts (line 34) put toward *helping* PWA/H are celebrated throughout the corpus. Conversely, the Financial Assistance category in the *treat* concordance set is much more negative; here, PWA/H are construed as dependent upon the system, and a threat to the national economy (lines 35 and 36). This is a revival of the *topos* of strain explored in all three Katrina chapters, wherein the ‘others’ are described in terms of their ‘neediness’, particularly financial. Despite not being coded within the I category, this is one aspect of the poverty/financial strain discourse surrounding others in Act of God reporting, whereby one social group is construed as a threat to the in-group members by being reported as drawing resources out of the system, without the ability to replenish them. As in line 35, the cost of others ultimately falls to the average in-group citizen, or the ‘nation’ at large (this is the “the nation’s medical bill”). While grants and private fundraisers are positively appraised as *helping* the needy, allocation of public funds to *treat* PWA/H is negatively appraised as a threat endorsed by the government.

33. The grants help people with AIDS who live on \$700 a month or less to pay rent, phone bills and other expenses. (*San Francisco Chronicle*, 3 May 1989)
34. Sir Elton has championed the fight, establishing the Elton John AIDS Foundation in 1992, which has spent more than \$35 million in 55 countries to help **people with AIDS** and prevent the spread of HIV. (*Washington Post*, 11 April 2002)
35. The nation’s medical bill for *treating people with AIDS* or HIV infection will reach \$5.8 billion this year and will almost double to \$10.4 billion in 1994, a federal study estimates. (*Seattle Post-Intelligencer*, 29 November 1991)
36. There is no question that *treating people with AIDS* is extremely costly, Dr. Feinberg said. (*New York Times*, 20 April 1992)

**Figure 7-9: Instances of the Financial Assistance agent in concordance lines where *help* and *treat* collocate with PWA/H**

The Government and its Legislation is another common agent in the *help* concordance lines (over 8% of the total). The late 1980s and early 1990s are dominated by discussion of the government’s need to *help* PWA/H, though [perceived inefficiency in] efforts to do so are met with negative reactions, leading to ‘chaos’ and demands from citizens (line 37), and efforts to *help* HIV being positively appraised and indicating good leadership (line 38). By the 2000s, it is no longer American parties and politicians, but foreign governments that are criticised for their inability or slow movement to begin *treating* PWAs (line 39), while legislation supporting *treatment* of PWA/H in America continues to be appreciated (40).

37. The [city] council's action prompted outrage from some members of the city's homosexual community, who turned the meeting into chaos demanding that the council "**help people with AIDS.**" (*Washington Post*, 22 December 1988)
38. Clinton's proposals would **help people with HIV**, the virus that causes AIDS, as well as people with kidney disease, spinal-cord injuries, multiple sclerosis, various types of mental illness and impairments caused by strokes and heart attacks. (*Seattle Post-Intelligencer*, 30 November 1998)
39. Last November, the [South African] government agreed to begin *treating* **people with AIDS**, but it's unclear when the program will begin. (*Boston Globe*, 2 July 2004)
40. [George W. Bush] also called for renewal of legislation known as the Ryan White Act designed to help *treat* **people with AIDS**. (*Boston Globe*, 1 February 2006)

**Figure 7-10: Instances of the Government/Legislature agent in concordance lines where *help* and *treat* collocate with PWA/H**

The final standard category here is Pets. With only one concordance line falling under this heading, I will not dedicate much time to its discussion. I only stop briefly to note that humans are not the only sentient beings to *help* PWA/H; a single concordance line (41, below) also indicates that pets are also able to give them emotional support. In this example, PWA/H are doubly passivated—a volunteer organisation **helps** pets who **help** *people with HIV*, while the agency of these people is passive, grammatically restricted to living (alone).

41. Slyly labelled "barkitecture with cattitude," the silliness is in its sixth year, benefiting Pets Are Wonderful Support, a volunteer organization that helps the pets who **help people with HIV** who live alone. (*San Francisco Chronicle*, 9 March 2001)

**Figure 7-11: Instances of the Pets agent in concordance lines where *help* and *treat* collocate with the PWA search terms**

Before moving away from discussion of *help* and *treat* and on to the other collocates in the USAS B category, it is important to note that, while I have compared these two collocates as near-synonyms, *treat* does occupy one semantic sense that *help* does not. Previous results showed the *help* had a more social aspect, despite *treat* being more frequently used with Individual agents. An additional, unique sense of *treat* does signpost a cross-over capability in this word; while *treat* is nearly always used in the medical, physical sense, an emotional, social sense is observed in over 10% of the concordance lines arising from this collocation. This sense is not used for individuals (doctors, practitioners, etc.) to discuss medical procedures or treatments, but can be found in calls to action for society to behave kindly to PWA/H, and *treat* them with positive emotional/social qualities, like respect (line 42), fairness and compassion (line 43). A much more detailed comparison between use of the physical/medical sense and

the emotional/social sense of *treat* will be provided in Chapter 8, where a higher frequency of concordance lines to analyse allows for more empirical description.

42. Bush has made occasional remarks about the need to *treat people with AIDS fairly and compassionately*. (*USA Today*, 15 June 1992)
43. "It's about *treating people with AIDS with respect*," said the Rev. Yvette Flunder of the City of Refuge Church in San Francisco. (*San Francisco Chronicle*, 9 June 2001)

Figure 7-12: Instances of the 'Emotional Sense' of the PWA/H collocate *treat*

The lemma collocate *serve* very strongly prefers the Organisation/Business agent (46 concordance lines), followed by Government/Legislation/ Programme (18) and Medical Treatment/Centre (16). Therefore, one of the dominant roles of corporate and charitable actors in association with PWA/H is one of service or supply of goods. The most frequent collective agents of *serve* indicated by collocates of the same spans and cut-offs in the newspaper section of COCA have to do with the food service industry (*eatery* MI 4.81, *café* MI 4.02, *waiter* MI 3.59, *waiters* MI 3.96, *restaurants* MI 3.20) and military workers (*gays* MI 4.07). However, collocates relating to the objects or ranges of *servicing* indicate another pattern in semantic use: various individuals also *serve*, usually either government documents such as *warrants* (MI 5.17) or *subpoenas* (MI 5.89), or periods of time, as in *stints* (MI 4.94) and *sentences* (MI 4.33). Use of *serve* is therefore quite variable. Looking at the semantic sense of food service, *serve* could also construe PWA/H as clients in a service transactions where an Organisation/Business agent (including charitable foundations) bestows goods and services upon them in a financial exchange. In linking this to the type of *serve* found in 'in the military', this might mean a [patriotic] civic duty being undertaken—indeed, the Government/Legislation/Program agent is the second-most frequent to use *serve*, signifying a link between these discourses.

*Assist*, the second most frequent verb lemma in this USAS semantic group, shows an even stronger preference for the Organisation/Business agent, being activated by this group in 36 of its 63 concordance lines. This verb has the semantic capacity for a greater level of interpersonal connection; the transaction in *serve* is very one-sided with the flow of goods going from actor A downward to actor B, but the definition of *assist* allows for the possibility of actor B then becoming empowered to become activated with this assistance. Surprisingly, in only seven of these cases are human agents *assisting* PWA/H; in an additional four, a call to action is delivered (to the individual reader). In an overwhelming majority of the concordance lines attached to *assist*, PWA/H are being



*assisted* by organisations, businesses, and charitable foundations, rather than by individuals.

This leaves us with the final three collocates passivizing *PWA/H*, all with relatively low frequency of occurrence. Government/Legislation/Program is the sole agent in each of the 45 concordance lines returned for the collocate *protect*. Of these, 23 explicitly state that laws and legislation are designed to *protect PWA/H* from discrimination. The Government/Legislation/Program agent also features in the majority of the *enable* concordance lines (six out of 10); most of which seem concerned with *enabling PWA/H* to qualify for government aid: Medicaid (4), “federally financed housing for the disabled” (1). One final concordance line helps *PWA/H* to carry on contributing to the economy—which was found to be critical in the earlier Katrina chapters—by *enabling* them to “continue working”. Finally, in five out of a total of nine cases, Individuals *counsel PWA/H*, making this one of the stronger interpersonal verb forms to be analysed. However, given the extremely low frequency of this result, not much will be made of its statistically insignificant preferences.

The last collocate in this category has been kept isolated from collocates showing *PWA/H* agency and collocates indicating lack thereof. *Deserve* is unique in that *PWA/H* directly precede the verb, but *deserve* does not directly implicate either agency or power. The value that *PWA/H* are construed as *deserving* or not *deserving* is judged and designated by a third party; the nature of *deserving* is that one is not achieving or obtaining some goal or item on their own, but being designated it in a possible world. Each of the 12 concordance lines arising from this collocation appear chronologically below; note that lines 46 and 50 appeared in two different publications (and therefore should be counted twice), in the case of 46, months apart. In line 47, it is not *PWA/H* themselves who are described as *deserving*, but “portraits” of them; this line will therefore be disregarded.

44. “A lot of people feel that **people with AIDS** *deserve it*,” she says. “But others, especially parents, will have compassion.” (*Boston Globe*, 21 September 1987)
45. In contrast, people who express the opinion that there should be compulsory screening for infection with the AIDS virus and quarantine of those who test positive also support workplace testing for other sexually transmitted diseases and drug use, and say, in the survey, that “**people with AIDS** *deserve it and are responsible for it*.” (*Newsday*, 2 August 1988)

46. During a course on disease, Franke said, the discussion inevitably came to AIDS. "The prevailing attitude in class threw me for a loop: **People with AIDS deserve to die**. Others would say, 'I want nothing to do with them.' There was hardly any show of compassion!" And yet, Terkel quotes Franke, "when I showed them a film on AIDS and the dying, they were disturbed. Some in the class cried openly. (*Boston Globe*, 9 October 1988 / *St. Louis Post-Dispatch*, 14 December 1988)
47. Though his penetrating and often astonishing portraits of people with AIDS deserve to be seen by as wide a public as possible, it's the in-your-face immediacy of Kearns's work that sets it apart; experiencing this difficult but ultimately uplifting material with a theater full of laughing, tear-stained folk seems an essential part of the proceedings. (*Washington Post*, 20 September 1991)
48. Yet at the expense of you, your family and friends, members of Congress and the press have bowed to extreme political pressure from the radical homosexual community so that homosexuals may obtain preferential treatment in medical research dollars. Don't get me wrong. **People with AIDS deserve our compassion just as much as others with life-destroying diseases**. Furthermore, I'm sad that a super athlete like Earvin "Magic" Johnson contracted the virus, but let's face reality: AIDS doesn't deserve favored-disease status. (*St. Louis Post-Dispatch*, 12 December 1991)
49. The final message: **people with AIDS deserve our compassion and respect**. (*New York Times*, 25 March 1992)
50. "The world is still handicapped in the battle against AIDS," [Dr. Jonathan Mann of the Harvard School of Public Health] concluded, "by debating whether it is OK to distribute condoms, exchange needles, give sex education; whether **people with AIDS deserve care that they obviously deserve**." (*New York Times*, 4 June 1992 / *Seattle Post-Intelligencer*, 4 June 1992)
51. "AIDS is not divine punishment," the text [of an international declaration against AIDS-based discrimination] reads. "**People with AIDS deserve the right to retain their jobs and homes and to travel freely**. The fight against AIDS must never be a fight against people who have AIDS." (*San Francisco Chronicle*, 18 May 1996)
52. I see nothing wrong with it considering most **people with AIDS deserve it**. How does a person come to a conclusion like this? Well, I do not use drugs intravenously, do not have unprotected sex and I am not a homosexual. I am capable of looking at my surroundings and realizing what is harmful and what is not. (*St. Louis Post-Dispatch*, 11 July 1998)
53. [Princess Diana] showed the world that **people with AIDS deserve not isolation, but compassion and kindness**. (*Chicago Sun-Times*, 31 August 2007)

Figure 7-13: All concordance lines arising from the collocation of *deserve* with *PWA/H*

In the 11 remaining lines, *PWA/H* are construed as *deserving* AIDS three times, and "to die" twice. In contrast, they are described as *deserving* compassion (3) (with respect (1) and kindness (1)), care (2) and "the right to retain their jobs and homes and to travel freely" (1, concordance line 51). It appears, superficially, that *PWA/H* are said to *deserve* negative things (the disease itself, and death) about as often as positive social responses

(compassion, care) and rights. However, four of the extended concordance lines show a compound meaning, which reiterates the case for deeper qualitative analysis.

First, I describe some immediate concordance lines that appear negative but which are more positive in extension. In line 44, while one speaker (an AIDS patient, it is revealed much earlier in the text) concedes that “A lot of people will feel that people with AIDS deserve it”, she states that others will have compassion. Line 45 is reporting results of a survey, and cannot be tied to journalistic stance, and the distance here is evidenced by the use of quotation marks. Line 46 shows a change of attitude, from a class believing that “people with AIDS deserve to die” to crying openly after seeing a video on AIDS and dying. Only one concordance line (48) has a positive prosody in close proximity to the node, with more negative prosodies appearing in extended context. While the writer allows that *PWA/H deserve* our compassion (“just as much as others with life-destroying diseases”), it is posited previously that it is at the expense of “you, your family and friends” that members of Congress “bowed to extreme political pressure from the radical homosexual community”, and later that “AIDS doesn’t deserve favored-disease status”. Here it seems that the definition of compassion does not extend to research funding, and does not include ‘the radical homosexual community’.

For the most part, however, the trend in concordance lines of *deserve* moves generally from more negative to more positive over time (with line 52 from 1998 being the anomaly). Most lines are uncomplicated in their prosody or stance, with repetition and rephrasing being features of both positive prosody (e.g. line 50: “**whether people with AIDS deserve care that they obviously deserve**”) and negative prosody (e.g. line 45: “**people with AIDS *deserve* it and are responsible for it**”).

In summary, though USAS category S was extremely populated with collocates, I came up against a familiar struggle in finding discussion focussing on *PWA/H* themselves, in contrast to the much more frequent appraisal and construal of social actors in their immediate context. For instance, focus was on the Agencies/Organisations/Centres *helping* them; named agents *treating PWA/H* were positively appraised, whereas doctors refusing treatment were negatively judged. As is symptomatic of both the risk society mindset (Beck, 1992) and the stigma of charity (Fothergill, 2003), it is the government who is most often indicated as ‘responsible’ for legislation and care of *PWA/H*, though this is at odds with resistance of the financial implications of their care for taxpayers. Instances where *PWA/H* are activated still implicate heavy agency of an additional party

(with more social capital) – they are *enabled* to do tasks, or deemed to *deserve* certain consequences or affordances due to their condition.

## **7.4 Summary**

As in previous chapters, this section is divided into two subsections: a summary of findings on the construal of *people with AIDS/HIV* as explored above, and an evaluation of the methods employed in this chapter.

### **7.4.1 Of findings**

*People with AIDS/HIV (PWA/H)* do not collocate with particularly novel broad semantic categories – proportionally co-occurring most often with USAS A (general and abstract terms), I (money and commerce in industry), N (numbers and measurement), and S (social actions, states and processes) categories. The lack of a proportionally significant representation of B category collocates is noteworthy, and indicates that *PWA/H* are construed more often in contexts of social processes than medical ones. USAS A category collocates associate *PWA/H* with segregation and isolation, though those concepts are in turn associated with extremism. The stigma of charity recurs here, most notably in category I collocates, where *PWA/H* are construed as *poor* and *needy*, and subject to debate about worthiness of public benefits, as do patterns of enumeration and the attendant threat *topos* in category N. Concordance lines arising from category S collocates were arguably the most interesting in that they were internally conflicted; it seemed widely agreed that something should be done to *help PWA/H*, though who exactly should be involved in helping (and in paying for this help before it became a strain) was hotly debated.

However, not terribly much can be said about what *PWA/H* themselves do, feel or say. Despite carrying out a search specifically for “people with (AIDS|A.I.D.S.|HIV|H.I.V.|GRID|G.R.I.D.)”, it was very difficult (nigh on impossible) to find instances in the corpus where this target group was the agent in any process types. As typified in the analysis of concordance lines for *isolate/quarantine* and *help/treat*, *PWA/H* are construed as the passivized recipients of the actions of others, and it is these others who are actually the subjects of evaluation and discussion in the newspapers in the corpus. Searching for traces of agency of the *PWA/H* themselves is a challenge that I take forward into Chapters 8 and 9.

### **7.4.2 Of method**

In response to shortcomings with the companion study for this section (Katrina chapter 4), I opted to downsample on the basis of broad USAS semantic categories, taking into account each collocate of every part of speech falling under the most highly represented as a proportion of all collocates. This has been a vast improvement, for several reasons.

Firstly, for the sake of ease of use, dependence on (erratic) part of speech tagging has been removed for the most important first downsampling pass. While the USAS system does have an inconsistency margin of about 15%, this was easily corrected in the manual disambiguation stage where automatically assigned collocates from positions other than the first were usually just shifted upwards. The shades of meaning in this system make it more flexible, particularly in contrast to the CLAWS tagger, which distinguishes parts of speech on a four-way split with little room for disambiguation not involving manual analysis of every single associated concordance line.

Weight has also been placed on the semantic sense of collocates rather than over reliance on part of speech. This allows domains evenly represented (but perhaps with only middling frequency) over multiple parts of speech to be highlighted as salient with equal, if not greater, importance than categories spiking in one part of speech category and tapering off in all others.

Finally, but quite importantly given the limitations of this medium, no unnecessary space has been wasted in the needless duplicated description of a word collocating with the naming strategy in two different parts of speech.

This method, based upon USAS broad semantic categorisation, will be carried forward into the remaining chapters and further refined. In Chapter 8 (as in Chapter 5), I explore the next four most frequent 'human' naming strategies in the HIV/AIDS corpus, comparing semantic preference of their collocates and the associated prosodies. Chapter 9 is a diachronic study (as in Chapter 6) of a single identity, in which the use of USAS semantic categories is critical in identifying salient patterns in nearly 30 years data exhibiting extremely varied discourses.

## ***Chapter 8: Semantic preference of frequent naming strategies for people with HIV/AIDS***

### **8.1 Introduction**

This chapter carries on from the previous one in analysing the discourses around people with AIDS/HIV (PWAs). As before, scope is narrowed as we proceed through the analysis chapters of a given corpus. Following on from the findings regarding the most frequent naming strategy in the HIV/AIDS corpus—*people*—I now explore a more complete set of four naming strategies of varying lesser frequencies, comparing and contrasting their collocates, semantic preferences, and semantic prosodies for heightened insight into both their individual meanings and the wider discourses of AIDS.

In tandem with HIV/AIDS Chapter 8, this chapter addresses research question 3, with emphasis on the items in boldface below:

*What are the most frequent naming strategies employed in Hurricane Katrina and AIDS/HIV reporting, and what can corpus-based discourse analysis of associated **attributes, transitivity, and cognitive metaphors** tell us about the in- and/or out-group qualities attached to the people that the naming strategies represent?*

In section 8.2 I detail the method utilized in this chapter, and in section 8.3, I specifically describe the process of pinpointing particular naming strategies for scrutiny. In response to methodological shortcomings arising in Chapter 5, several practical adjustments have been made. These are signposted throughout section 8.2 and 8.3 under the boldface titles, “Adjustment in methods”.

In keeping with the structure established in Chapter 7, the analysis section (8.4) is organized by USAS broad semantic category of collocates. Two USAS categories make their debut in this chapter: Category B: the body and the individual (section 8.4.2), and Category X: psychological actions, states and processes (section 8.4.6).

The final section contains a summary of results (8.5.1), a critique of the method (8.5.2), and a reflection upon the distinctions between this chapter and its sister, Chapter 5 (8.5.3).

## 8.2 Method: Targeting search terms

Following the method established in Chapter 5, frequent naming strategies were explored using both the two genitive forms surrounding the 'Act of God'. The first formula, the *of* genitive—"x of AIDS/A.I.D.S./HIV/H.I.V./GRID/G.R.I.D." (where x is a plural common noun)—is an exact replica of the syntax used in Chapter 5.

**Adjustment in methods:** A point of difference with the earlier examination of Katrina came in the form of the -s genitive. Whereas hurricanes and tropical storms are assigned human names under international convention (thereby lending themselves to a standardised use of the -s genitive, e.g. 'Katrina's victims'), diseases do not follow this pattern, particularly in the case of acronymic labels such as A.I.D.S. and H.I.V. Therefore, investigation of naming strategies could not be exactly mirrored between the two corpora. Instead, I have searched for compound noun phrases "**AIDS/A.I.D.S./HIV/H.I.V./GRID/G.R.I.D. x**", where the disease pre-modifies the (plural common noun) naming strategy, x.

As before, the most frequent 'human' nouns were then selected from the frequency lists resulting from these search phrases. I have restricted the search span of x to plural common nouns composing more than 0.5% of the search results in this position to target the most common combinations. This number is somewhat arbitrary, but does produce the top 28-31 results for each syntactical search.

In Tables 8-1 and 8-2, all 'human' results of each search type are presented. The pre-modified search form was more productive than the *of*-genitive form in terms of exposing naming strategies. Naming strategies referring specifically to PWAs are in boldface. Of the top 31 plural common nouns in this position, five refer to PWAs in the "**AIDS/A.I.D.S./HIV/H.I.V./GRID/G.R.I.D. x**" form (e.g. *AIDS victims*, *HIV patients*), while only one of the top 28 results in the **x of AIDS/A.I.D.S./HIV/H.I.V./GRID/G.R.I.D.** search string (*victims of AIDS*) was a group of PWAs. Human results for both searches are shown below. The resulting set of four—the lemma forms of *patient*, *victim*, *sufferer*, and *carrier*—form the basis of analysis in this chapter.

	Search result	Number of occurrences	Percentage of search	Human actor?	PWA/H?
1	AIDS patients	18,759	25.0%	✓	✓
3	<b>AIDS victims</b>	5,008	6.7%	✓	✓
5	AIDS activists	2,929	3.9%	✓	
9	AIDS researchers	1,331	1.8%	✓	
10	<b>AIDS sufferers</b>	1,322	1.8%	✓	✓
13	AIDS experts	1,066	1.4%	✓	
15	<b>HIV patients</b>	879	1.2%	✓	✓
20	AIDS orphans	649	0.86%	✓	
25	AIDS carriers	453	0.6%	✓	✓
28	AIDS specialists	405	0.5%	✓	

Table 8-1: Frequency of 'human' plural common nouns in the AIDS/A.I.D.S./HIV/H.I.V. x form

	Search result	Number of occurrences	Percentage of search	Human actor	Person with AIDS
3	<b>victims of AIDS</b>	616	5.2%	✓	✓
21	partners of AIDS	78	0.7%	✓	

Table 8-2: Frequency of 'human' plural common nouns in the x of AIDS/A.I.D.S./HIV/H.I.V. form

As indicated in both Table 8-1 and 8-2, several other 'human' plural common nouns appeared in frequency lists. However, as this research is focussed on the representation of PWAs, ancillary roles, e.g. *AIDS activists* and *partners of AIDS*, where the infection status of the named person is ambiguous, have been excluded. Further research into the representation of those affected by AIDS/HIV (*orphans, partners*) or working in the medical or social activism fields (*activists, researchers, experts, specialists*) would likely be quite interesting, though this is not included in the scope of the current project.

**Adjustment in methods:** Once the most frequent 'human' plural common nouns were identified in the Katrina corpus with the aid of genitive syntactic structures, these portions of the search terms were stripped out, allowing for a greater variety of supplementary identities to be thrown up by the corpus (e.g. *black victims* as opposed to only *Katrina's victims*). This was viable given the extremely limited diachronic scope of the Katrina corpus, but is no longer a possibility when taking into account the much longer diachrony and diversity of texts in the HIV/AIDS corpus. Therefore, though I refer to naming strategies in their short forms (e.g. *patient*), each search term is completely embedded in the surrounding phraseology (e.g. *AIDS patient*) to ensure that even in such a complex corpus, naming strategies are as targeted and descriptions are as accurate as possible.

Once these naming strategies (*patients, victims, sufferers, and carriers*) were identified, it became necessary to compress results, both in terms of whether each naming strategy



preceded or proceeded the disease name, and whether it occurred in phraseology with *AIDS*, *A.I.D.S.*, *HIV* or *H.I.V.* Searches adhered to the following syntax in CQPweb, where *x* is one of the naming strategies above:

$$(((AIDS|HIV|H.I.V.|A.I.D.S.|GRID|G.R.I.D.) \{x\}_N^*)|(\{x\}_N^* \text{ of} \\ (AIDS|HIV|H.I.V.|A.I.D.S.|GRID|G.R.I.D.)))$$

Frequency figures derived from this set of search criteria can be found below.

### **8.3 Exploring naming strategies**

#### **8.3.1 Incidence of naming strategies**

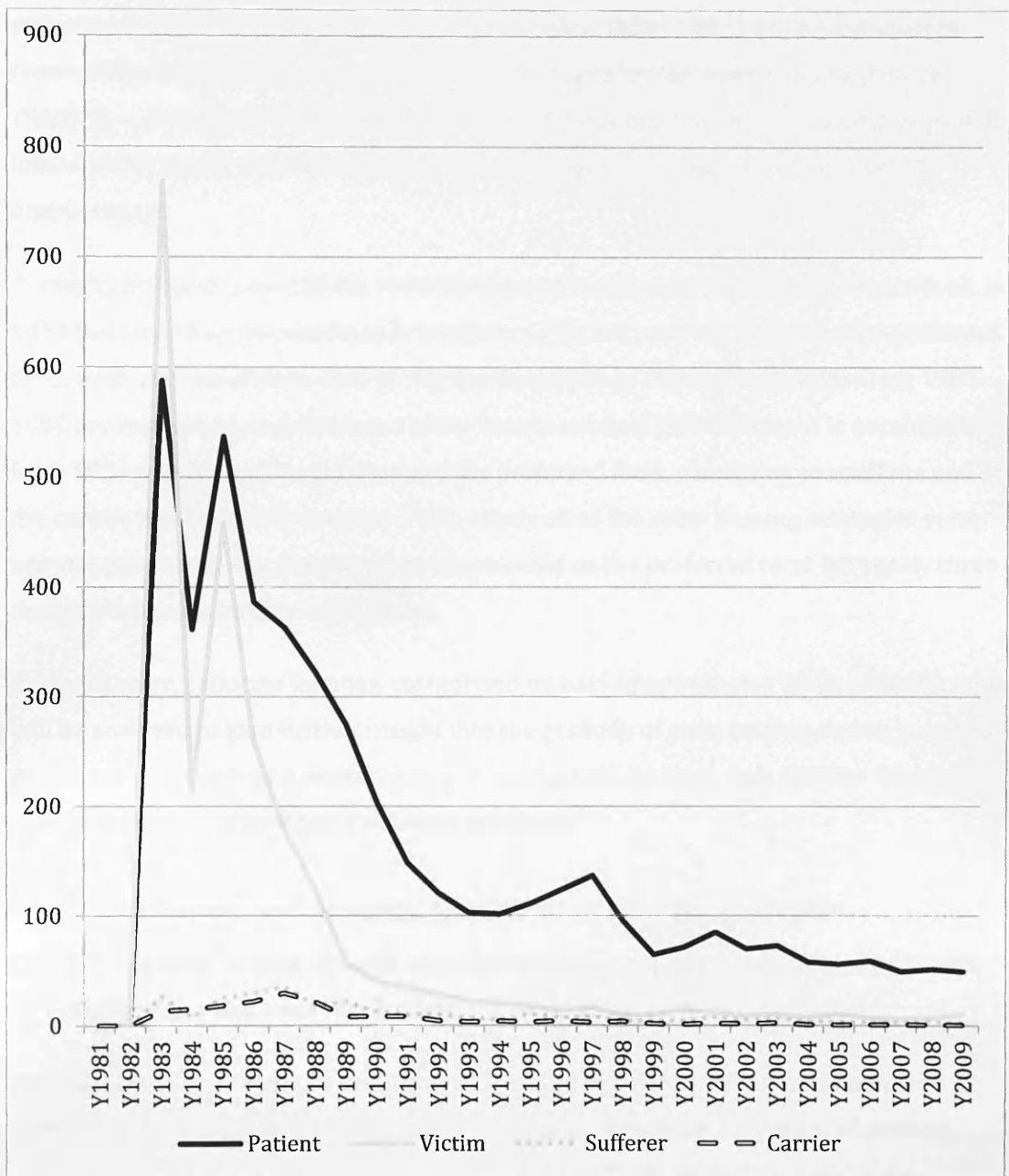
Overall, *patient(s)* is the most popular naming strategy of the four, appearing in the phraseology above 140.26 times per million words in the corpus. Significantly less frequent: *survivor(s)* at 42.09/million, *sufferer(s)* at 9.76/million, and *carrier(s)* at 5.83/million. Each of these naming strategies is most frequent (in terms of raw frequencies, proportional frequencies and dispersion) in the 1980s, followed by the 1990s, and then the 2000s. This indicates that naming strategies became less simplistic as time went on, and that identities of PWAs were not as tightly connected with the disease itself (as conveyed in genitive and compound nominal grammatical structure). The shifting construal of one group of PWAs over time will be explored further in Chapter 9.

For now, it is interesting to note that *carrier(s)*, always the most infrequent of the naming strategies being analysed in this chapter, became drastically less frequent between the 1980s and 1990s, and was nearly extinct by the 2000s. In this instance, the frequencies derived from the corpus are able to quantify the rate of decline in naming strategies as they become obsolete. Here, it is clear that early discussion on the disease was characterised by a misunderstanding of the relationship between HIV and AIDS; it was believed that one could be a *carrier* (and distributor) of the virus without contracting the syndrome. The emergence of this pattern is encouraging (at least as it relates to the validity of this study), as it suggests that a good proportion of early, biologically unsophisticated discourse has been digitised and archived.

		1980s	1990s	2000s	Overall
<b>Hits in category</b>	patient(s)	11,115	8,474	3,013	22,602
	victim(s)	4,851	1,510	422	6,783
	sufferer(s)	802	557	214	1,573
	carrier(s)	617	277	45	939
	<b>Words in decade:</b>	<b>35,286,877</b>	<b>77,681,146</b>	<b>48,176,901</b>	<b>161,144,924</b>
<b>Dispersion</b>	patient(s)	5,563	6,053	2,269	13,885
	victim(s)	3,442	1,365	399	5,206
	sufferer(s)	652	504	197	1,353
	carrier(s)	476	223	41	740
	<b>Texts in decade:</b>	<b>41,548</b>	<b>81,587</b>	<b>43,441</b>	<b>166,576</b>
<b>Freq./ million in category</b>	patient(s)	314.9	109.09	62.54	140.24
	victim(s)	137.47	19.44	8.76	42.09
	sufferer(s)	22.73	7.17	4.44	9.76
	carrier(s)	17.49	3.57	0.93	5.83

**Table 8-3: Frequency of naming strategies by decade—expressed as raw hits, dispersion over total texts, and frequency per million words in the category**

While it is interesting (and illuminating) to consider the broad frequencies of naming strategies by decade, the granular nature of corpus mark-up in combination with the powerful indexing of CQPweb allows for a much deeper look at diachronic patterns. Reviewing relative frequencies per year (see Figure 8-1 below) exposes a noteworthy irregularity in usage of naming strategies.



**Figure 8-1: Frequency of naming strategies by year—expressed as frequency/million words in the category**

At this point it is important to note that no matches for any of the four naming strategies were found in the 1981 and 1982 texts. During this time, pre-dating the establishment of nomination for the disease itself, PWA/H were labelled with a large variety of generalised naming strategies, though the most frequent are *patient(s)*, *victim(s)*, and *case(s)*, not pre- or post-modified with any of the disease markers used to eliminate false positives from the corpus. Due to the low frequency of texts in 1981-1982, and the strong overlap in naming strategies that do and do not show pre- and post-modification,

statements made about the naming strategies below (after 1983) will be assumed to reasonably extrapolate to the same naming strategies for the emergent era (before 1983). As a more detailed diachronic analysis of the representation of social groups will follow in Chapter 9, the focus of this chapter is frequent naming strategies over the entire corpus.

Though *patient(s)* is overall the most frequent of the four naming strategies analysed, in 1983 (when all four nominations first appear in the corpus), *victim(s)* was the preferred form. Both of these show a drop in relative frequency in 1984, though texts from 1983-1985 are infrequent, and incidence alone here is not a reliable finding. It is notable that from 1984 onward, *patient(s)* becomes the preferred form, remaining so until the end of the corpus timeframe. By the year 2009, nearly all of the other naming strategies were nearing extinction. The development of *patient(s)* as the preferred term for nearly three decades warrants further exploration.

In this chapter, collocate lemmas, categorised by part-of-speech and USAS semantic tags, will be analysed to gain further insight into the prosody of each naming strategy, and to draw conclusions about how the choice of one naming strategy over another impacts AIDS/HIV discourse around PWAs more generally.

### **8.3.2 Collocation and semantic preference of naming strategies**

Collocates of each naming strategy were calculated on a span of +/-3, with significance thresholds of  $LL \geq 10.83$  and  $MI \geq 3$ , mirroring those set in all previous chapters.

**Adjustment in methods:** The minimum threshold of collocation has been adjusted to reflect the change in corpus and the difference in standardized frequency of naming strategies between the two corpora. The naming strategies in Katrina were more generalized, with searches performed for just *resident(s)*, *victim(s)*, *evacuee(s)*, and *survivor(s)*. These returned 640.3/million, 334.4/million, 300.8/million, and 91.4/million results, respectively. By contrast, the naming strategies in the HIV/AIDS corpus are much less generalized, appearing in search strings that specify the disease (e.g. *AIDS patients*). Though the AIDS/HIV corpus, with over 161 million words, is 439% larger than the Katrina corpus, which has fewer than 37 million words, the change in search strings has resulted in dramatically lower frequencies. The most popular naming strategy (*patient*) appears 140.3 times per million words, while the least frequent (*carrier*) only occurs 5.8 times per million. As a result, the minimum frequency of

collocation has been lowered to 10 to allow for a comparable variety of relationships to be thrown up by the corpus.

Other than lowering the minimum frequency of collocation, all other settings and methods for calculating these relationships have been maintained. Proportions of resulting collocations are shown in Table 8-4 below. The highest proportion preference for each naming strategy is in boldface; preferences of semantic categories are in italics. Those semantic categories not containing collocates for any naming strategy are shaded out. The names of semantic categories represented by >10% of the collocates of any given naming strategy are in all capitals—the contents of six categories will be analysed below.

As in previous chapters, some USAS broad categories contain a large number of collocates that are not related to one another to the same degree. Therefore, when sets of more closely semantically related collocates present themselves within a broad semantic USAS category, subcategories are created to ease discussion. These appear under subheadings within a headed USAS broad category section.

#### **8.4 Collocational profiles of naming strategies**

In this section, I present analysis of the items categorized under the USAS broad semantic categories representing the highest proportion of collocates overall. The parameter for inclusion as a 'salient' domain is representation of at least 10% of collocates for at least one naming strategy. Full collocation tables for each naming strategy can be found in the appendices: more detailed collocation information for *patient* can be found in Appendix I, for *victim* in Appendix J, for *sufferer* in Appendix K, and for *carrier* in Appendix L.

Categories are covered alphabetically: collocates from B (the body and the individual) are analysed in section 8.4.1; G (government and public) in section 8.4.2; H (architecture, housing and the home) in 8.4.3; S (social actions, states and processes) in 8.4.4; and finally, X (psychological actions, states and processes) in 8.4.5.

Please note that in tables of this section, collocates shared between two or more naming strategies will appear in italics for emphasis. Those that have been disregarded for reasons described in the text have been struck out. For USAS categories A, B, H, and S, further subcategories have been created for ease of description across groups of like collocates. Analysis of these appear under subheadings with tables further contextualizing the proportions that a particular number of collocates represent of a USAS broad category and of all collocates of a given naming strategy overall.

##### **8.4.1 Category A: general and abstract terms**

Category A contains the third highest proportion of collocates of the four naming strategies being explored in this chapter, averaging over 10%. As illustrated below in Table 8-5, it is *patient(s)* and *carrier(s)* that show the highest semantic preference for this domain, though over 5% of collocates of both *victim(s)* and *sufferer(s)* also fall into this category.

Naming strategy	Collocates from USAS category A: GENERAL AND ABSTRACT TERMS	% total collocates
patient(s)	<b>quarantining</b> [LL=309.62, MI=7.369], <b>quarantine</b> [LL=298.101, MI=5.09], <b>exclusively</b> [LL=119.817, MI=4.342], <b>isolate</b> [LL=107.477, MI=3.603], <b>shun</b> [LL=90.537, MI=4.194], <b>confidentiality</b> [LL=80.885, MI=3.385], <b>isolating</b> [LL=72.042, MI=6.111], <b>severely</b> [LL=59.414, MI=3.343], <b>commonly</b> [LL=56.212, MI=3.332], <b>solely</b> [LL=29.876, MI=3.077], <b>classify</b> [LL=29.238, MI=3.418]	12.5%
victim(s)	<b>quarantine</b> [LL=473.827, MI=6.923], <b>quarantining</b> [LL=372.58, MI=9.025], <b>plight</b> [LL=70.615, MI=4.793]	6.3%
sufferer(s)	<b>plight</b> [LL=85.501, MI=6.576]	6.7%
carrier(s)	<b>quarantine</b> [LL=174.793, MI=8.07], <b>potential</b> [LL=95.516, MI=5.025]	14.3%

Table 8-5: Collocates for each naming strategy from USAS category A

These collocates will be discussed under various subheadings: Exclusion (section 8.3.1.1), Plight (8.3.1.2), Confidentiality (8.3.1.3), Classification (8.3.1.4), and Exclusivizers/Particularizers/Maximizers/Comparisons (8.3.1.5).

#### 8.4.1.1 Exclusion

The theme of exclusion or social sequestering explored in Chapter 7 recurs here, featuring frequently in collocates from category A. Some form of the lemma *quarantine* collocates with *patient(s)*, *victim(s)* and *carrier(s)*, with *patient(s)* additionally collocating with *isolate*. While these collocations also appear initially to indicate a discourse of exclusion, close inspection reaffirms that collocates such as *quarantining*, *quarantine*, *isolate*, and *isolating* appear in concordance lines where proposals of such actions are negatively appraised, associated with the extremely far right, or called outright irrational, e.g. “hysterical calls for testing or *quarantining AIDS patients*” (*San Francisco Chronicle*, 27 July 1987). Please refer back to section 7.3.1 for discussion of the implications of co-occurrence of PWA naming strategies and *quarantine* and *isolate*.

Social actor	EXCLUSION A category collocates	% A category collocates	% overall collocates
<i>patient(s)</i>	<i>quarantining, quarantine, isolate, shun, isolating</i>	45.5%	5.2%
<i>victim(s)</i>	<i>quarantine, quarantining</i>	66.7%	4.2%
<i>sufferer(s)</i>	-	0.0%	0.0%
<i>carrier(s)</i>	<i>quarantine</i>	50.0%	7.1%

Table 8-6: USAS Category A subcategory ‘Exclusion’ collocates by naming strategy

I focus instead on the novel collocate to arise in this batch of PWA naming strategies. For instance, *patient(s)* favours the exclusion pattern above the others, leading to unique

collocations with both *isolate* and *shun*. While *quarantine* and *isolate* additionally occupy the semantic space of medical treatment and disease, *shun* is emotionally charged and more closely related to sociological interaction with communities. Upon examining the concordance lines, of 23 cases where *patient(s)* collocates with *shun*, *patient(s)* are construed 15 times as being *shunned* not by members of their immediate social networks, but by medical workers (as in line 1 below). This is largely a pattern of the 1980s and early-to-mid-1990s, though there is some disagreement about when this so-called *shunning* stopped. See, for instance, concordance line 2 below (from 1991) where it was stated that half of the nation's physicians would *shun* AIDS patients if they could. In concordance line 3 (from 1994), it's claimed that nearly a decade has passed since this *shunning* has occurred. However, by 2002 (concordance line 4), this *shunning* becomes the issue of other, perhaps 'exotic' countries like India. This pattern of 'geographically distancing' problematic behaviour was commented upon in sections 6.3.2.2 and 7.3.1, will be revisited in Chapter 10.1.3.

1. The reminder reflected a growing number of reports of doctors in the United States, Britain and other countries *shunning* **AIDS patients** out of fear and ignorance. (*New York Times*, 15 July 1986)
2. Half the nation's front-line physicians would *shun* AIDS patients if they could, a survey published yesterday has found. (*Boston Globe*, 28 November 1991)
3. Less than a decade after many **AIDS patients** were *shunned* by the medical establishment, many in New York State are enrolled in hospital programs that offer what experts say is a model for the comprehensive care that all chronically ill patients should receive. (*New York Times*, 22 February 1994)
4. In India, HIV-prevention workers have been harassed, and impoverished **AIDS patients** have been *shunned*, while the disease spreads unchecked into the general population. (*San Francisco Chronicle*, 17 November 2002)

**Figure 8-2: Sample concordance lines showing the collocation between *shun* and *AIDS patients* where doctors/medical programs are the agents**

In four cases, *shun* appears in the passive form with the agent concealed (see, for instance, lines 5 and 6 below), indicating factuality of this state. In one additional line (7), it is "us" who is told not to do the *shunning*, though the parameters for inclusion in this group are not clear, and appear to cover all non-PWAs. Only one concordance line states that a person's family was doing the *shunning*, and not for having AIDS, but for being bitten by someone who did (see concordance line 8 below). Twice it is PWAs who *shun*, though the objects of the action are not human—they are construed as *shunning* treatment options (as in line 9).



5. But without an AIDS house, where do **AIDS patients** who are shunned and too sick to work go? (*Chicago Sun-Times*, 31 January 1988)
6. It is a fine line they walk between the gravity of their situation and the desire to hope, their feelings of being discriminated against and being cared for. **AIDS patients are shunned**. (*St. Louis Post-Dispatch*, 12 February 1989)
7. You can argue, and many people do, that it's in the public interest to portray HIV as a virus that can strike anyone, any time. So that none of us feels safe, so that none of us shuns AIDS patients and blames their terrible illness on them. (*The Wall Street Journal*, 7 January 1998)
8. Hospital police officer Lofton Johnson, 58, testified in \$3 million suit against hospital that family has *shunned* him since **AIDS patient** bit him. Johnson's wife says she's divorcing him because of AIDS threat. (*USA Today*, 12 December 1989)
9. Many **H.I.V. patients shun drugs** and use alternative therapies like herbal compounds, vitamins and spiritual approaches. (*New York Times*, 13 March 1994)

**Figure 8-3: Sample concordance lines showing the collocation between *shun* and *AIDS patients* with agents other than doctors/medical establishments**

The use of a negative value-laden verb such as *shun* in conjunction with medical professionals contributes to a negative prosody surrounding accessibility of treatment through those who were 'meant' to retain open minds and professional demeanours, despite the state of public reaction (which was largely negative in the 1980s). This blame on establishments, while obscuring individual citizen fault, is a feature of risk society discourse appearing frequently in the AIDS/HIV corpus.

#### 8.4.1.2 Plight

*Victim(s)* and *sufferer(s)* also collocate with *plight*, which is absent from the longer list of USAS A Category collocates of *patient(s)*.

Social actor	PLIGHT A category collocates	% A category collocates	% overall collocates
<i>patient(s)</i>	-	0.0%	0.0%
<i>victim(s)</i>	<i>plight</i>	66.7%	2.1%
<i>sufferer(s)</i>	<i>plight</i>	100.0%	6.7%
<i>carrier(s)</i>	-	0.0%	0.0%

**Table 8-7: USAS Category A subcategory 'Plight' collocates by naming strategy**

Two conversations surrounding the *plight of victim(s)* arise in the corpus. Of the 15 concordance lines, in seven cases the *plight of victim(s)* is publicized/underscored/dramatized/ put into perspective/raised into public awareness (see concordance lines 10-13 and 17 below), even if this is in a brash or noisy manner (e.g. line 11). An additional seven times, the *plight of victim(s)* is the object of emotional/mental processes, such as sympathizing/considering/ producing

empathy/having more on his mind (see concordance lines 12-13 below). As further indication of risk society rhetoric, lines 14-16 below indicate blame toward institutions: it is the church and the government who are viewed as *uncaring* and *unsympathetic*. An interesting parallel in phraseology also arises in concordance line 14, where the journalist refutes an implication that the pope “does not care about the plight of AIDS victims”. This phrase, and the sentiment in line 16, is reminiscent of allegations in the Katrina corpus that George Bush (or the entire government) did not care about certain—namely black—people. Post-dating Kanye West’s statement by two and a half years, this reaction—questioning the immeasurable level of sentiment of public figures regarding social groups—seems to have entered the cultural lexicon as an aspect of risk society discourse.

10. She also recognized early the need to publicize the *plight* of **AIDS victims**, and won tougher enforcement against fathers who stop paying child support. (*The New York Times*, 3 October 1985)
11. ACT UP celebrates 10 years with protest NEW YORK - ACT UP, the activist group whose brash tactics seek to publicize the *plight* of **AIDS victims**, marked its 10th anniversary Monday with a noisy demonstration in the heart of the city's financial district. (*The New Orleans Times-Picayune*, 25 March 1997)
12. Her well-publicized story already has drawn much-needed publicity to just how at-risk everyone is to AIDS and produced empathy for the *plight* of **AIDS victims**. (*Seattle Post-Intelligencer*, 27 March 1992)
13. The community activists in Los Angeles who organized the traditional Mexican procession, a candlelight re-enactment of Joseph and Mary's biblical search for shelter, known as a *posada*, hoped this jarring piece of symbolism would underscore the *plight* of Hispanic **victims of AIDS**. (*The New York Times*, 19 December 1989)
14. The implication of the cartoon that the pope does not care about the *plight* of **AIDS victims** was insulting and outrageously untrue. (*The Washington Post*, 28 March 2009)
15. Since the fatal disease is often a consequence of homosexual conduct, which the Roman Catholic Church condemns, many people, especially in the homosexual community, believe the church and its hierarchy are unsympathetic to the *plight* of **AIDS victims**. (*Chicago Sun-Times*, 5 September 1987)
16. Many expressed concern over what they see as an uncaring, if not a punitive, government attitude toward the *plight* of **AIDS victims**. (*The Wall Street Journal*, 8 June 1987)
17. David Falcone, an AIDS activist from New York, cradles Kathy Chou of ACT-UP to dramatize the *plight* of female **AIDS victims**. (*USA Today*, 3 October 1990)

Figure 8-4: Sample concordance lines of *plight* collocating with *AIDS victims*

The only cases of additional attributive adjectival pre-modification of *victim(s)* belong to existing minority groups, two belonging to known risk groups, and one counterintuitive

to the expected victim: *Hispanic* (line 13), *African*, and *female* (line 17). This further contributes to the prosody of deviancy doubling, explored in previous chapters.

In reviewing the concordance lines for *sufferer(s)*, it becomes apparent that five of 12 collocations with *plight* are not true collocations, but reprinted iterations of a passage containing the phrase, "...ignorance at the anguish of AIDS and **HIV sufferers**, the plight of the *homeless*, the isolation of *lepers*...", where *plight* actually belongs to *the homeless*, and *anguish* belongs instead to *AIDS and HIV sufferers*. After discounting these instances, the collocation of *plight* with *sufferer(s)* drops below the minimum frequency of 10, and is therefore discarded with an additional caveat of the import of looking at co-occurrences in context before deeming them true collocations.

#### **8.4.1.3 Confidentiality**

Only one collocate – *confidentiality* itself – is suitable for this subcategory. Its co-occurrence with *AIDS/HIV patient(s)* is relatively low, resulting in only 28 lines, all of which positively evaluate *patient* confidentiality. This pattern is illustrated in line 18 below, where the state [New York] is appraised as protecting *confidentiality* with "commendable consistency". However, nearly all of the extended texts indicate that *confidentiality* is a main sticking point in debates about maintaining registers of PWAs in America. Meanwhile, the state has protected the *confidentiality* of **AIDS patients** with commendable consistency for years. (*Newsday*, 18 January 1998)

18. However, it has drawn intense emotional opposition based on a perception that patient confidentiality could be harmed - even though Illinois' record in protecting the *confidentiality* of **AIDS patients**, whose names have been collected for years, is unblemished. Doctors share the concern about confidentiality. But 28 states have implemented name reporting without compromising patient privacy. (*Chicago Sun-Times*, 12 October 1998)

#### **Figure 8-5: Sample concordance lines of *confidentiality* collocating with *AIDS patients***

Discussion of *confidentiality* is inextricably intertwined with negotiation of protecting privacy (e.g. by using a system identifying HIV-infected patients by a code rather than a name) and being able to act quickly to protect public interest (e.g. by switching to a name reporting system, thereby doing away with the anonymity of codes and opening up PWAs to the possibility of privacy breaches). It appears that while *confidentiality* as a concept is well acknowledged, *confidentiality* as an ongoing practice is more controversial.

#### 8.4.1.4 Classification

Classification, or the identification of possible or certain PWAs, is a rare subcategory in this USAS broad category; see section 8.4.6 (X: psychological states) for a more complete discussion on this phenomenon. Categorised under A, only two collocates contribute to the discourse of classification: *classify* with *patient(s)* and *potential* with *carrier(s)*.

Social actor	CLASSIFICATION A category collocates	% A category collocates	% overall collocates
<i>patient(s)</i>	classify	9.1%	1.0%
<i>victim(s)</i>	-	0.0%	0.0%
<i>sufferer(s)</i>	-	0.0%	0.0%
<i>carrier(s)</i>	potential	50.0%	7.1%

Table 8-8: USAS Category A subcategory ‘Classification’ collocates by naming strategy

The collocation of the lemma *classify* with *patient(s)* indicates a schema present within the corpus. This collocation is quite infrequent, occurring with the minimum threshold of 10 instances. *AIDS patients* are classified by medical parameters or by risk group. Examples of the former can include either definitions for purposes of gaining access to legal protection or government benefits (e.g. “disabled” or “all those with severe immune-system impairment”) or definitions of stage (e.g. “terminal”). Classes of risk group include items such as “homosexual transmission”, “heterosexual transmission”, “IV drug users”, and “no known risk factors”.

While *classify* does not collocate with *carrier(s)*, and *potential* does not collocate with *patient(s)*, the uses of these two words in connection to the naming strategies in concordance lines are not so discrete. In seven out of 19 concordance lines, members of specific risk groups are named as *potential carrier(s)*: sex workers (3), foreigners (3), homosexual/bisexual men (2), drug users (2), and black people (1). The other sense of *potential carrier* is a person who has not yet been tested for AIDS/HIV, which corresponds to the medical sense detailed in the *classify* collocation. Though *carrier(s)* are not always *classified* overtly by their risk groups, their *potential* for becoming a *carrier* is covertly linked to the same risk factors. I argue that the use of a word like *potential*, which signifies suspicion rather than surety, is in fact more damaging to the social groups, particularly when taking into account that those discussed in these reports are not necessarily PWAs.

#### 8.4.1.5 Exclusivizers/particularizers/maximizers/comparison

Collocation with items from this subcategory is the unique purview of the naming strategy *patient(s)*, and all contribute to social sequestering. Each of 12 instances of

collocation between *solely* and *patient(s)* appear in description of medical centres dedicated specifically (or *solely*) to the treatment of PWAs. However, things that are construed as belonging *exclusively* to *patient(s)* are slightly more diverse.

Social actor	EXCLUSIVIZERS ETC. A category collocates	% A category collocates	% overall collocates
<i>patient(s)</i>	exclusively, severely, commonly, solely	36.4%	4.2%
<i>victim(s)</i>	-	0.0%	0.0%
<i>sufferer(s)</i>	-	0.0%	0.0%
<i>carrier(s)</i>	-	0.0%	0.0%

Table 8-9: USAS Category A subcategory ‘Exclusivizers etc.’ collocates by naming strategy

In 29 concordance lines, 20 cases occur wherein medical treatment centres/practices/nursing homes are *exclusive* to PWAs. Two occurrences each indicate that housing scenarios and charitable funds are additionally available *exclusively* to PWAs. Nevertheless, it is not only beneficial services which are *exclusive* to PWAs; in four concordance lines, certain cancers are seen *exclusively* in AIDS *patients*, and in one more, they are the exclusive consumers of an experimental drug. A similar semantic profile surrounds *commonly* in its collocation with *patient(s)*. In 20 concordance lines, illnesses, infections, and afflictions are described as *commonly* occurring in (or killing) AIDS *patient(s)* 18 times; in the remaining two concordance lines, drugs *commonly* used to treat PWAs are listed. Overall, collocates from this subcategory (*solely, exclusively, commonly*) construct the universal experience of the PWA as something discrete from the non-PWA majority, defined by segregation in access to medical care facilities and housing, and to types of treatment.

Though *severely* does emerge here as a collocate, it is as a modifier indicating degree of illness where this term is most salient. Therefore, discussion of this particular word will be found in the next section, USAS Category B: The body and the individual.

#### 8.4.2 Category B: The body and the individual

The next USAS broad semantic category to be analysed in this chapter has the second highest average frequency of collocates per naming strategy. In an analysis of AIDS/HIV reporting, it is not surprising to find that the semantic domain of the body and the individual features so highly—it is much more surprising that this did not come up previously. The *PWA/H* naming strategy is framed more within social experiences, as opposed to *patients* and *victims*, whose construction is much more focussed upon

medical aspects of their conditions. As a point of illustration, nearly 35% of all collocates of *patient(s)*—the most frequent naming strategy—fall under the B category.

Naming strategy	Collocates from USAS category B: THE BODY AND THE INDIVIDUAL	% total collocates
<i>patient(s)</i>	<b>treat</b> [LL=7641.663, MI=5.442], <b>azt</b> [LL=645.228, MI=4.003], <b>hospice</b> [LL=598.956, MI=4.427], <b>ill</b> [LL=369.781, MI=3.314], <b>afflict</b> [LL=311.806, MI=4.782], <b>bed</b> [LL=283.851, MI=3.279], <b>nursing</b> [LL=278.155, MI=3.518], <b>pneumonia</b> [LL=201.961, MI=3.635], <b>hospitalize</b> [LL=159.053, MI=4.224], <b>transplant</b> [LL=148.537, MI=3.017], <b>hospitalized</b> [LL=138.631, MI=5.932], <b>pneumocystis</b> [LL=137.226, MI=5.052], <b>blindness</b> [LL=135.59, MI=4.878], <b>pediatric</b> [LL=131.306, MI=3.576], <b>cytomegalovirus</b> [LL=114.503, MI=5.741], <b>kaposi</b> [LL=111.271, MI=4.124], <b>antiretroviral</b> [LL=106.2, MI=3.991], <b>lymphoma</b> [LL=103.093, MI=4.223], <b>marrow</b> [LL=98.316, MI=3.532], <b>cmv</b> [LL=92.523, MI=5.559], <b>retinitis</b> [LL=86.058, MI=5.525], <b>opportunistic</b> [LL=79.137, MI=4.215], <b>leper</b> [LL=63.674, MI=4.912], <b>anemia</b> [LL=56.1, MI=3.568], <b>chemotherapy</b> [LL=53.242, MI=3.013], <b>medical</b> [LL=50.589, MI=4.705], <b>sarcoma</b> [LL=50.429, MI=3.459], <b>ulcer</b> [LL=49.812, MI=3.908], <b>unapproved</b> [LL=42.238, MI=4.415], <b>neurological</b> [LL=39.085, MI=3.688], <b>ddi</b> [LL=37.634, MI=3.236], <b>ribavirin</b> [LL=36.061, MI=3.944], <b>hospitalization</b> [LL=35.68, MI=3.276], <b>susceptible</b> [LL=28.343, MI=3.147], <b>dementia</b> [LL=27.403, MI=3.283]	34.4%
<i>victim(s)</i>	<b>treat</b> [LL=443.546, MI=3.796], <b>carrier</b> [LL=178.05, MI=4.766], <b>dying</b> [LL=177.426, MI=4.753], <b>acquire</b> [LL=128.717, MI=3.929], <b>hospice</b> [LL=66.124, MI=3.484], <b>handicapped</b> [LL=49.72, MI=4.653], <b>hemophiliac</b> [LL=46.955, MI=3.962], <b>suffering</b> [LL=34.952, MI=3.417]	16.7%
<i>sufferer(s)</i>	<b>treat</b> [LL=96.479, MI=3.736]	6.7%
<i>carrier(s)</i>	<b>infect</b> [LL=47.624, MI=3.461], <b>victim</b> [LL=29.907, MI=3.254]	14.3%

Table 8-10: Collocates for each naming strategy from USAS category B

A series of subcategories have been created to ease discussion of this broad range of results. Groups of collocates will be discussed in the following subcategories: conditions; drugs and treatment; medical services and sites; illness; and other groups.

#### 8.4.2.1 Conditions

The only naming strategy to have numerous collocates in the CONDITIONS subcategory is *patient(s)*, though this accounts for over 42% of the naming strategy's collocates in the B category, and nearly 15% of its overall collocate types. Collocation with highly technical medical terminology contributes to the lexical priming of *patient(s)*, who

discursively occupy a greater portion of the biological experience of the disease than the sociological or psychological ones.

Social actor	CONDITIONS B category collocates	% B category collocates	% overall collocates
<i>patient(s)</i>	pneumonia, pneumocystis, blindness, cytomegalovirus, kaposi['s sarcoma], lymphoma, cmv, retinitis, opportunistic, anemia, sarcoma, ulcer, neurological, dementia	42.4%	14.6%
<i>victim(s)</i>	acquire	12.5%	2.0%
<i>sufferer(s)</i>	-	0.0%	0.0%
<i>carrier(s)</i>	-	0.0%	0.0%

**Table 8-11: USAS Category B subcategory 'Conditions' collocates by naming strategy**

A large number of collocates in this subcategory are also extremely specialized and technical terms and acronyms such as *CMV* would not be accessible to the average reader—a contributing factor to distance created between the experiences of the PWA and the news consumer. Fewer than half of the collocates here might be expected to belong to the lexicon of the average reader, removing them linguistically from the construed experience of *patient(s)* in the articles contained in the AIDS/HIV corpus. *Victim(s)* only collocates with a single lemma from this subcategory: *acquire*. The large majority (34 out of 36) of instances contributing to collocation can be attributed to in-text definitions of AIDS as “*acquired* immune deficiency syndrome”; these all occur in or before 1989. Here, *victims* are collocating with classification of their own chief syndrome, as opposed to description alongside the variety of other related conditions. This indicates that *victim(s)* are victimised by AIDS/HIV specifically, whereas *patient(s)* are patients of medicine more generally, as indicated by the following section.

#### **8.4.2.2 Drugs and treatment**

Like the previous subcategory, *patient(s)* is the only naming strategy with collocates belonging to the class of drugs and treatment. These two subcategories dovetail, though they are not identical, with the previous collocates of conditions constituting the causes leading to the current ones of drugs and treatment. Likewise, the collocates in this subcategory are of highly specialised, technical language.

Social actor	DRUGS AND TREATMENT B category collocates	% B category collocates	% overall collocates
<i>patient(s)</i>	azt, antiretroviral, unapproved, ddi, ribavirin, chemotherapy, transplant, marrow	21.1%	7.3%
<i>victim(s)</i>	-	0.0%	0.0%
<i>sufferer(s)</i>	-	0.0%	0.0%
<i>carrier(s)</i>	-	0.0%	0.0%

Table 8-12: USAS Category B subcategory ‘Drugs and Treatment’ collocates by naming strategy

Two acronyms, *DDI* and *AZT*, appear to similar distancing effect as discussed above. Of 32 instances of collocation between *patient(s)* and *marrow*, 23 of these refer to experimental procedures by which baboon bone marrow was *transplanted* into human PWAs. While this thesis is not reactionary in scope, I theorise that readers might additionally feel alienated from PWAs when confronted with medical discourse of scientific advances marrying animal and human biology. A greater proportion of the DRUGS AND TREATMENT subcategory could conceivably belong to the layman’s lexicon, with items like *transplant* and *chemotherapy* being widely well-known, with *marrow* perhaps only slightly less so.

#### 8.4.2.3 Medical services and sites

Social actor	MEDICAL SERVICES AND SITES B category collocates	% B category collocates	% overall collocates
<i>patient(s)</i>	<i>treat</i> , medi-cal, <i>hospice</i> , nursing, hospitalize, hospitalized, bed, hospitalization	23.4%	8.3%
<i>victim(s)</i>	<i>treat</i> , <i>hospice</i>	25.0%	4.1%
<i>sufferer(s)</i>	<i>treat</i>	100.0%	7.1%
<i>carrier(s)</i>	-	0.0%	0.0%

Table 8-13: USAS Category B subcategory ‘Medical services and sites’ collocates by naming strategy

In the MEDICAL SERVICES AND SITES subcategory, a greater distribution of collocates amongst the naming strategies begins to appear. The lemma of the verb *treat* collocates with *patient(s)*, *victim(s)*, and *sufferer(s)*—being the only B category collocate for the latter. However, in close analysis, it becomes clear that this while this word has been semantically tagged into the B category for its medical sense (*to treat* as in *to give medical care to*), it is also used in a behavioural sense (*to treat* as in *to behave toward or deal with*). The medical sense is the more frequent nonetheless; in collocations with *patient(s)*, *treat* is medical 1299 times and behavioural 23 times; with *victim(s)*, medical 94 times, behavioural 36 times; with *sufferer(s)*, medical 18 times, behavioural 11 times.



Again, the greatest variety of collocates is found with *patient(s)*, including healthcare sites (*hospices*, [hospital] *beds*, and *nursing* [homes]), services (*Medi-Cal*<sup>20</sup>), and processes (*treat*, *hospitalize*, *hospitalized*, *hospitalization*) all collocate. *Victim(s)* collocate only with one healthcare site—*hospice*—which offers holistic end-of-life care. This contrasts to *patient(s)*, which collocates with several *hospital* root words—a site which focuses on aggressive medical treatment rather than palliative care. Further, over a third of the *treat* collocates of *victim(s)* are used in the behavioural sense, meaning that *victim(s)* are comparatively proportionally more often discussed as objects not of medical treatment but of behavioural conduct towards them. Compare, for instance, concordance line 19 below (the medical sense) to line 20 (the behavioural sense).

19. Scientists from developing countries lamented the resources that AIDS will drain away from other pressing priorities, their inability to buy even basic technology to screen blood for the AIDS virus, the impossible expense of *treating AIDS victims with new drugs such as AZT*, which costs \$10,000 per patient per year. (Boston Globe, 8 June 1987)
20. Since last July, the proportion of Americans who believe that **victims of AIDS** should be *treated with compassion* has jumped to 87 percent, from 78 percent, the poll found. (New York Times, 22 November 1987)

Figure 8-6: Sample concordance lines for *victim(s)* collocating with *dying* and *suffering*

At the far end of this cline is *sufferer(s)*, whose only collocate in medical services and sites subcategory is *treat*, and over half of the instances of this collocation are of the behavioural sense. *Sufferer(s)* are very frequently constructed in the frame of their social treatment rather than their medical treatment.

#### 8.4.2.4 Illness

The illness subcategory has moderate spread across three of the four naming strategies under analysis, with *victim(s)* showing a strong preference (37.5% of all B category collocates are contained within this subcategory) and *sufferer(s)* showing a dispreference (with no collocates whatsoever coded here).

---

<sup>20</sup> Medi-Cal is the California state healthcare service.

Social actor	ILLNESS B category collocates	% B category collocates	% overall collocates
<i>patient(s)</i>	ill, afflict, susceptible	6.1%	2.1%
<i>victim(s)</i>	dying, suffering, handicapped	37.5%	6.1%
<i>sufferer(s)</i>	-	0.0%	0.0%
<i>carrier(s)</i>	infect	50.0%	7.1%

Table 8-14: USAS Category B subcategory 'Illness' collocates by naming strategy

The most common collocate of *patient(s)* in the ILLNESS subcategory is *ill*, with a collocational frequency of 132 (including 4 false collocations). In 115 of these cases, *ill* is itself modified with one of the adverbs in table 8-15 below; in nearly half of the cases, *patient(s)* are described as *terminally ill*. In an additional five cases, PWAs are described as being *too ill* to care for themselves, and in one case, *very ill*. Use of the word *ill* in connection with AIDS is redundant and nearly always leads into a descriptor of degree and/or severity. See the illustrative example in concordance line 21 below, where AIDS patients are redundantly described as "terminally ill".

21. They said this would insure that terminally ill AIDS patients would receive the compassion and care that hospitals rarely can provide. (*New York Times*, 24 December 1985)

Figure 8-7: Sample concordance line of AIDS patient collocating redundantly with *ill*

An exception is *mentally ill*, a relatively infrequent (6) fixed bigram. Nearly all tie back to end-of-life prosody, with infrequent adverbial modifiers *desperately* and *terribly* offering the only obviously emotive overtone.

Number	Adverb modifier of <i>ill</i>	Frequency of occurrence
1	terminally	55
2	seriously	15
3	severely	11
4	critically	9
5	chronically	7
6	<u>mentally</u>	6
7	acutely	4
7	desperately	4
9	gravely	2
10	sufficiently	1
11	terribly	1

Table 8-15: Frequency of adverbs modifying the collocation of *patient(s)* with *ill*

In each of 64 of 66 total collocations of *afflict* with *patient(s)*, the *patient(s)* are the direct objects, who are *afflicted* by various (named) AIDS/HIV-related conditions. In the two other cases, the tagger has misidentified the instance of *afflict* as belonging to the verb

lemma, e.g. in “The condition is frequently found in AIDS patients afflicted with cytomegalovirus retinitis”, it is still the cytomegalovirus retinitis which is the named afflicter, as it were, but the afflicting is an attribute here. *Affliction*, then, is an extension of the medical lexicon, wherein *patient(s)* are construed as being attacked or victimised by the named agents: abstract conditions. An extremely similar pattern occurs with *susceptible*, which does not then call for extended repetitive discussion.

Moving to the next naming strategy, collocates of *victim(s)* from the ILLNESS subcategory include two participles as attributive adjectives: *dying* and *suffering*. This particular group is modified with processes as attributes, indicating that their *dying* and *suffering* is enduring—a perpetual and unchangeable part of their being. These are among the most emotionally charged concordance lines of the corpus, in my opinion. Those who assist the *suffering* and *dying* offer “comfort and help” (line 22) to PWAs, whose current conditions are described as truly miserable: they struggle “to survive under unbearable hardships” (line 23) and “terrible physical sufferings” (line 24).

22. The nomination letter goes on to recount an incident in which DeRoche left a support-group retreat to dash back to Seattle and remain at the beside (on a mattress on the hospital room floor) of a *dying AIDS victim* and to offer comfort and help. (Seattle Post-Intelligencer, 29 January 1992)
23. “Many of the leading charities in our city,” he added, “depend on these contributions to ease the suffering of those struggling to survive under unbearable hardships: the traumatized homeless, the abandoned children, the *dying victims of AIDS* and tuberculosis, the homebound elderly in need of hot meals, the new immigrants with no place to go. (New York Times, 29 November 1992)
24. It is quite tragic that she can use incomprehensible euphemisms such as “lesions” that are “blooming on his body” to describe the terrible physical sufferings of an *AIDS victim*. She says “the time has come” and “you can feel it, in a hundred little ways year after year,” alluding to a future time when it will be more commonplace to be openly gay. (New Orleans Times-Picayune, 21 April 1994)

**Figure 8-8: Sample concordance lines for *victim(s)* collocating with *dying* and *suffering***

Another enduring state arising in the collocates is *handicapped*, which collocates with *victim(s)* only 11 times; of these, 6 discuss passage of acts which protect AIDS/HIV *victim(s)* under an extension of *handicapped* discrimination laws. The rest link together PWAs and the *handicapped* in ways similar to the deviancy doubling previously discussed. Again, PWAs are construed as *victim(s)* of their bodies and circumstances, much like the *handicapped* and other relatively powerless minority groups (see, for instance, line 25 below).

25. Two weeks earlier, his opponent, Senator Lowell P. Weicker Jr., announced his candidacy in Farmington by calling himself a champion of women, minority group members, AIDS victims, the handicapped and those “on the outside of America looking in.” (*New York Times*, 28 February 1988)

Figure 8-9: Sample concordance line for *carrier(s)* collocating with *infect*

Standing apart from these discourses, *carrier(s)* has only one collocate in this subcategory (constituting half of its overall collocate count in the USAS B category): *infect*. In each of the 16 concordance lines presenting this collocation, it is the *carrier(s)* who are the agents of infection. Named patients of the infection process include “ordinary people” (who we must assume exclude PWAs) as in line 26 below, and a *wife* (after the husband’s infidelity; concordance line 27). Several discourses emerge in the *infect* concordance lines. In lines 26-28, we see examples of the discourse of personal risk, wherein *AIDS carrier(s) infect* individual people. In line 28, they might even be actively “encouraged” to do so. Line 29 indicates a discourse of corporate risk; both in terms of carriers infecting other employees, but also as hurting the image of a company. The positioning of company image alongside human health or lives suggests an equivalence between the two, which is troubling.

26. Others proclaimed that “**AIDS carriers infect ordinary people**” and “No special privileges for gays.” (*Chicago Sun-Times*, 30 July 1986)
27. After the bar pickup and the mirror message, he became an **AIDS carrier** and *infected his wife*. (*Chicago Sun-Times*, 21 December 1986)
28. A company operating out of a medical office here is advertising and performing a test for exposure to the AIDS virus and offering clients “a photo ID verifying negative test results.” AIDS researchers say the test can be unreliable, sometimes generating false positives or false negatives and possibly encouraging AIDS carriers to infect others. (*San Francisco Chronicle*, 7 July 1986)
29. Firms fear **HIV carriers** can *infect other employees and hurt a company’s image*. (*Wall Street Journal*, 16 April 1993)
30. If one **HIV carrier** *infects* eight others, and they all rely on Medicaid for lifetime health-care costs, the average cost of more than \$119,000 per patient adds up to about \$1 million shelled out by taxpayers. (*Denver Post*, 9 February 1997)
31. In hearings earlier this month, commission member Dr. William A. Walsh said guidelines are needed to protect people from being *infected* by the AIDS carrier or patient who knowingly tries to spread the virus. “It seems to be that the health of the public is losing in this fight,” Walsh said, asking a panel of experts if “they could please give us a set of recommendations of when is it appropriate and proper to impose criminal prosecution” in such cases. (*Washington Post*, 19 April 1988)
32. The next night we are back to AIDS, when the intrepid district attorneys on ABC’s “Equal Justice” (Jan. 9, 10-11 p.m. EST) prosecute an **AIDS carrier** who has *infected* a young woman. (*Wall Street Journal*, 31 December 1990)

Figure 8-10: Sample concordance lines for *carrier(s)* collocating with *infect*

The focus on the economic costs of ‘AIDS carriers’ is continued in line 30 which references a different aspect of moral panic, keying into capitalist discourse. Here the folk devils are a financial risk if they *rely* on government healthcare—they lead to taxpayers ‘shelling out’ millions to care for them. “*The carrier*”—a genericised grammatical term indicating a single identity and homogenous aim—is construed as a social and legal risk in line 31, where the “health of the public” is being engaged in a metaphorical battle with PWAs and losing. Hypothetically, “the *carrier*” here “knowingly tries to spread the virus”, in the face of a seemingly helpless government. Concordance line 32 shows how these various elements—the AIDS carrier as a villainous agent, dramatically infecting a young (more innocent) woman, being brought to justice by “intrepid district attorneys” have become a cultural trope. In this line, AIDS is also positioned simply as a “topic” that the average listener can “come back to”, but does not have to deal with on a personal or daily basis.

#### 8.4.2.5 Other groups

The last subcategory in the USAS B broad category is other groups, or naming strategies for social actors with a semantic prosody of the body embedded, as detected by the USAS semtagger. The presence of other identities as collocates of three out of four of the naming strategies here does indicate a certain level of deviancy doubling that compels me to carry out a deeper investigation.

Social actor	OTHER GROUPS B category collocates	% B category collocates	% overall collocates
<i>patient(s)</i>	leper	3.0%	1.0%
<i>victim(s)</i>	carrier, hemophiliac	25.0%	4.1%
<i>sufferer(s)</i>	-	0.0%	0.0%
<i>carrier(s)</i>	victim	50.0%	7.1%

Table 8-16: USAS Category B subcategory ‘Other groups’ collocates by naming strategy

*Patient(s)* collocates with only one other group in this subcategory—this is *leper(s)* with a total of 13 collocations, the majority of which are in posthumous reference to the charity work of Princess Diana (8), or a report of a speech made by the pope (3). Princess Diana and Mother Teresa are positively evaluated for their work with PWAs and lepers, as illustrated in sample concordance lines 33 and 34, where the former is proven to be “an angel” for touching the infirm, and the latter is reported as having “truly recognized Jesus” in deviant, helpless groups. While concordance line 35 does show a true collocation in that *AIDS patients* and *lepers* do co-occur within the set parameters and clause boundaries, this does not show ideological linking so much as reporting of

crowd attendance of the leprosarium and religious centre; this is indicated by the third list member, *health workers*. Concordance line 36 is the only example where the *AIDS patients* and *lepers* are directly correlated; here an extreme far-right French politician negatively evaluates the former in the frame of the latter, though the journalist distances himself by using the reporting term “likened” and a direct quotation pattern.

33. [Princess Diana]’d already proved herself an angel among **AIDS patients** and *lepers* when she reached out and touched them to show the world that it was not only safe to touch the infirm, but also a way to offer healing dignity. (*San Francisco Chronicle*, 1 September 1997)
34. Hickey said Mother Teresa “truly recognized Jesus” in *lepers*, in **AIDS patients**, in the homeless and in neglected infants. (*Washington Post*, 8 September 1997)
35. The pontiff’s appeal on behalf of political prisoners came later Saturday, during a dramatic evening pilgrimage to a hallowed Cuban shrine for the sick - the leprosarium and religious center of St. Lazarus, outside Havana, where he spoke to health workers, *lepers* and **AIDS patients**. (*New Orleans Times-Picayune*, 25 January 1998)
36. In 1987 [Jean-Marie Le Pen] likened **AIDS patients** to *lepers* “sweating the virus through their pores,” and said that he was “not thrilled at the idea of Arab boys having sex with girls from Strasbourg.” (*New York Times*, 6 May 2002)

**Figure 8-11: Sample concordance lines for *patient(s)* collocating with *leper***

In this subcategory, we also see the interesting interaction between naming strategies themselves: *victim(s)* and *carrier(s)* both collocate with one another. Though this confirms that they are distinct social groups in the discourse of AIDS/HIV, their majority positions on either side of conjunctions such as “and” and “or” do not overtly indicate the points of difference that have motivated journalists to list two separate nominations when one might have sufficed. This suggests a move toward conflation of the terms, wherein *carrier(s)* are *victim(s)* and vice-versa, though it was only *carrier(s)* who performed the act of *infecting* above.

Finally, *victim(s)* collocates with *hemophiliacs*, a social group which marked by lack of blame not realised by those explored above. In four of 13 concordance lines, the *hemophiliacs* mentioned are minors, which adds an additional layer of innocence to the construal of their characters. See for instance line 31 below, where the stoning of the vehicle is even more emotionally charged due to its underage occupant, prefaced by the appraisal of discrimination as “ugly, irrational and born of ignorance”.

37. Discrimination is ugly, irrational and born of ignorance. It caused Lake City, Tenn., school parents to stone the vehicle carrying 12-year-old hemophiliac and **AIDS victim** DeWayne Mowery. (*USA Today*, 5 August 1988)

Figure 8-12: Sample concordance line of *victim(s)* collocating with *hemophiliac*

None of the collocations in this subcategory create a positive prosody around PWAs or even soften the negative prosody that is carried across texts with them. They appear alongside other HIV+ individuals or social pariahs (e.g. *lepers*), and it is not the last time in this chapter that long lists of ‘others’ are constructed with PWAs at the centre: see section 8.4.4.1 for more.

### 8.4.3 Category G: Government and public

Collocates from the USAS G category are the most evenly distributed in the entire chapter, with a common (and unfortunately negative) discourse running across *patient(s)*, *victim(s)*, *sufferer(s)*, and *carrier(s)*. Only two items below are unique, with each of the others shared between two or more naming strategies.

Naming strategy	Collocates from USAS category G: GOVERNMENT AND PUBLIC	% total collocates
<i>patient(s)</i>	<i>discrimination</i> [LL=441.077, MI=3.799], <i>discriminate</i> [LL=82.202, MI=3.599]	2.1%
<i>victim(s)</i>	<i>discrimination</i> [LL=1259.973, MI=6.116], <i>discriminate</i> [LL=235.513, MI=5.883], <i>protect</i> [LL=213.555, MI=3.552], <i>prohibit</i> [LL=77.159, MI=4.003]	8.3%
<i>sufferer(s)</i>	<i>discrimination</i> [LL=243.304, MI=5.914], <i>protect</i> [LL=44.413, MI=3.455], <i>ban</i> [LL=43.557, MI=3.766]	20.0%
<i>carrier(s)</i>	<i>discrimination</i> [LL=95.598, MI=5.465], <i>protect</i> [LL=64.908, MI=4.297]	14.3%

Table 8-17: USAS Category G collocates by naming strategy

The abstract noun *discrimination* collocates with all of the naming strategies, with *patient(s)* and *victim(s)* additionally collocating with the verb form. However, *discrimination* is a far more frequent collocation, with frequencies and text distributions as follows:

	<i>discrimination</i>		<i>discriminate</i>	
	Collocate freq.	In no. of texts	Collocate freq.	In no. of texts
<i>patient(s)</i>	129	122	26	26
<i>victim(s)</i>	193	171	38	31
<i>sufferer(s)</i>	39	38	-	-

Table 8-18: Frequency and distribution of *discrimination* and *discriminate* collocating with *patient(s)*, *victim(s)*, and *sufferer(s)*

Appearing in clauses as recipients of the verb form *discriminate*, *patient(s)* and *victim(s)* are more clearly impacted by the process than in clauses by which the process has been nominalised into *discrimination* and paired with PWAs as an additional participant. The use of the verb form alludes to the question of agency (who is discriminating) while the use of the noun form can be used to obscure the other (agentive) human actors.

Of the 129 concordance lines featuring the collocation of *patient(s)* with *discrimination*, sources and/or sites of discrimination are given only marginally more frequently (66 times) than they are ellipited (63 times). However, when sources/sites are named, it is nearly always the circumstances of discrimination that are named rather than the (human) perpetrators of such.

As illustrated in Table 8-19 below, the most frequent circumstantial locations of discrimination are the workplace (66) and the housing market (41), with the former sometimes appearing conflated with abstract entities such as “businesses” and “employers” (nonhuman/nominalised). These are very frequently related to bills being debated and passed at various levels in the US government, wherein the two sites were often linked litigiously.

Source/site of <i>discrimination</i>	<i>patient(s)</i>	<i>victim(s)</i>	<i>sufferer(s)</i>	Total
Employment	26	30	10	66
Housing	16	17	8	41
Healthcare	14	10	2	26
Education	2	4	4	10
Government legislation/services	5	4	1	10
Human agents	4	2	0	6
Public/society	3	2	0	5
Transportation	0	2	0	2
Communications	0	2	0	2
Business/credit	2	0	0	2
Military	0	1	0	1
Private drug rehabilitation centers	0	0	1	1

**Table 8-19: Frequency of various sources/sites of *discrimination* by appearance in concordance lines of congruently realized PWA naming strategies**

Therefore, one direct effect of using the abstract, nominal form *discrimination* appears to be that other elements in the syntactical structure—namely, the actors—also become more abstract, as they are no longer the agents in active processes. In sample concordance line 38 below, doctors are described as having a “moral obligation” to treat PWAs; lack of care is considered a form of discrimination (line 39), and results in



legislation protecting PWAs from the public (line 39), and other institutional care figures (line 40). The 'general public' seems to be an expected source of discrimination in line 41, compared to health care workers, whose discrimination is marked by the work "even". Such personal discrimination, alongside institutionalised acts, are finally outlawed (see line 42).

38. Edmund Pellegrino, director of the Center for the Advanced Study of Ethics at Georgetown University, said doctors had a "moral obligation" to treat AIDS patients without discrimination. (*St. Louis Post-Dispatch*, 15 November 1989)
39. The study, the third to be produced by the 15-member commission since it was set up in November, 1988, focused on discrimination against AIDS patients by the public and by health care workers. (*Chicago Sun-Times*, 22 August 1990)
40. Legislation protecting **AIDS patients** from discrimination by doctors, medical technicians, firefighters and other health and law enforcement professionals cleared the House of Delegates today as part of a legislative package that also would toughen penalties for intentionally transmitting the deadly disease. (*Washington Post*, 24 March 1989)
41. But the report focused mainly on discrimination against AIDS victims by the general public and even by health-care workers. (*Washington Post*, 22 August 1990)
42. In a speech to begin the memorial service at the Bandstand, Mayor Koch announced that the city would prosecute any business or person that discriminated against AIDS victims if such discrimination is found to be illegal. (*New York Times*, 14 June 1983)

**Figure 8-13: All concordance lines showing human agents as sources of *discrimination***

Discourses surrounding *discrimination* mainly have to do with discussion of attempts to *ban* or *prohibit* associated practices, which accounts for these related collocates in the same USAS semantic category. As it is legislative practices which can effectively do so, the government and, by extension, its legislative practices, is again a major focus of this risk society sub-discourse.

#### **8.4.4 Category H: Architecture, housing and the home**

As demonstrated below in Table 8-20, each of the congruently realized human naming strategies collocate with *homeless*, another instance of deviancy doubling.

Naming strategy	Collocates from USAS category H: ARCHITECTURE, HOUSING AND THE HOME	% total collocates
<i>patient(s)</i>	<i>homeless</i> [LL=693.967, MI=3.626], <i>occupy</i> [LL=215.932, MI=4.548], <i>residence</i> [LL=162.717, MI=3.62], <i>residential</i> [LL=59.716, MI=3.063]	4.2%
<i>victim(s)</i>	<i>homeless</i> [LL=358.692, MI=4.14], <i>residence</i> [LL=90.75, MI=4.208]	6.3%
<i>sufferer(s)</i>	<i>homeless</i> [LL=121.681, MI=4.513], <i>housing</i> [LL=24.833, MI=3.079]	13.3%
<i>carrier(s)</i>	-	0.0%

Table 8-20: USAS Category H category collocates by naming strategy

In the subsections below, the various iterated collocations with *homeless* will be considered separately before I move on to the semantic categories of residence, and discussion of *occupy*.

#### 8.4.4.1 Homeless

Three of the naming strategies—*patient(s)*, *victim(s)*, and *sufferer(s)*—collocate with *homeless*. However, the nature of automated collocation obscures a certain level of detail in the relationship between the node and the collocate.

Social actor	HOMELESS H category collocates	% H category collocates	% overall collocates
<i>patient(s)</i>	<i>homeless</i>	25.0%	1.0%
<i>victim(s)</i>	<i>homeless</i>	50.0%	2.1%
<i>sufferer(s)</i>	<i>homeless</i>	50.0%	6.7%
<i>carrier(s)</i>	-	0.0%	0.0%

Table 8-21: USAS Category H subcategory 'Homeless' collocates by naming strategy

Intervening punctuation, conjunctions, and determiners could alter the meaning created; there is a large ideological different between, for instance, "*homeless AIDS victims*" (a compounded naming strategy to represent one social group) and "*AIDS victims and the homeless*" (a compounded naming strategy representing two social groups). Though simply limiting collocation to the -1 position ("*homeless AIDS victim*", etc.) would give a fair estimation of the number of times this collocation is indeed a pre-modification, more complex cases such as "*HIV and AIDS patients who are homeless or at risk of becoming homeless*" had to be identified through close reading.

A manual analysis has been undertaken to gain a better understanding of the interaction between the node words and this particular collocate. As shown in Table 8-18 below, parallel structures (e.g. "*the homeless and AIDS victims*") are marginally more frequent than compound structure (e.g. "*homeless AIDS victims*"), which suggests that this is a

collocation stemming from overlapping perceived social values rather than common attributes of a single social group. Uniquely, *victim(s)* far prefers the parallel structure over the compound structure, indicating that *homeless* people are less likely be construed as *victim(s)* themselves, than to be likened to them through a set of reported circumstances.

	patient(s)	victim(s)	sufferer(s)	Total
Parallel structures	101	64	13	178
Compound structures	112	29	15	156
False collocations with <i>homeless</i>	3	0	0	3
Total collocations with <i>homeless</i>	215	93	28	336

**Table 8-22: Frequency of collocation with *homeless* by various naming strategies, by parallel and compound structuring**

A total of 178 concordance lines featured parallel structures wherein PWAs were constructed in tandem to the *homeless* social group. However, many clauses from this parallel category also contained several other social groups. The top 15 most frequent social groups additionally appearing in tandem constructions have been collated in Table 8-19 below.

		patient(s)	victim(s)	sufferer(s)	Total
	the homeless	101	64	13	178
1.	drug abusers/alcoholics	20	5	1	26
2.	children	11	9	1	21
3.	immigrants/refugees/illegal aliens/migrant workers	10	4	2	16
4.	the poor	6	5	1	12
5.	prisoners	8	2	0	10
5.	the elderly	6	3	1	10
7.	the mentally ill	5	2	2	9
7.	the disabled	4	5	0	9
9.	victims of family violence (women and children)	3	4	1	8
10.	hospital/nursing home patients	6	1	0	7
11.	unemployed	1	2	1	4
12.	hungry	2	0	0	2
13.	homosexuals	0	1	0	1
14.	the overweight	0	1	0	1
15.	minorities	0	1	0	1

**Table 8-23: Frequency of additional social groups appearing in parallel structures**

In parallel structures, PWAs are linked with the stigmatized group *the homeless*, as well as with a variety of traditionally victimized or powerless roles, including children (2nd),

the elderly (5th), the disabled (7th), victims of family violence (9th), and hospital/nursing home patients (10th). Some of the most frequent associations are with social groups that are both powerless and customarily stigmatized and systematically othered including drug abusers/alcoholics (1st), immigrants/refugees/illegal aliens/migrant farm workers (3rd), the poor (4th), prisoners (5th), and the mentally ill (7th). *Homeless* PWAs, already doubled (or tripled, if type of transmission has been given or assumed) in deviancy, have a prosody of an undignified and disempowered other, quite often appearing in long strings of deviants such as those listed above. A few, including children, the poor, the elderly, and the mentally ill, also contributed to threaded deviancy strings in the Hurricane Katrina corpus.

In compound *homeless* [PWA] structures, discourse is dominated by discussion of housing availability and policy. The exact chain of events leading to homelessness (e.g. failure on the part of governmental policy, lack of access to healthcare, ostracism by family and peers) is only ruminated upon once in 157 compound *homeless* PWA concordance lines. An excerpt of this 1989 article from the *New York Times* is listed in concordance line 43 below. Similarly, only two concordance lines discussed the possibility of homelessness (see line 44 below for an example). The vast majority of articles dealt with PWA homelessness as an issue without context, which we will shortly see led to a discourse dominated by lack of educated discussion of trends or possibility of controlled effects.

43. Some **AIDS patients** are *homeless* to begin with, and city policy forbids discharging them from a hospital to a shelter as inappropriate. Others lose their homes, ejected by relatives, roommates or landlords. Sometimes they can no longer afford the rent, because of medical costs or because they have lost their jobs from illness or discrimination. (New York Times, 3 January 1989)
44. The organization will use a \$25,000 grant to pay for programs to keep **AIDS patients** from becoming *homeless*. (St. Louis Post-Dispatch, 21 June 2006)

**Figure 8-14: Concordance lines indicating past or future states of homelessness**

Though the compound *homeless* PWA concordance lines lack context of causation and background information relating to the circumstances of PWA homelessness in America, they do not lack information on current behaviours of *homeless* PWAs. All concordance lines offering clues as to the circumstances of contraction in *homeless* PWAs cited intravenous drug use (lines 45-48, 50, 51, with two additionally implying or listing gay/bisexual acts (lines 45-46) or prostitution (line 49). Concordance line 46 assigns ownership of HIV to the patients with the use of the possessive pronoun in “their

disease” as a substitute for the more commonplace determiner “the disease”, further indicating culpability in contraction.

45. Yet O'Connor offered to open a residence for *homeless victims of AIDS*, most of whom are gays, bisexuals or users of illegal intravenous drugs. (*Newsday*, 3 February 1986)
46. Unlike gay HIV patients who generally contracted their disease from unprotected gay sex, *homeless HIV patients* generally contracted their disease from dirty drug needles. It is therefore a fact that a substantial number of the residents of the proposed shelter are likely to be current or former drug addicts. (*Denver Post*, 3 February 2007)
47. Commissioner Curran observed: “Yonkers General has been one of the county's most cooperative hospitals in dealing with hard-to-place patients, including *homeless AIDS patients with a drug addiction*.” (*New York Times*, 11 January 1987)
48. The center is treating a growing number of *homeless AIDS patients who are IV drug users*. (*New York Times*, 3 August 1987)
49. For about a year, the city's HIV/AIDS Services Administration has taken over 35 rooms at the Royal York and filled them with *homeless AIDS patients*. Many have histories of drug abuse and prostitution. (*New York Times*, 8 December 2002)
50. Problems are also surfacing in other agencies, especially when the **AIDS patients** are *homeless intravenous drug users*. (*San Francisco Chronicle*, 14 June 1989)
51. The abundance of *homeless AIDS patients*, many of them drug addicts, in S.R.O.'s on the Upper West Side caused an uproar until the city agreed to a cap. (*New York Times*, 31 December 1995)

**Figure 8-15: Concordance lines indicating additional activities of *homeless* PWAs**

An additional discourse of strain and threat also arises in the compound *homeless* PWA concordance lines. Much like the displaced victims of Hurricane Katrina, while Americans are construed as theoretically (and often literally) caring for PWAs, they do not want this care to extend past charity into, for instance, their own neighbourhoods. The ‘rubbish’ metaphor encountered in the Hurricane Katrina metaphor also resurfaced in this instance. Residents stated that “their neighborhood had become a dumping ground for *homeless AIDS patients* and others needing shelter and social services” [emphasis added]. In concordance lines 48-49 above, a hint of the same containment discourse of ‘growing’ and ‘filling’ with numbers begins to appear. Concordance line 50 shows a familiar strain on a government agency coping with a difficult charge (including the threat/water metaphor “surfacing”), while concordance line 51 describes a community outcry against the settlement of the ‘others’ in a local neighbourhood. These discourses are nearly identical to those observed in Chapter 6, and seem to show the interplay between risk society (a strain and failure of a governmental body) and moral

panic (a resurgence in fear of a folk devil, and demand for protection from the governing bodies).

#### 8.4.4.2 Residence/Residential/Housing

Related to the discussion of homelessness in the corpus is that of residential needs of PWAs. Some aspect of this semantic subcategory collocates with each naming strategy save for *carrier(s)*, indicating that this identity is not construed in terms of needing dedicated *residential* care. Collocates in this subcategory are listed in Table 8-24 below.

Social actor	RESIDENCE H category collocates	% H category collocates	% overall collocates
<i>patient(s)</i>	<i>residence, residential</i>	50.0%	2.1%
<i>victim(s)</i>	<i>residence</i>	50.0%	2.1%
<i>sufferer(s)</i>	<i>housing</i>	50.0%	6.7%
<i>carrier(s)</i>	-	0.0%	0.0%

Table 8-24: USAS Category H subcategory 'Residence' collocates by naming strategy

For the most part, collocates in this subcategory carry with them a surprisingly positive semantic prosody. As seen in sample concordance line 52 below, care in *residential settings* is construed as something to be encouraged “as an alternative to hospitalization”. The same distinction is made clear in line 53; the remodelled convent “isn’t a hospice or hospital”, but a *residence* for PWAs. While the same lack of agency as previously discovered does continue to typify these lines (e.g. it is not PWAs who seek or establish their own *residences*, but these are provided to them once they “no longer can care for themselves”), it appears that these are a positively evaluated substitute for more traditional medical facilities, most likely because *residences* are funded by charities, not-for-profit organisations, NGOs, or private donations, in contrast to publically funded hospitals.

52. Clearing the Senate were measures designed to: Guarantee the anonymity of people tested for AIDS. Encourage care in nursing homes and other *residential settings* for **AIDS patients** as an alternative to hospitalization. (*Chicago Sun-Times*, 27 June 1987)
53. The remodeled convent at 1840 Grant St. isn't a hospice or hospital, but a *residence* for **AIDS patients** who no longer can care for themselves. (*The Denver Post*, 30 July 1995)
54. Despite a history of opposition from some community members, several Lincoln-Lemington residents sought to make clear yesterday that the new Corpus Christi *Residence* for **AIDS victims** has the support of some of its neighbors. “I live 12 doors away and I feel very comfortable with it,” said Sarah Gaines of Churchland Street, a board member of the facility, which has accepted one resident thus far. (*Pittsburgh Post-Gazette*, 29 July 1994)

55. The often brutal nature of this new budget politics was made particularly clear Thursday when a Republican, Rep. Christopher Shays (Conn.), attempted to get \$186 million restored for a *housing* program assisting **AIDS sufferers** and offset the cost by taking the same amount from NASA's science and technology projects account. (*Washington Post*, 18 March 1995)

**Figure 8-16: Concordance lines illustrating the collocation between USAS H subcategory 'Residence' collocates and various naming strategies**

The exception to this *housing*, which is the sole and unique collocate of *sufferer(s)* in this subcategory. Each of the 10 concordance lines arising from this collocation refer to government legislation, which does activate the more negative discourse of financial burden and strain on the taxpayer (e.g. in line 55 above).

In reviewing the concordance lines of collocates from this section, it becomes evident once more that citizens (or perhaps just readers of newspapers) are assumed to ideologically support care for PWAs, just so long as this does not infringe in any way on government spending or otherwise monopolize institutional resources. The final H category collocate—*occupy*—is even more marked in this sense.

#### **8.4.4.3 Occupy**

The unique collocation of *occupy* with *patient(s)* arises almost entirely from discussion of allocation of hospital beds and other medical resources; in 49 concordance lines, only twice do *patient(s)* *occupy* something other than medical treatment services. The discourse that arises from this collocation is one of shortage—with a certain number, percentage, or ratio of health services dedicated to PWAs, costs for (non-PWA) taxpayers are conveyed as increasing whereas services available to them are assumed to be decreasing. In concordance lines 56-57 below, PWAs are counted on the basis of growing numbers and financial need. The familiar *topos* of numbers (56, 57, 58) and strain (57) recur here, construing the PWA occupation as a threat. The dirty water metaphor reappears as well – in line 58, a hospital is “swamped” with AIDS patients. This collocation lends itself to an extremely negative semantic prosody of threat, strain, and pollution.

56. By 1991, **AIDS patients** will *occupy* 14,000 hospital beds each day at a cost of \$3.5 billion a year. Does that mean hospital care is going to be more expensive for the rest of us? (*USA Today*, 7 December 1987)

57. Nevertheless, there are increasing signs that these systems of care, developed largely on an ad hoc basis, are becoming severely strained. New York City reports long waits in the emergency departments of overcrowded hospitals where **AIDS patients** *occupy* a growing proportion of inpatient beds... There is a catastrophic shortage of facilities to care for intravenous drug users. (*Newsday*, 21 June 1988)
58. New York's flagship city hospital, Bellevue, is swamped with **AIDS patients**, who *occupy* 27 percent of its general beds. (*Boston Globe*, 27 December 1990)

Figure 8-17: Concordance lines illustrating the collocation between *occupy* and *patient(s)*

Category H collocates contributed to a discourse of threat and strain around PWAs, and symptoms of deviancy doubling with *homeless* people and further othered groups also appeared. The next USAS broad semantic section, S, contains a number of additional identity markers. It will be interesting to see how PWAs are construed as interacting in a social capacity.

#### 8.4.5 Category S: Social actions, states and processes

USAS S is the runaway preference of collocates for the four naming strategies being discussed here. This category is the first preference for both *victim(s)* and *sufferer(s)*, accounting for 25.0% and 20.0% of their total collocates. On average, 17.2% of all collocates belong to this semantic domain, which underlines its significance in the construal of PWAs.

Naming strategy	Collocates from USAS category S: SOCIAL ACTIONS, STATES AND PROCESSES	% total collocates
<i>patient(s)</i>	<b>care</b> [LL=3855.428, MI=3.699], <b>advocate</b> [LL=1432.859, MI=4.478], <b>advocacy</b> [LL=175.239, MI=3.451], <b>counsel</b> [LL=152.557, MI=3.513], <b>behalf</b> [LL=129.534, MI=3.54], <b>comfort</b> [LL=68.068, MI=3.018], <b>visiting</b> [LL=58.083, MI=3.286], <b>counsele[d]</b> [LL=47.942, MI=4.25]	9.4%
<i>victim(s)</i>	<b>homosexual</b> [LL=318.897, MI=3.93], <b>contact</b> [LL=173.833, MI=3.358], <b>memorialize</b> [LL=159.843, MI=7.474], <b>quilt</b> [LL=150.067, MI=4.469], <b>behalf</b> [LL=141.463, MI=4.791], <b>advocate</b> [LL=91.092, MI=3.048], <b>commemorate</b> [LL=68.057, MI=5.181], <b>deserve</b> [LL=55.161, MI=3.527], <b>dedicate</b> [LL=54.532, MI=3.798], <b>assist</b> [LL=48.468, MI=3.243], <b>tribute</b> [LL=40.039, MI=3.373], <b>embrace</b> [LL=39.514, MI=3.34], <b>casual</b> [LL=38.356, MI=3.641]	25.0%
<i>sufferer(s)</i>	<b>advocate</b> [LL=68.683, MI=4.113], <b>homosexual</b> [LL=64.856, MI=3.806], <b>fellow</b> [LL=51.999, MI=4.507]	20.0%
<i>carrier(s)</i>	<b>partner</b> [LL=47.19, MI=3.776], <b>contact</b> [LL=36.65, MI=3.74]	14.3%

Table 8-25: USAS Category S category collocates by naming strategy



Collocates around the discussion of the care and advocacy for living PWAs (section 8.4.5.1) and the commemoration of the dead (8.4.5.2), as well as public negotiation of private matters such as sexuality (8.4.5.3) and belonging (8.4.5.4) comprise this subsection.

#### 8.4.5.1 Care and advocacy

The subcategory of *care* and *advocacy* is the most frequent within the USAS S semantic field, encompassing a wide variety of meanings including emotional care (*visiting, comfort*), general care (*assist, care*), institutionalized assistance (*counsel, counsele[d], behalf*), and advocacy (*advocate, advocacy*). These are spread across three naming strategies – *patient(s)*, *victim(s)*, and *sufferer(s)* – though notably, *carrier(s)* are not the recipients of social care or advocacy.

Social actor	CARE AND ADVOCACY S category collocates	% S category collocates	% overall collocates
<i>patient(s)</i>	<i>care, advocate, advocacy, counsel, behalf, comfort, visiting, counsele[d]</i>	100.0%	8.3%
<i>victim(s)</i>	<i>behalf, advocate, deserve, assist</i>	30.8%	8.3%
<i>sufferer(s)</i>	<i>advocate</i>	33.3%	6.7%
<i>carrier(s)</i>	-	0.0%	0.0%

Table 8-26: USAS Category S subcategory 'Care and Advocacy' collocates by naming strategy

As shown in Table 8-26 above, *advocate* is a collocate of *patient(s)*, *victim(s)*, and *sufferer(s)*, with *patient(s)* additionally collocating with the abstract noun form *advocacy*. *Advocate* is an interesting collocate because it introduces an alternative social group into the discourse, one which is often more powerful as a result of being organised and well-connected, particularly with other powerful entities (such as the media). However, the presence of this, in addition with indicators such as *behalf* (collocating with both *patients* and *victims*) and previously analysed collocates *deserve* and *assist* signpost a familiar configuration in these concordance lines. The care and advocacy subcategory collocates create discourses around the people surrounding PWAs, empowering others to *advocate*, to work on their *behalf*, and to *assist* them, debating what they may or may not *deserve*. Once more, these collocates reveal little about PWAs except their lack of agency (see also the discussion of S category collocates in sections 4.3.8, 5.5.3.3, 6.3.1.2, 7.3.4, and 9.3.3 for similar findings).

Rather than dwelling on these familiar findings, I turn instead to novel collocations of *victim(s)*, subcategorized under a heading called commemoration.

### 8.4.5.2 Commemoration

A variety of words in the semantic field of commemoration collocate with *victim(s)*: *memorialize, quilt, commemorate, dedicate, and tribute*. This is the only naming strategy with collocates from the USAS S category of a commemorative semantic meaning.

Social actor	COMMEMORATION S category collocates	% S category collocates	% overall collocates
<i>patient(s)</i>	-	0.00%	0.0%
<i>victim(s)</i>	memorialize, quilt, commemorate, dedicate, tribute	38.5%	10.4%
<i>sufferer(s)</i>	-	0.0%	0.0%
<i>carrier(s)</i>	-	0.0%	0.0%

Table 8-27: USAS Category S subcategory ‘Commemoration’ collocates by naming strategy

While this might indicate that *victim(s)* are construed as more deserving somehow of memorialisation, over-interpretation here must be avoided. The most likely reason for a heightened frequency of commemoration collocates for *victim(s)* is that a sense of this naming strategy can include conclusion of life in a way that the others do not. In other words, PWAs labelled *patient(s)*, *sufferer(s)* and *carrier(s)* are more likely to still be living, and therefore would not be referred to in the frame of posthumous remembrances. This is very similar to findings in the matched Katrina Chapter 5.5.2.1, where *native* collocated with *resident*, and terms such as *son* and *daughter* collocated with *survivor*, meaning that they had survived a deceased family fatality of Hurricane Katrina.

59. In a Memorial Day observance, members of the Chicago gay and lesbian community paid *tribute to victims of AIDS* with a candlelight march and church service. (*Chicago Sun-Times*, 26 May 1987)
60. About 2,000 worshipers filled the Cathedral Church of St. John the Divine yesterday for “AIDS: A Service for Hope” - one of the city’s largest gatherings dedicated to AIDS victims and those who care for them. (*Newsday*, 9 November 1987)
61. Families and friends have created more than 25,000 3-by-6-foot panels to *memorialize AIDS victims*. (*Chicago Sun-Times*, 24 November 1993)
62. Families and friends of AIDS victims make the *quilt* sections and donate them to the project. (*Washington Post*, 1 April 1999)
63. Many walkers brought their families along, carried signs, and wore shirts *commemorating AIDS victims*. (*Boston Globe*, 7 June 1999)

Figure 8-18: Sample concordance lines showing ‘commemoration’ collocates of *victim(s)*

The similarities do not stop there. As in the Katrina corpus, when PWAs have died (as when PAKs had died), they are strongly associated with family members and loved ones in a way that is not true when they are living. Agents in concordance lines above include

“members of the Chicago gay and lesbian community (59). “those who care for [PWAs]” (60), “family and friends” (61 and 62), and families (63). In this sample of media discourse, it is only after death that PWAs pass from the realms of medical institutions and political advocacy back into the dominions of their own social circles, regaining some semblance of their own humanity, though of course agency will forevermore be focused elsewhere.

### 8.4.5.3 Sexuality

Sexuality accounts for a number of collocates in this category, and as a subcategory, is represented across nearly all of the naming strategies. Interestingly, *patient(s)* does not collocate with this subcategory, indicating that once a PWA is diagnosed with HIV/AIDS and assumes the identity of a *patient*, the form of transmission is no longer central to their condition. Expanding upon that, *homosexual* emerges as a collocate of both *victim(s)* and *sufferer(s)*, but not of *patient(s)* or *carrier(s)*. *Victim(s)* additionally collocates with [sexual] *contact*, *casual* [sex], and *carrier(s)* additionally collocates with [sexual] *partners* and [sexual] *contact*.

Social actor	SEXUALITY S category collocates	% S category collocates	% overall collocates
<i>patient(s)</i>	-	0.0%	0.0%
<i>victim(s)</i>	<i>homosexual, contact, casual</i>	23.1%	6.3%
<i>sufferer(s)</i>	<i>homosexual</i>	33.3%	6.7%
<i>carrier(s)</i>	<i>partner, contact</i>	100.0%	14.3%

Table 8-28: USAS Category S subcategory ‘Sexuality’ collocates by naming strategy

This method of transmission is clearly a salient aspect of the discourse around PWAs, but also one that has undergone radical change over time. For instance, consider the concept of “AIDS carriers” and the abandonment of this identity after the 1980s. For this reason, sexuality – and in particular homosexuality – forms the basis of an entire diachronic study, to be found in Chapter 9. Please refer forward for a greatly expanded discussion including each of these collocates and very many more, analysed through the lens of time.

### 8.4.5.4 Groups and belonging

Groups and belonging is the last and the smallest of the subcategories in USAS broad category S, containing just two collocates. These are obviously not particularly well distributed, with neither *patient(s)* nor *carrier(s)* discussed in the scope of groups and belonging. This subcategory accounts for one of the 13 S category collocates of *victim(s)*,

but one of only three S category collocates of *sufferer(s)*, making it quite significant in this manner. Both collocates are unique and will be taken in turn.

Social actor	GROUPS AND BELONGING S category collocates	% S category collocates	% overall collocates
<i>patient(s)</i>	-	0.0%	0.0%
<i>victim(s)</i>	embrace	7.7%	2.1%
<i>sufferer(s)</i>	fellow	33.3%	6.7%
<i>carrier(s)</i>	-	0.0%	0.0%

Table 8-29: USAS Category S subcategory ‘Groups and Belonging’ collocates by naming strategy

The first collocate that I will discuss is the one occurring with the higher frequency. *Embrace* collocates with *victim(s)* a total of 14 times, though two of these are false collocates (occurring beyond a sentence boundary) that have been miscalculated, and have therefore been discarded. In the 12 remaining lines, five cases use *embrace* in the physical sense (as in to hug or hold something/someone), six cases use *embrace* in the conceptual sense (as in to include something/someone without reservation), and in one case, it is unclear whether the physical or conceptual sense is indicated.

In each of the five cases of the physical sense, *embracing* is appraised positively; three of these are in connection with Princess Diana (as in line 64 below), who is singled out for praise throughout the HIV/AIDS corpus. However, none of conceptual cases of *embrace* are actually completed – each of them exists in a future state, often one which is unsure. For instance, though it is recommended that the church “should” *embrace* AIDS victims in 1987 (65), in 1988 a survey stating that there is “no way” to get the public committed to doing so is reported in the *Washington Post* (66). Even in 2006, influential people and institutions are negative appraised for being slow to *embrace* AIDS victims (67). In the one case that was unclear (shown in line 68), *embracing* victims of AIDS is listed with another item that apparently indicates an amazing feat of generosity – the ability to say “I love you” to homosexuals. This says quite a bit about the reporting of the time.

64. Seeing Diana in Harlem *embracing* a 2-year-old **AIDS victim** alone can make anyone respect and love her. (*New Orleans Times-Picayune*, 6 September 1997)
65. Cardinal Arns last month declared that, despite the church’s condemnation of homosexuality, it should embrace AIDS victims “as human beings.” (*Boston Globe*, 21 March 1987)
66. There’s no way to politically get 51 percent of the American public committed to *embracing AIDS victims*. (*Washington Post*, 13 July 1988)
67. Influential black people and black secular and religious institutions have been slow to embrace black **AIDS victims**, demand government aid or speak out. (*San Francisco Chronicle*, 5 June 2006)

68. He didn't change many people's minds, but he did touch hearts because of his generosity. He's not afraid to *embrace* **victims of AIDS**. He's not afraid to say to homosexuals, "I love you." (*USA Today*, 21 September 1987)

Figure 8-19: Sample concordance lines showing *embrace* collocating with *victim(s)*

*Fellow* was a welcome surprise in the list of collocates in the USAS S category, because it indicated that finally PWAs themselves would be those with voices in the texts, reflecting on the experiences of other PWAs through their own lens. This collocate is extremely infrequent – occurring just 12 times – reinforcing again how rare the PWA voice itself is within the corpus. And despite the camaraderie indicated by the adjective *fellow*, not all 12 instances are positive or community building. One narrative (reproduced in part in concordance line 69 below) is published in two separate newspapers, and in this, *fellow AIDS sufferers* are imagined as “laughing...looking up...with eyes that say, ‘You’ll be here soon enough’”. In another (line 70) *fellow H.I.V. sufferers* are said to feel betrayed by basketball player Magic Johnson, who is described as seeming “hurt” when the point is pressed. However, eight concordance lines do seem more positive, and in these, metaphors are often employed, indicating a shared conceptualization of the disease and of PWAs’ position in society. For instance, in line 71, *fellow AIDS sufferers* are rallied with metaphors of battle (“confront the disease”) and aggression (“shatter the silence”), empowering them to speak out. In line 72, a metaphor of merging (“becoming one”) is employed, reinforcing the conceptualization of *fellow* PWAs as a single unified entity, rather than fractured (and therefore less powerful) individual agents.

69. At times, I think my *fellow AIDS sufferers* are laughing at me, looking up from their beds with eyes that say, “You’ll be here soon enough.” (*New York Times*, 20 December 1992)
70. But if Magic feels betrayed by fellow players, so, too, do many *fellow H.I.V. sufferers* feel betrayed by him. They argue that by leaving competition, he has given legitimacy to what they view as irrational fears. Magic seemed hurt when that point was pressed in the interview. (*New York Times*, 19 November 1992)
71. El-Amin urges her *fellow AIDS sufferers* to confront the disease, shatter the silence. (*Washington Post*, 10 June 2001)
72. In the article and the interview, Mr. Brodkey talked of becoming one with his fellow AIDS sufferers, whose courage he praised. (*New York Times*, 17 June 1993)
73. But there is a hint of homophobia in his suggestion that the revelation was a blemish on his reputation. Nevertheless, his compassion for his *fellow AIDS sufferers* leads him to seek to understand the experiences of homosexuals, and he comes out firmly in support of tolerance and equal protection for gays, including the use of boycotts to secure equal rights. (*San Francisco Chronicle*, 27 June 1993)

Figure 8-20: Sample concordance lines showing ‘commemoration’ collocates of *victim(s)*

However, like all identities, there is not one clear-cut identity of PWA, and this comes to the fore in the concordance lines of *fellow*. In line 73 above, we see one HIV+ man who has not contracted the virus through sex with a man grappling to support and to understand *fellow* PWAs who have had an experience alien to his own (former) identity. Though these are his *fellow AIDS sufferers*, they are not identical in their circumstances, and it is for this reason he is able to show “compassion” for some who could be considered on his own level.

Unfortunately, this is the last collocate leading to concordance lines of PWA agency in this chapter. I shift now to the last USAS broad category to be analysed, Category X, where people surrounding people with AIDS firmly take back the reins.

#### 8.4.6 Category X: Psychological actions, states and processes

Distribution across USAS Category X is extremely variable, with one naming strategy (*sufferer*) showing no semantic preference for collocates here at all, and another (*carrier*) showing its strongest preference overall, with 28.6% of all collocates falling under the USAS X broad category. A good number of the collocates here are shared; indeed, *patient(s)*, *victim(s)* and *carrier(s)* share some permutation of three lemmas (*known*, *identify*, and *suspected*) and have only one unique collocate each. This could indicate that while preference for the USAS X category is varied, collocation with items from this semantic field activate similar discourses.

Naming strategy	X category collocates	% total collocates
<i>patient(s)</i>	<b>known</b> [LL=71.677, MI=3.287]	2.1%
<i>victim(s)</i>	<b>identify</b> [LL=85.315, MI=3.026], <b>known</b> [LL=83.794, MI=4.571], <b>innocent</b> [LL=64.732, MI=4.108], <b>suspected</b> [LL=49.084, MI=4.608]	10.4%
<i>sufferer(s)</i>	-	0.0%
<i>carrier(s)</i>	<b>suspected</b> [LL=179.529, MI=8.25], <b>identify</b> [LL=161.072, MI=5.558], <b>knowingly</b> [LL=100.544, MI=8.023], <b>known</b> [LL=79.287, MI=6.636]	28.6%

Table 8-30: Collocates from USAS Category X: psychological actions, states and processes

Each of the three collocates recurring across multiple naming strategies deals with the concept of [public] identification of people with HIV/AIDS. This is in contrast to the X category collocates of the naming strategies in Katrina chapter 5 (e.g. *urge*, *heed*), which did involve interaction with the people beyond hunting and identifying them. This serves as the starting point of analysis in this section.

The concept of a *known* AIDS/HIV *patient* is an early one, with all concordance lines arising from this collocation predating 1991. In only three instances, *known* refers to the risk factor or exposure to the disease, rather than to the PWA (see, for instance, line 74 below). Conversion from 'know' to *known* also obscures the agent—who 'knows' the status of these PWAs? In the *patient* concordance lines, *known* indicates the presence of counts, e.g. the number of *known AIDS patients* currently on record or being treated (e.g. lines 75 and 76 below). In one of these (64), *known* is used to contrast with the unknown number of Americans infected. *Victim(s)* has a not entirely dissimilar preference; while numbers do not occur as frequently in the 19 concordance lines of this collocation, *known* does denote that there are an unknown number of PWAs that cannot be included or tallied in current reports. For instance, in line 77, the number of *known* AIDS victims in schools is counted, though this does not (and cannot) include unknown cases. In line 78, the death of the first *known victim* is reported, but this phrasing does not discount the possibility of another, earlier fatal case. This is also the only concordance line of *known + victim(s)* from after 1993, seemingly belonging to the 'derision regarding the primitive actions of foreign governments' discourse previously discussed. In terms of semantic prosody, it is really *carrier(s)* which stands alone in this section. *Known carrier(s)* are consistently construed as dangerous (see line 79 below) or deviant (line 80) in each of the 11 concordance lines of this collocation. This is most closely related to the semantic prosody of the collocates itself; the top 'human' nominal collocates in the +1 position in the newspaper section of the COCA corpus are *terrorists*, *criminals*, and *pedophiles*. Giving PWAs the attribute of *known* also imbues them with a negative association with criminality and (sexual) deviance.

74. Dr. Groopman of Boston said he wasn't surprised by the possible oral exposure because he has documented cases of **AIDS patients** whose only *known exposure* to the virus has been through intimate oral contact with their partners. (*Wall Street Journal*, 21 May 1987)
75. The 1,231 dental professionals in the study included 1,104 dentists and 127 dental hygienists. About 15 percent reported having treated **known AIDS patients**. (*Washington Post*, 5 June 1987)
76. For every **known AIDS patient**, 15 other Americans -- 1.5 million or more -- are infected with the AIDS virus. (*Boston Globe*, 2 August 1989)
77. The boy is one of five **known AIDS victims** enrolled in California schools. (*San Francisco Chronicle*, 8 October 1985)
78. After [China's] first **known AIDS victim** died in 1985, hospital authorities burned his belongings and even the furniture he had used. Many health officials willfully ignored AIDS, calling it a "foreigners' disease". (*Newsweek*, 4 December 2000)
79. In addition, corrections officers in some parts of the country have begun recently wearing special protective clothing-much like surgical gear-designed to protect them from **known AIDS carriers**. (*Washington Post*, 17 May 1985)

80. In Columbia, S.C., Terry Lee Phillips, 27, a *known AIDS carrier* from Miami, is accused of intentionally trying to transmit AIDS to a woman he was charged with raping. (*USA Today*, 6 July 1987)

Figure 8-21: Sample concordance lines showing *known* collocating with various naming strategies

The next collocate – *suspected* – carries with it a very similar preference for criminality. The top nominal collocates in the +1 position in the newspaper section of the COCA corpus are: *terrorist, militants, insurgents, gunman, thief, bombers, hijackers, and killers*. *Suspected* collocates with *victim(s)* only 11 times, and two of these are false collocates arising from the phrase “**AIDS victims** or *suspected carriers*”. Discarding them lowers the collocation total below the minimum threshold, and this will be disregarded; however, both instances are covered in discussion of *carrier(s)*. Despite being much lower frequency overall, the collocation between *carrier(s)* and *suspected* results in 19 concordance lines, which are markedly more negative, in keeping with the more negative prosody of *carrier(s)* overall. In a series of three articles in late 1985, the *Chicago Sun-Times* describes the hunt for a prostitute who is *suspected* of being an AIDS carrier, and considered a risk of spreading the disease (see lines 81 and 82 for a sample). Sex work is often othered, and the attitude of the journalist or publication is exposed in the choice of phrasing in hair colour in an unusual resurrection of the dirty water metaphor: she is “dishwater blond” (line 82). *Suspected* AIDS carriers are also wanted in connection with aggravated assault (83) and sexual assault of a very young child (84). While reporting of criminal activity is clearly extremely salient for newspapers, the inclusion of *suspected* qualities of the accused seems incredibly dangerous, and in my opinion, reckless. It is likely this semantic preference for the domain of criminality that leads both *victim(s)* and *carrier(s)* to further collocate with *identify*, or the process of others seeking out and finding ‘suspects’ in the crime of HIV+ status.

81. Chicago police are searching for a 29-year-old North Side prostitute who is a *suspected AIDS carrier* and whose regular customers reportedly include a number of men from the north suburbs. (*Chicago Sun-Times*, 8 December 1985)
82. The woman fit the description of a white woman who is 5 foot 9, 110 pounds with dishwater blond hair and is a *suspected AIDS carrier*, Town Hall Capt. John Grentzner said. (*Chicago Sun-Times*, 23 December 1985)
83. Prosecutors in Albuquerque will ask a grand jury Thursday to indict a suspected AIDS carrier on a charge of aggravated assault - spitting on two police officers. (*USA Today*, 22 July 1987)
84. The case of a suspected AIDS carrier charged with sexually assaulting a 6-year-old girl has authorities in a quandary over the implications of crime and AIDS striking together. (*Denver Post*, 11 January 1990)

Figure 8-22: Sample concordance lines of the collocation between *suspected* and *carrier(s)*



Another unique collocate of *carrier(s) – knowingly* – also contributes to the discourse of criminality surrounding this naming strategy. Each of the 11 concordance lines are found in texts spanning from 1987 to 1991, making this phenomenon relatively restricted in time frame of the corpus. Of these, nine discuss the possibility of criminalizing the act of PWAs *knowingly* passing the virus onward (see line 85). The additional two construe PWAs as a financial burden (e.g. line 86), a discourse already explored at length in both the Katrina chapters and the HIV/AIDS chapters.

85. Ashcroft said the law permitted state prosecution of **AIDS carriers** who *knowingly* infect others. (*St. Louis Post-Dispatch*, 2 June 1988)

86. How, in the name of all that is rational, can we adopt a policy to admit **AIDS carriers** and *knowingly* give entry to a new source of contagion? The growing dollar costs of AIDS are financially overwhelming our hospitals. (*USA Today*, 2 August 1991)

**Figure 8-23: Sample concordance lines of the collocation between *knowingly* and *carrier(s)***

For the final collocate in this category, we remain in the semantic domain of crime and punishment, but turn from construal of guilt to its opposite: the concept of *innocent victim(s)*. This construction is tautological normally, and therefore implies that ‘guilty’ victims of AIDS also exist in contrast. *Innocent victim(s)* occurs 17 times in the corpus, though it is largely contested, with only 4 cases appearing in the wholly uncontested sense (e.g. line 87 below). Much more frequent is the metadiscussion of whether a distinction between *innocent* and “guilty” *victims* of AIDS can or should actually be made (88) or emotionally acted upon. In several cases, *innocent* appears as a foil for guilt, wherein parties deemed responsible for their conditions are distinguished and vilified (see lines 89 and 90 below). For instance, in line 89, the writer sympathizes with and prays for *innocent victims*, hoping that they are helped soon, but argues against taxpayer relief for “those who infect one another” through “sodomy or injecting themselves with needles”, construing this transmission group alone as the financial burden upon the nation. It is unclear how the writer proposes that “help” be divided among risk groups. The same strategy (whereby “so-called” *innocent victims* are contrasted to so-called ‘guilty’ *victims*) is also employed to indicate political shortcomings, as in line 91, where the concept of guilt or innocence is implicitly disparaged. In a juxtaposing strategy, *innocent* also appears to highlight similarity between groups, as in line 92, where children, haemophiliacs and gay men are all named as equally *innocent*.

87. In 1984, AIDS was even more mysterious and frightening than it is today. Certainly, Ryan [White] was an *innocent* and unwitting **victim of AIDS**. (*Chicago Sun-Times*, 16 January 1989)

88. For some, her accusation to Congress – “I did not do anything wrong, yet I am being made to suffer” - was a call to distinguish between innocent and guilty victims of AIDS and to allot compassion and contempt accordingly. (*St. Louis Post-Dispatch*, 14 December 1991)
89. I sympathize with and pray for the innocent victims of AIDS and pray that they can be helped soon. However, those who infect one another, be it by practicing sodomy or injecting themselves with needles, do this aware of consequences thereof; therefore, it ill behoves taxpayers supporting any relief for them. (*Chicago Sun-Times*, 17 November 1987)
90. Why are all **victims of AIDS** treated as *innocent* victims when so many are responsible for their condition by their own actions? (*Wall Street Journal*, 6 August 1992)
91. AIDS has also torn the wrapper off this country’s hatred of homosexuals. During the plague’s 11 years, two Presidents have pointedly reserved their sympathies for the so-called **innocent victims of AIDS** -- babies and hemophiliacs. (*New York Times*, 26 April 1992)
92. Children are innocent victims of AIDS, but so are hemophiliacs and gay men. (*San Francisco Chronicle*, 24 July 1989)

Figure 8-24: Sample concordance lines of the collocation between *innocent* and *victim(s)*

Generally speaking, while some instances of X category collocates (e.g. *known*) occur simply to contrast to the number of “unknown” PWAs in the world, the more frequent discourse activated through these collocations is that of criminality. Certain PWAs—namely, those who have contracted the disease through homosexual sex or intravenous drug use—are construed as criminal, in contrast to (young) haemophiliac PWAs—who are *innocents*.

## 8.5 Summary

In this chapter, four naming strategies of PWAs—*patient(s)*, *victim(s)*, *sufferer(s)* and *carrier(s)*—were explored through the scope of collocates from highly populated semantic categories. As is discussed in 8.5.1 below, analysis of collocates from these categories revealed extensions of patterns previously explored, as well as novel construal of PWAs. Methods were adapted from the sister Katrina chapter 5, and the efficacy of these changes will be reviewed below in section 8.5.2.

### 8.5.1 Of results

Categories which were identified as ‘salient’ for containing over 10% of collocates for any given naming strategy included: B (the body and the individual), G (government and public), H (architecture, housing and the home), S (social actions, states and processes), and X (psychological actions, states and processes). Several of these categories appeared for the first time in analysis, notably B.

Though B (the body and the individual) is an expected part of AIDS/HIV discourse, it is surprising that this did not surface in Chapter 7. Its appearance here is likely due to the extended window span and greater variety of naming strategies allowing for different meanings to emerge. Likewise, use of a naming strategy like *people with AIDS* might indicate a reluctance to medicalise this social group, in contrast to the constructions of *victims*, *patients*, etc. Use of specialised, field-specific terminology in the B category collocates, particularly co-occurring with *patient(s)*, alienates the experience of the PWA from the experience of the reader. Discussion of medical services and sites are framed in a discourse of shortage or denial of service, linked most frequently to *patient(s)* and *victim(s)*. *Patients occupying* hospital beds are described in terms of the *topoi* of numbers and strain, just as people affected by Katrina were construed in 5.5.3.1 as they evacuated the Gulf Coast and relocated to other areas. *Residential* care is positively appraised as an alternative to hospital stays for *patient(s)*, *victim(s)* and *sufferer(s)*, as these are privately funded and do not contribute to the discourse of governmental strain usually typifying PWAs (and PAKs).

*Discrimination* was found to haunt each of the naming strategies and PWA identities studied thus far, though concordance analysis showed that human agents are rarely the source of this. Most often, PWAs are construed as experiencing institutional *discrimination*, e.g. in employment, housing, healthcare, and education. This could be due in part to the implicit criminalizing prosodies created by collocation with terms such as *known* and *suspected*. Deviancy doubling occurred again in across all naming strategies save *carrier(s)* (though this group was vilified in its own right). AIDS/HIV *patient(s)*, *victim(s)*, and *sufferer(s)* were all associated with *homeless* people in parallel and compound structures (with *victims* preferring the parallel structure, indicating that the *homeless* are not considered *victims* of their circumstances), in addition to drug abusers, immigrants, the poor, the elderly, the mentally ill, victims of family violence, the unemployed, etc. Unlike the Katrina corpus examples, these are not temporarily homeless, but persistently *homeless*. *Patient(s)* are also grouped with another group of social pariahs: *lepers*; whereas *victim(s)* are associated with more *innocent* groups, such as *haemophiliacs*.

Analysis of collocates in the care and advocacy subcategory underscored steadfast focus on the agency of people around PWAs in the corpus, at the sacrifice of discourse empowering PWAs themselves being represented in the media (as in the Katrina section 5.5.3.4). The notable exception to this is the collocation between *fellow* and *sufferer(s)*,

whose concordance lines do show PWAs as complex agents who may either support or betray other *sufferer(s)*.

The most positive semantic prosody in this USAS broad category comes from an expected and unhappy place. Due to possible semantic sense of having died, *victim(s)* collocates with a subcategory of commemoration, in which they are more positively associated with family and loved ones than in concordance lines in which they are still living.

### **8.5.2 Of method**

Three major methodological changes have been made in response to shortcomings of Katrina Chapter 5, the companion to this chapter. These will be enumerated and evaluated in turn below.

The first, primary organization of collocates by USAS broad semantic category rather than part of speech category, has continued to prove advantageous. Please refer back to section 7.4.2 for a longer discussion on the ramifications of this adjustment.

The second alteration was maintaining extended portions of naming strategies in search strings, e.g. searching for “(AIDS|HIV|H.I.V.|A.I.D.S.|GRID|G.R.I.D.) {victim}” in the AIDS/HIV corpus versus “{victim}” in the Katrina corpus. This has been more effective in pinpointing the precise identities under investigation and automatically clearing some noise, e.g. “breast cancer *victims*” in the Katrina corpus. However, it has had the side effect of dramatically lowering the frequency of all naming strategies, and eliminating variations outside of this formula. Given the extremely inflated size of the corpus, and the greatly lengthened span of diachrony, however, this was a necessary variation.

In response to the decreased frequency of naming strategies, the minimum frequency of collocation has also been lowered to allow for a greater variety of meanings to emerge. This has been done with the belief that minimum frequency thresholds are largely arbitrary, and are best utilized as methods of downsampling data to ‘manageable’ proportions.

Experimentation with thresholds is a theme that will be carried forward through the discussion and development of the method of the next chapter, 9. This is a close study of a single identity group—gay people—and is paired with Katrina chapter 6 (on *refugees*). As this is the final analysis section in the thesis, Chapter 9 is the culmination of the

method crafted throughout the series of analyses thus far, and will be followed by Chapter 10 with an in-depth appraisal of its efficacy, limitations, and possible future applications.

## ***Chapter 9: Diachronic analysis of an identity group in the AIDS/HIV corpus***

### ***9.1 Introduction***

This, the final chapter in the AIDS/HIV analysis section, was designed as a counterpart to Chapter 6 of the Hurricane Katrina study; in other words, as a diachronic analysis.

However, the significantly extended sampling time frame of this corpus has necessitated some strict selection. It was my original intention to explore the construction of the “four H’s” of AIDS/HIV (Grmek et al., 1990, p. 31) here, though diachronic analysis of many of these groups has proven problematic. Two of the so-called four H’s are no longer considered core risk groups; the inclusion of Haitians was highly problematized in the early- to mid-1980s, and modern blood testing technology developed in the mid- to late-1980s removed much of the risk to haemophiliacs receiving transfusions in America. As a result, I have settled upon the most enduring risk group in AIDS/HIV discourse: gay men. Like previously in Katrina Chapter 6, here I emphasize a diachronic approach to corpus-based critical discourse analysis to address the following research question:

4. What evidence does corpus-based discourse analysis of the two corpora give of **changing or resistant discourses surrounding othered social groups over time?**
  - a. What kind of contextual or cultural changes result in shifting preference for terms or interest in discourses surrounding certain identities?

Throughout this chapter, I explore changes in the discursive construction of gay people, as highlighted by shifting preferences for semantic collocation and (in some cases), differing discourses constructed around consistent collocates. Though the corpus was not collected using this social group as a seed, the identity has proven extraordinarily salient, and pilot studies indicate that discourses around gay people have shown features of both risk society and moral panic discourses at various points in the diachrony of the corpus. In this chapter, I will take the opportunity to explore wider implications of methods of construing one social group via these (and other) discourses, while noting when this diverges from direct discussion of AIDS/HIV. Focus will be on collocates related to depiction of in- and out-groups; those unrelated to this will be only briefly described. As this is the final analysis chapter and the culmination of the

refinement of this type of semantic collocation as a method of downsampling, I will also make frequent reference to methodological choices and reflect upon the effects of implementing these.

The issue of downsampling particularly requires some explaining at the start of this chapter, particularly in order to account for articles which might mention HIV/AIDS and gay men at different points and not necessarily link the two together.

An overview of the method (including adjustments made since the sister Katrina chapter 6) can be found in section 9.2 below. In section 9.3, I provide analysis of each collocate tagged into USAS broad categories representing a high proportion of all collocates. The most salient semantic categories for this social group are A: general and abstract terms (section 9.3.1), G: government and public (9.3.2), and S: social actions, states and processes (9.3.3). In the summary section (9.4), I will provide an overview of results (9.4.1) and critique the method used in this chapter (9.4.2).

## **9.2 Methodology**

In the following section, I will describe the method by which I have targeted search terms (9.2.1) and identified key semantic categories of collocates (9.2.2). Where applicable, I will flag any adjustments in method made between Katrina Chapter 6 and this, its sister analysis.

### **9.2.1 Targeting search terms**

A search for node words denoting homosexuality and bisexuality—including some slurs—was carried out in the AIDS/HIV corpus using the following search string:

```
(gay|gays|homosexual|homosexuals|homo|homos|bisexual|bisexuals|bi|fag|fags|faggot|faggots)
```

This returned 121,242 matches in 30,254 different texts. Frequency of occurrence for these search terms can be found in Table 9-1 below.

No.	Search result	No. of occurrences	Percent of all results for search string
1	gay	81,726	67.4%
2	homosexual	13,160	10.9%
3	gays	10,210	8.4%
4	homosexuals	10,107	8.3%
5	bisexual	4,418	3.6%
6	bisexuals	508	0.4%
7	homo	248	0.2%
8	faggot	219	0.2%
9	fag	186	0.2%
10	fags	129	0.1%
11	faggots	122	0.1%
12	bi	112	0.1%
13	homos	97	0.1%

**Table 9-1: Frequency of node words associated with homosexuality, shown in raw hits and as a percentage of all of the search results**

As this is a study of the changes in the most frequent naming strategies over time, only the most common terms will be used for further analysis. The top four naming strategies, *gay*, *homosexual*, *gays*, and *homosexuals*, comprise over 95% of all of the matches in this set, and are the most closely semantically related. Therefore, these will form the basis of the diachronic study moving forward.

A simple case-insensitive CQPweb search for the following was carried out:

(gay|gays|homosexual|homosexuals)

This returned 115,203 matches in 29,745 texts (out of 161,144,924 words in 166,576 texts), occurring 714.9 times per million words. However, approximately 6% of a random thin of these results included initial capitalization, indicating the node word's use as part of a proper noun phrase, e.g. in place names ("the Gay Men's Health Crisis Center in New York City"), association/group names ("Penn State's Coalition of Lesbian, Gay, Bisexual and Transgendered Graduate Students") and personal names ("Ms. Gay"). To allow for a more accurate focus on the representation of *gay people*—the aim of this chapter—a case-sensitive search for the same term has been used. This results in the same loss of sentence-initial occurrences of the node as discussed earlier in Section 4.2.2, which remains an acceptable (if unfortunate) forfeit for the sake of enhanced precision of results in line with research aims.

The case-sensitive search returned 100,997 matches in 26,984 texts. This search string will be referred to hereon as 'the gay set'. Distribution and dispersion information per decade can be in Table 9-2 below.



Decade	Words in category	Hits in category	Dispersion in texts	Frequency/million words in category
1980s	35,286,877	29,133	8,410 out of 41,548	825.6
1990s	77,681,146	48,590	12,847 out of 81,587	625.51
2000s	48,176,901	23,274	5,727 out of 43,441	483.09
Total:	161,144,924	100,997	26,984 out of 166,576	626.75

Table 9-2: Frequency of node words associated with homosexuality, shown by distribution and dispersion over three decades

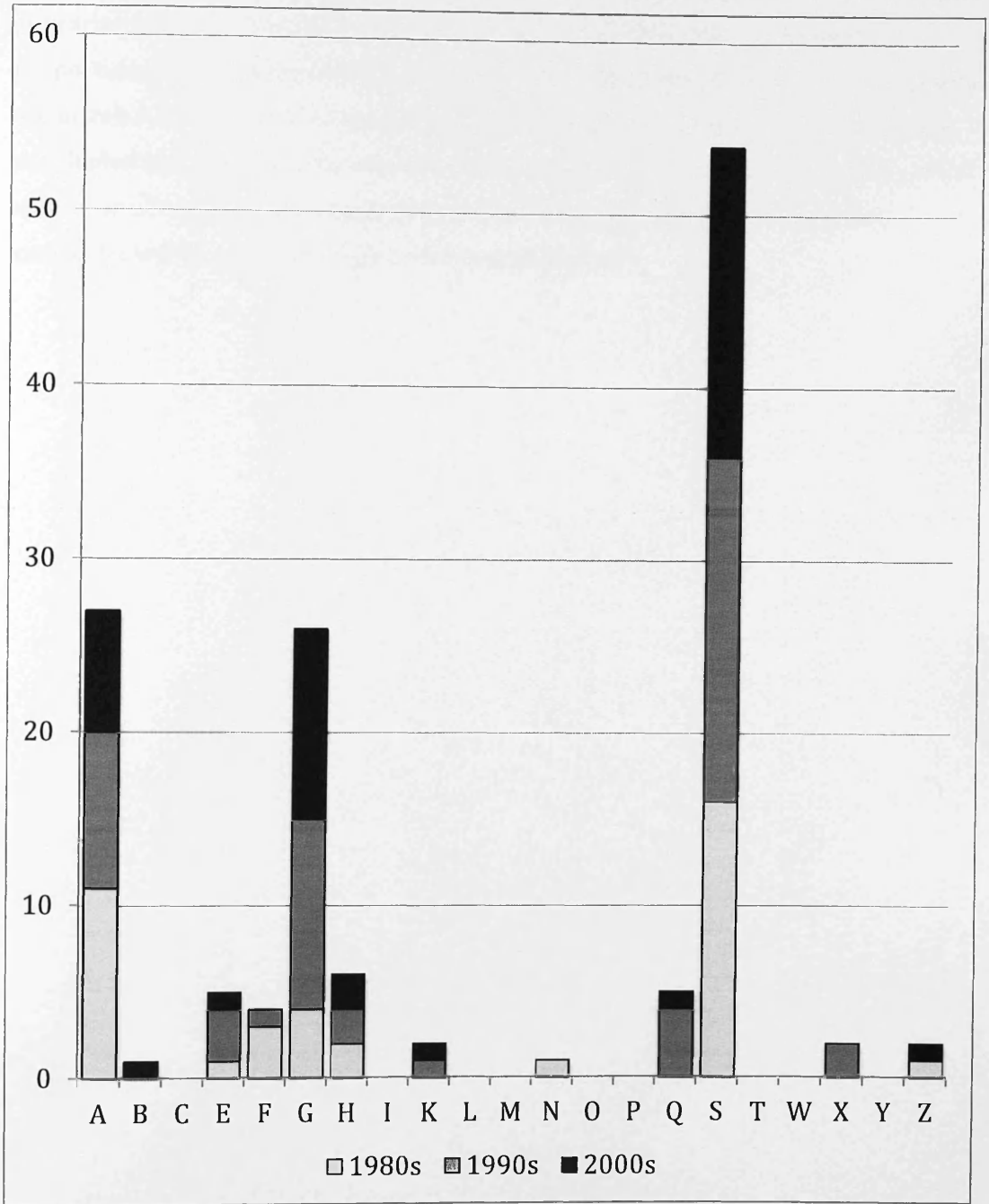
On the basis of declining occurrence of the most frequent node words associated with homosexuality, it appears that the topic (and the social group) was discussed in news texts less over time in conjunction with HIV/AIDS. The nodes appear nearly twice as frequently in texts from the 1980s as they do in texts from the 2000s, with a nearly constant rate of change between. Such a dramatic decrease in frequency could indicate a decline in social salience, and an evolution of the moral panic beyond this social group. In this chapter, I explore discourses of gay people throughout the 1980s, 1990s, and 2000s, to search for traces of shift in construal accompanying this shift in frequency.

### 9.1.1 Identifying key semantic categories of collocates

As in chapters 5, 6, and 8, collocates of the node words denoting social group have been generated on the basis of a +/-3 span, with thresholds of  $MI > 3$  and  $LL > 10.83$ .

**Adjustment in methodology:** Some alterations in collocation settings have been made in response to fundamental differences between the Katrina and AIDS/HIV corpora, and the consequent nature of diachronic analyses within these resources. Sheer size has made an extraordinary difference, with the gay set search string returning over 100,000 matches, compared to approximately 2,000 results from *refugee* in the Katrina corpus. Due to this extreme frequency difference, I have raised the minimum frequency of collocation to 50 to ensure that all MI scores over 3 guarantee LL scores exceeding 10.83. Another major factor to account for is the span of diachrony. In Chapter 6, I analysed the changing construal of *refugee* over one month; in this chapter, the timespan is nearly three decades. Therefore, unlike previous chapters, collocates for the search terms have been generated independently for each decade, and will be referred to both by decade and collectively.

As has become standard, each decades' collocates were run through WMatrix in parallel to obtain semantic tags, and these were manually disambiguated in the usual way. Visual representation of collocates by USAS broad semantic category can be found in Figure 9-1 below.



**Figure 9-1: Count of collocates in each semantic category, stacked by time**

In looking at this figure, it becomes clear that categories A (General and abstract terms), G (Government and public), and S (Social actions, states and processes) dominate the collocates in nearly every decade. However, this visualization does not account for those items collocating with the node words in multiple decades (effectively double-counting these items). Table 9-3 below shows the frequency of collocates occurring in each USAS category as a percentage of all collocates in the same decade, but also indicates an

overall percentage across the whole corpus represented by the category when collocates appearing in multiple decades have been de-duplicated. Once more, the highest proportion preference for each decade is in boldface; preferences of semantic categories are in italics. Those semantic categories not containing collocates for the search string are shaded out. As before, the names of semantic categories represented by >10% of the collocates of any given naming strategy are in all capitals—the collocates of these categories will be analysed in full in the following chapter.

	1980s	1990s	2000s	De-duplicated collocates
A GENERAL AND ABSTRACT TERMS	28.2%	17.0%	16.3%	19.5%
B the body and the individual	0.0%	0.0%	2.3%	1.2%
C arts and crafts	0.0%	0.0%	0.0%	0.0%
E emotion	2.6%	5.7%	2.3%	4.9%
F food and farming	7.7%	1.9%	0.0%	3.7%
G GOVERNMENT AND PUBLIC	10.3%	20.8%	25.6%	22.0%
H architecture, housing and the home	5.1%	3.8%	4.7%	2.4%
I money and commerce in industry	0.0%	0.0%	0.0%	0.0%
K entertainment, sports and games	0.0%	1.9%	2.3%	1.2%
L life and living things	0.0%	0.0%	0.0%	0.0%
M movement, location, travel and transport	0.0%	0.0%	0.0%	0.0%
N numbers and measurement	2.6%	0.0%	0.0%	1.2%
O substances, materials, objects & equipment	0.0%	0.0%	0.0%	0.0%
P education	0.0%	0.0%	0.0%	0.0%
Q language and communication	0.0%	7.5%	2.3%	4.9%
S SOCIAL ACTIONS, STATES AND PROCESSES	41.0%	37.7%	41.9%	36.6%
T time	0.0%	0.0%	0.0%	0.0%
W world and environment	0.0%	0.0%	0.0%	0.0%
X psychological actions, states & processes	0.0%	3.8%	0.0%	2.4%
Y science and technology	0.0%	0.0%	0.0%	0.0%
Z names and grammar	2.6%	0.0%	2.3%	0.0%

Table 9-3: Collocates of the gay set by decade, categorised into USAS semantic categories, and expressed as a percentage of the overall number of collocates per decade

The table above does give an indication of the overall patterns of the changing landscape of semantic field associated with the social group, but the top three categories (A, G, and S, as above) are relatively static in represented percentage over time. These three categories are clearly the most salient: the A category contains 19.5% of all de-duplicated collocates; the G category contains 22.0%; and the S contains 36.6% of all collocates, dominating the reportage on this group. No other category nears 20%, or even exceeds 5%. As such, categories A, G, and S form the basis of analysis in this chapter.

The shifting diachronic frequency of collocates in semantic categories does not speak very loudly of the data; this is still best utilized as a 'pointer' indicating the most salient semantic fields under analysis. The most telling aspect will be the shift in collocates themselves. To visualize this, Venn diagrams will be incorporated to illustrate the logical relationship between items (here: all collocates, and collocates in the A, G, and S broad semantic categories) in the sets (here: decades). A Venn diagram of all collocates of *gay/gays/homosexual/homosexuals* by decade appears in Figure 9-2 below. These are listed alphabetically, as differences in statistical values by decade prevent a descending list by confidence measure (LL value) of the type included in previous chapters. Full tables of all collocates with MI and LL values, classified by USAS semantic category and decade, can be found in Appendix M (1980s), Appendix N (1990s), and Appendix O (2000s).

The Venn diagram in Figure 9-2 below does show a clear progression in discourse, with many words unique to each decade or occurring in two adjacent decades, but with only one word shared between the 1980s and 2000s excluding the 1990s. There is also a clear core set of c-collocates (Gabrielatos & Baker, 2008) depicted in the middle of the diagram, which do not change over time. In considering those collocates occurring in two or three decades, I am mindful of the danger in assuming that a word always activates the same discourse; for instance, one word might be used in an uncontested way early in the corpus, and again later on to contest or critique a concept. Therefore, I will again be performing close qualitative analyses to get a better understanding of the senses of these collocates, sensitive to the possibility of changes over time.

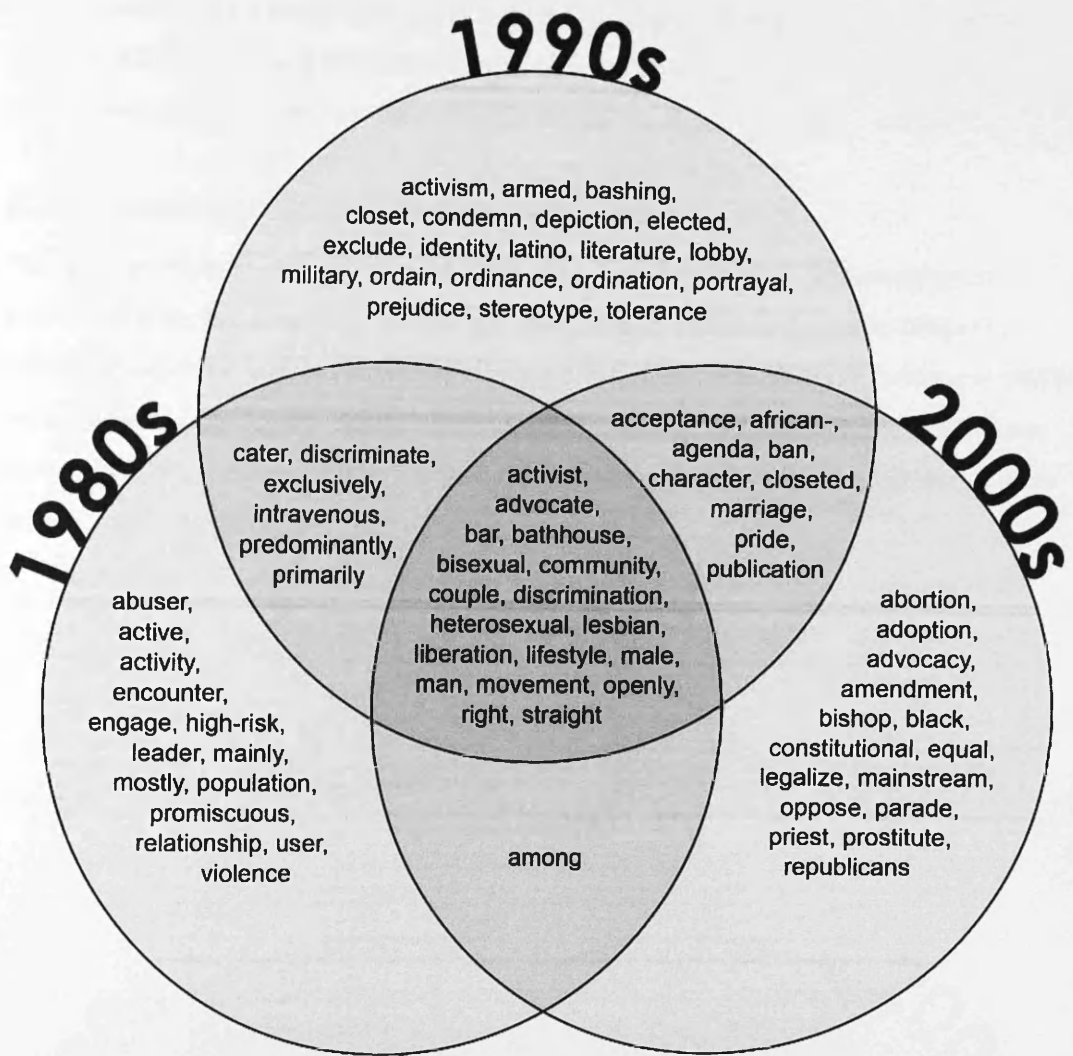


Figure 9-2: Venn diagram of collocates from all USAS broad semantic categories for *gay/gays/homosexual/homosexuals* in the AIDS/HIV corpus depicted over time

Even de-duplicated, the relatively large number (83) of total collocates under analysis makes such a large, generalized view of these quite convoluted; it is difficult to spot patterns in the data over time without further categorization. Moving forward, smaller Venn diagrams containing collocates from specific USAS semantic categories will be incorporated to illustrate finer details. Diagrams and discussion on collocates from the most frequent USAS semantic categories (A, G, and S) comprise the following section.

### 9.3 Analysis

The following section contains analysis of all collocates belonging to the three USAS broad semantic categories containing a 'substantial' proportion (exceeding 10%) of all collocates of the search string. The categories to be covered are USAS Category A:

general and abstract terms (section 9.3.1), USAS Category G: government and public (section 9.3.2), and USAS Category S: social actions, states and processes (section 9.3.3). Within each semantic section, discussion is organized diachronically.

### 9.3.1 Category A: General and abstract terms

The nature of Category A, containing “general and abstract terms”, has resulted in collocates from this category occupying a more diverse range of semantic fields than would be expected or experienced in viewing other categories. As such, collocates will be arranged and discussed in subcategories where relevant. Analysis will generally move chronologically, culminating in discussion of ‘c-collocates’ (Gabrielatos & Baker, 2008) remaining static across the subcorpora.

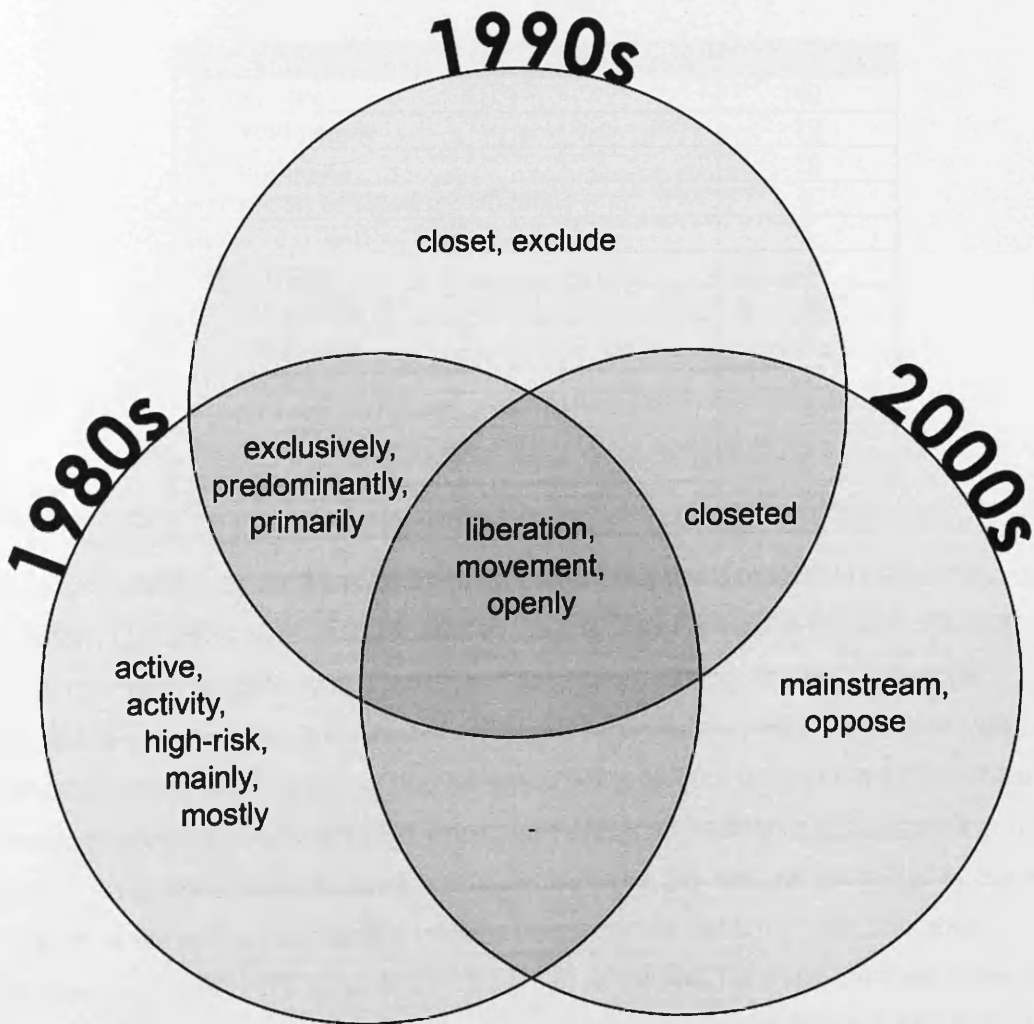


Figure 9-3: Venn diagram of USAS Category A collocates for *gay/gays/homosexual/homosexuals* depicted over time

Instances of collocation over time can be viewed in Figure 9-3 above. Note that the most populous subcategory is in the 1980s, with no collocates occurring in the 1980s and 2000s but not the 1990s.

### 9.3.1.1 USAS A category collocates unique to the 1980s

I will begin with collocates unique to the 1980s subcorpus (*active, activity, high-risk, mainly, mostly*). It is immediately clear that journalists explicitly negotiate the concept of risk during this period. The term *high-risk* collocates with the gay set 68 times, occurring in descriptions of risk groups which have broadly been previously discussed. In the texts from the 1980s, drug users were the most frequently associated risk group, followed by bisexuals and haemophiliacs, and an assortment of other, lower-frequency groups (see Table 9-4 below).

Associated risk group	Frequency
Drug users	42
Bisexual people	12
Hemophiliacs	10
Recipients of blood transfusions	5
Sexual partners of those infected	5
Sex workers	3
Haitian people	2
African people	1
Minorities	1
People with many sexual partners	1
People with other STDs	1
Health care workers	1

**Table 9-4: Social groups appearing in concordance lines where the gay set collocates with high-risk**

A sample of this—by far the most frequent pattern—appears in sample concordance line 1 below. The 1980s were characterized by description of groups at risk and affected, leading both to heightened awareness and heightened anxiety. In line 2 below, the spread of fear and panic is illustrated as “people begin to fear [HIV] irrationally”, with those perceived to be outside of the risk group being derided for seeking advice on their status. However, by the late 1980s, a discourse of fear around this particular risk group (gay men) is still present. In line 3, a group of high-risk gay men are described as having “harbored the AIDS virus silently” while testing negative on blood tests. This shows elements of both a moral panic (a deviant group possessing the power to undermine the stability of society) and risk society (medical tests and modern technology had failed to identify the risk and the illness in time to stop transmission to others). Here, the PWAs are construed as possible perpetrators of a crime; the fact that they had “harbored” the



virus (“silently”) is “disturbing”. In the COCA corpus (in the same +/3 span, MI at least 3, in the newspapers section), the top resulting things that are “harbored” are terrorists, [ill, bad, bitter, hateful] feelings, and illegal [immigrants]. This contributes to a discourse of social disorder surrounding the high-risk group.

1. *High-risk* groups included **homosexual** or bisexual men, intravenous drug abusers, hemophiliacs, female prostitutes, heterosexuals from Haiti and some central African countries where heterosexual spread of the virus occurs frequently, and people who received many blood transfusions between 1983 and 1985, when routine testing of blood for HIV began. (Washington Post, 22 April 1988)
2. School of Medicine calls *afrAIDS* a form of mass hysteria. “When something like this is so much in the news, people begin to fear it irrationally.” Prevalent among *high-risk gay* men since the early '80s, AIDS anxiety has only recently spread to heterosexuals, researchers say. “Over 75 percent of the calls we get now are from heterosexuals,” says Eric Engstrom of the Minnesota AIDS project. (USA Today, 26 June 1987)
3. In a new study that some scientists say contains the most disturbing AIDS news in a long time, a team of researchers from California has documented that nearly one-quarter of a group of 133 *high-risk gay* men harbored the AIDS virus silently in their systems for as long as three years, all the while testing negative on standard blood tests to detect antibodies to the virus. (Boston Globe, 1 June 1989)

Figure 9-4: Sample concordance lines where the gay set collocates with *high-risk*

Membership within the *high-risk* group is largely determined through being an ‘*active* homosexual’ or participating in ‘homosexual *activities*’. As these are semantically (and etymologically) related collocates, they will be discussed in turn. Firstly: *active* collocates with the gay set 162 times. Manual coding of concordance lines shows clear patterns in the data. In 81 instances, the bi-grams *active gay/gays/homosexual/homosexuals* are actually tri-grams, additionally modified by ‘sexually’ to be explicitly clear about the nature of the active state (see line 4 below). This state of being *active* is linked plainly to risk. Even in the 37 concordance lines where ‘sexually’ does not modify *active*, but this is implied (see, for instance, illustrative examples 5 and 6 below), other indicators of sexuality occur in the extended context (e.g. “sexual activities” and “bisexuality”). In these lines, elements of risk society and moral panic also appear: unaware in the 1980s of exactly how the virus spread, gay people are described as abstaining altogether (line 5), an indicator of fear arising from risk society, and blood banks refused donations from *active homosexuals* (line 6), a symptom of moral panic. Not all concordance lines of *active* are related to sexual practices, however. In 18 concordance lines, social actors are *active* politically (as in line 8 below); in an additional 13, they are *active* in gay rights issues or the rights movement. Some social actors are also construed seven times as being *active*

in the 'gay community' or taking part in other community-building activities, such as promoting gay-friendly travel (see line 9 for an example). However, sexually *active* is by far the majority pattern, and one that ties in with the next collocate to be discussed: *activity*.

4. Sexually active homosexuals are prone to a host of diseases: syphilis, gonorrhoea, genital herpes, hepatitis, amebiasis (one of the most common diseases in what doctors call the "gay bowel syndrome") and infections caused by fungi and protozoa usually seen only in the tropics. (*New York Times*, June 2 1983)
5. There is such division over which sexual practices are safe that many previously *active homosexuals* are no longer engaging in sexual activities at all. (*New York Times*, 17 December 1985)
6. AIDS is of course a disease that has hitherto been associated mostly with homosexuals, but it has gotten into the general population via bisexuality, intravenous drug use, and blood banks (which are forbidden to ask blood donors if they are *active homosexuals*). (*Chicago Sun-Times*, 13 July 1985)
7. But in New York it took a group of politically active gays to go into the mayor's office and start things rolling. (*Washington Post*, 19 May 1987)
8. A New York geneticist *active in the gay rights movement* has been named by President Reagan to a 13-member AIDS panel, administration sources said yesterday. (*San Francisco Chronicle*, 21 July 1987)
9. *Active in the gay and lesbian community*, Meagher openly acknowledged his sexual preference to the Judicial Nominating Council. (*Boston Globe*, 16 February 1989)

Figure 9-5: Sample concordance lines where the gay set collocates with *active*

*Activity* is the only noun present in the collocates of the 1980s. This is somewhat unexpected, given the flexibility of the words *homosexual* and *gay* as either common singular nouns or adjectives. Its presence is doubly surprising given the oblique aspect of the term; "homosexual activities" is euphemistic, polite, and even vague. Explicitly including a term such as "[unprotected] homosexual sex" in reportage, for instance, would be more socially constructive in adequately describing risk activities but this has been avoided, leading to some additional modifiers, detailed below.

This collocation occurs 309 times, of which the overwhelming majority (286, 92.56%) is with *homosexual*. *Activity* is rarely modified directly by *gay*; this collocation occurs 16 times (5.18%), but only three cases of *gay* being used as the sole attribute adjective appear. A table of n-grams where *gay* collocates with *activity* can be found below. Unlike *homosexual*, *gay activities* are not necessarily to do with sexual practices. The most frequent phraseology, *gay sexual activity*, explicitly marks this aspect. N-grams in positions 4, 5, and 6 in Table 9-5 also discuss *sexual activity/ies* being practiced either at *gay* 'places' or within the community. Positions 7 and 8 are instances where *activities*

does not denote sexual activities at all, but rather more general goings-on in the community.

	n-gram	Frequency
1.	gay sexual activity	6
2.	gay activity	3
3.	gay or bisexual activity	1
4.	sexual activity at gay clubs and bathhouses	1
5.	wide-open sexual activities practiced in gay bathhouses	1
6.	sexual activity within the gay community	1
7.	gay rights activities	2
8.	gay community activity/ies	1

Table 9-5: n-grams resulting from the collocation of *gay* with *activity*

The relatively high occurrence of phrases to do with sexuality compared to the low frequency of other activities is related to differences in frequency of the search nodes. The plural *homosexuals* collocates with *activities* seven times, five of which are modified to clarify *sexual activities*, and two of which are unrelated to sexuality. *Gays* and *activities* do not collocate statistically significantly. This distribution shows a marked preference for the more medical, formal term *homosexual* in conjunction with *activities*. There is evidence that *homosexual activities* are necessarily sexual ones, whereas *gay activities* might involve social processes not having to do with sex.

Analysis of the concordance lines allows for a more detailed view of media depiction of *activities* associated with homosexuals. In the 1980s, these seemed primarily to do with risk of various descriptions. Concordance lines 10, 11, and 12 below show how risk of contraction of AIDS/HIV is linked directly to homosexual activities. Concordance line 10 includes heterosexual activities here, but qualifies this with “certain types”, leaving it ambiguous as to whether this is meant to indicate anal sex of a homosexual or heterosexual nature, or unprotected sex of any kind, between any partners. Shared use of IV drug equipment is also mentioned here. Statistics are included in concordance line 11, contributing to the *topos* of numbers discussed in earlier chapters. It is interesting to contrast the combined figure (of an alarming, moral panic-inducing 89%) here—taking into account both homosexual activity and intravenous drug use—and the separate figures provided in concordance line 12, which distinguishes these risks from one another, and therefore presents the statistic of contraction associated with homosexual activities (74%) as a lower, more precise, and slightly less alarming one. The discussion of the spread of the virus amongst inmates, who are seemingly guilty of deviant behaviour, and have partaken in deviant activities, is also mentioned in this concordance

line, but as no research was available at the time, very little data is given. *Homosexual activities* are construed as moral or ethical risks in lines 13, 14, and 15 below. In line 13, a ban on government spending for educational materials encouraging *homosexual activities* is upheld; here, it seems that tax money cannot be justifiably allocated to educational materials regarding immoral behaviour. Concordance line 14 displays this sentiment more strongly, incorporating the viewpoint of a religious leader who constructs homosexual activities as both a health risk and a moral “perversion”. The letter to the editor in concordance line 15 links both homosexual activity and casual sex to health risk, using the word *vice* to indicate an ethical shortcoming.

10. The AIDS virus infects persons who expose themselves to known risk behavior, such as certain types of **homosexual and heterosexual activities** or sharing intravenous drug equipment. (*Newsday*, 3 February 1987)
11. In America, the political, economic and emotional responses to the spread of AIDS have turned on one central fact: about 89 percent of those who have been diagnosed as having AIDS contracted it through **homosexual activity** or intravenous drug use. (*Washington Post*, 17 January 1989)
12. AIDS is most readily transmitted by unprotected **homosexual activity**. According to the Centers for Disease Control in Atlanta, 74 percent of all AIDS infections are transmitted in this manner and 17 percent by the sharing of needles by intravenous drug users. Unprotected homosexual activity in prison, although undocumented, is believed to be relatively common. (*New York Times*, 17 December 1989)
13. Sen. Jesse Helms did win approval of continuing a ban on the use of public money for educational materials that “promote or encourage directly homosexual activities.” (*Washington Post*, 3 May 1988)
14. Caffarra, who is regarded as close to Pope John Paul II, called **homosexual activity** perverse and said that the use of condoms to stem the spread of AIDS was not morally justified. (*St. Louis Post-Dispatch*, 15 November 1989)
15. To admit that abstinence from homosexual activity - not sodomy with condoms - is the most effective way to avoid AIDS would force Mayor Koch to admit also that casual sex is a perilous vice. (*New York Times*, 23 November 1987)

**Figure 9-6: Sample concordance lines where *homosexual* collocates with *activity***

Both of the remaining collocates unique to the 1980s—*mainly* and *mostly*—also relate to risk, this time via the *topos* of numbers, discussing the statistics of infection rates in relationship to the *homosexual* risk group. This linguistic feature serves two purposes: primarily, it reports facts relating to infection and disease; secondarily, it defines the social group at risk, one to which the average reader does not belong. This is another feature of social sequestering, previously discussed in Chapters 4.3.5.1, 6.3.4, and 8.4.1.1. A similar pattern is revealed in analysis of USAS A category collocates shared between the 1980s and 1990s, below.

### 9.3.1.2 USAS A category collocates shared between the 1980s and 1990s

The *topos* of numbers and negotiation of risk also appear as steady semantic field collocating with *gay/gays/homosexual/homosexuals* through the early-to-mid portions of the AIDS/HIV corpus. *Exclusively*, *predominantly*, and *primarily* all appear as (the only) collocates both in the 1980s and 1990s, and could serve the function of describing and conveying the indicators of risk. However, their appearance over time has allowed for inspection of possible change.

Closer inspection of the concordance lines arising from these collocations shows that it is really *primarily* that functions as a risk marker, both in the 1980s and 1990s. In 1980s reporting, as in concordance lines 16 and 17 below, it is reported that [male] homosexuals (and bisexuals – noting the lack of adoption of the term ‘men who have sex with men’) are *primarily* affected. However, in the 1990s, this begins to become problematized. In lines 18 and 19 below, we see that these skewed perceptions are negatively represented. Note the re-emergence of a water metaphor in line 18, despite the contrary argumentation problematizing the placement of “homosexuals and largely black or Hispanic drug addicts” as “outsiders”. In line 19, the conception of these groups as still being *primarily* affected is labelled a “myth”. Contrary to my initial impression, however, the remaining two collocates sharing this time period and grammatical class do not create similar semantic impressions.

16. The cause of the underlying immune disorder, which came to public attention a year ago, has not been determined. It had been reported earlier that the disorder affected primarily male homosexuals and people who use drugs intravenously. (*New York Times*, 9 July 1982)
17. For unknown reasons, Kaposi's sarcoma in AIDS patients in the United States has struck primarily among homosexuals and bisexuals, who account for 73 percent of the AIDS cases. (*New York Times*, 16 June 1986)
18. Despite the horrifying statistics, people with AIDS have remained largely invisible, even in television movies that exploit just about every other disease known to the world at large. The reason has been obvious: The first waves of AIDS cases primarily involved homosexuals and largely black or Hispanic drug addicts, decidedly outsiders in television's traditional pantheon of acceptable social images. (*New York Times*, 11 April 1991)
19. The new studies also underscore the myths about AIDS that still prevail among many Americans and Europeans, who see the epidemic as affecting primarily gay and bisexual men and users of injection drugs. (*San Francisco Chronicle*, 5 July 1996)

Figure 9-7: Sample concordance lines where *primarily* collocates with the node words

*Predominantly* also shows some shift in semantic preference over time. In the 1980s this refers to risk groups such as “male homosexuals and drug addicts”, and can be wielded

as either a force of social sequestering and comfort (e.g. the disease *predominantly* affects a minority group of which the average reader is not a part) or negated or called into doubt as a form of risk society (e.g. while the disease *predominantly* affects one group, this is not to say that it will not soon spread to others, as in line 20 below). In the 1990s, *predominantly* more often refers to a group of people (as in line 21), most of whom are gay. This discourse does not relate to risk but to proportions of identity. By the 2000s, this pattern has solidified, and *predominantly* refers to geographical locations populated by a group which has a high proportion of gay people (line 22).

20. Researchers have said they fear that the disease, which has *predominantly affected* male **homosexuals** and drug addicts, is spreading to the general population. (*Washington Post*, 25 October 1985)
21. Later Tsongas met with a *predominantly gay* audience at a local restaurant to talk about health care. (*Washington Post*, 16 March 1992)
22. The beach at Cherry Grove on Fire Island -- billed as America's oldest *predominantly gay* and lesbian settlement, dating back to the 1920's -- is a place where gays and lesbians have long felt at home. (*New York Times*, 26 August 2001)

Figure 9-8: Sample concordance lines where *predominantly* collocates with the node words

*Exclusively* is not used as often to negotiate proportions of risk groups, likely due to its semantic force – it would have been unreasonable and irresponsible throughout the 1980s and 1990s for doctors or journalists to report that one group was at risk, at the exclusion of all others. Rather, this term describes proportions of human behaviour categorized under sexuality, as in lines 23 and 24 below. Social actors (some who have acquired AIDS) self-identify or are identified as being *exclusively* homosexual, in contrast to bisexual (line 23) or on some other position on the Kinsey scale (line 24). While the use of this word might evoke discourses of risk—particularly in terms of sexuality—it is not explicitly doing so.

23. Among white males who acquired AIDS through homosexual contact, they said, 87 percent reported they were exclusively homosexual and 13 said they were bisexual. (*Boston Globe*, 17 May 1987)
24. Its introductory chapter updates Kinsey's 0-6 scale, which denotes the range of human behavior from exclusively heterosexual to exclusively homosexual, to place more emphasis on sexual and romantic feelings. (*Boston Globe*, 16 December 1990)

Figure 9-9: Sample concordance lines where *predominantly* collocates with the node words

While it initially appeared that the 1980s and 1980s would be dominated by a discourse of counting and categorising risk groups, in fact much of the discourse arising from these collocates has had to do with description of identity rather than definition of risk.

### 9.3.1.3 USAS A category collocates unique to the 1990s

There are only two A Category collocates unique to the 1990s: *exclude* and *closet*.

First, for *exclude*, I have coded each concordance line of the concordance between *exclude* and the gay set to ascertain common actions and roles which these social actors are being prevented from completing or occupying. The resulting frequency counts can be found in Table 9-6 below, ranked by order; of the 63 concordance lines, two were categorized N/A as the *excluding* referred to non-human social actors (e.g. when interest groups “exclude ‘gay’ from their names”, *Boston Globe*, 26 April 1992).

being <i>excluded</i> from...	Frequency
military service	13
teaching/scouting	10
immigrating to America	7
anti-discrimination legislation	6
marching in a St. Patrick's Day parade	5
religious affiliation/service	5
politics	5
jobs	3
fostering/adopting children	3
donating blood	2
clinical studies/treatment	2
marriage	1
N/A	2

Table 9-6: Frequency of actions and roles in concordance lines of *exclude*

Much of the discourse of exclusion here deals with *excluding* in a legal sense, for instance, excluding gay people from military service (13 instances), immigration (7) or anti-discrimination legislation (6), though others are more closely related to religious affiliation (5) or social exclusion, e.g. from marching in a St. Patrick's Day parade (5), participating in politics (5), taking jobs (3) fostering/adopting children (4) and marrying (1). Legal exclusion (exemplified in line 25-27) tends to be construed as risk-based (asserting that gay people are more likely carriers of HIV and therefore less attractive as immigrants and soldiers), and this sort of exclusion is problematized (and often eventually overturned) in the 1990s. The more insidious exclusion is the religious/social sort, which indicates underlying assumptions about difference and deviancy of this identity group that would lead to their rejection from in-groups. The conception of heterosexual people as being somehow more pious (line 28) or moral (line 29) reinforces existing moral panics as a method of delineating in- and out-group membership, rather than drawing upon risk society rhetoric. This is also the case for the next collocate, related to secrecy and perceived deviancy.

25. Neither the Defense Department nor the plaintiff, a gay midshipman who sued the United States Naval Academy over discrimination against homosexuals, raised the issue of AIDS. But Judge Gasch said the Government's policy of excluding homosexuals "is rational in that it is directed, in part, at preventing those who are at the greatest risk of dying of AIDS from serving." (*New York Times*, 10 December 1991)
26. Lawyers handling such cases estimate that since the United States in 1990 repealed a law that excluded all homosexuals from immigration, about 20 homosexuals have applied for sanctuary based on claims of gay persecution abroad. (*New Orleans Times-Picayune*, 17 August 1993)
27. Instead of excluding homosexuals from protected status, Question 1 asks whether civil-rights safeguards should be conferred solely on the basis of certain characteristics. Sexual orientation is pointedly excluded from that list. (*New York Times*, 5 November 1995)
28. Biblical teachings have played a pivotal role in the consistent decision by most major Christian and Jewish denominations to exclude homosexuals from the pulpit and prohibit religious marriages of gays and lesbians. (*Boston Globe*, 22 March 1998)
29. For scouts, the policy of excluding homosexuals is a core belief, Ron Carroll told the Human Rights Commission as the hearing against the Boy Scouts opened. It has been so since the first scout troop was chartered in this country in 1910, he said, so central to its definition of "morally straight" that the organization never saw the need to spell it out until relatively recently. (*Washington Post*, 2 August 1998)

Figure 9-10: Sample concordance lines where *exclude* collocates with the node words

The final collocate unique to the 1990s—*closet*—is tagged as the verb lemma of an adjectival collocate—*closeted*—which is a collocate shared between the 1990s and 2000s. The *closet* is described as "a key metaphor of the gay liberation movement of the 1960s and onwards" by Jo Eadie, as "'COMING OUT OF THE CLOSET' denoted an acceptance of one's own homosexuality or bisexuality, and an unembarrassed public discourse about it" (2004: 32). Sedgwick (1990) posits that the EPISTEMOLOGY OF THE CLOSET is a critical component of many Western cultures, and this is defined by "a mode of inquiry in which knowing the truth is paramount, having secrets is taken as proof of weakness or corruption, and the state and its institutions assume that they must regulate what is known, and what may be made public" (Eadie, 2004, p. 32). Therefore, the discourse of the *closet* is also one of embarrassing secrecy and deviancy imposed by a heteronormative society (Eadie, 2004, p. 32).

During manual analysis of the 73 resulting concordance lines of the collocation between *closet* and the gay set, additional CLAWS tagging errors have been exposed that would discount this collocate from analysis. Fortunately enough (for the purposes of comprehensive analysis), *closeted*—the most frequent realization of this incorrectly



identified lemma—is itself a collocate lemma shared between the 1990s and 2000s. Therefore, extended discussion appears in the next section.

#### **9.3.1.4 USAS A category collocates unique shared between the 1990s and 2000s**

*Closeted* is in fact the only USAS A category collocate shared between the 1990s and 2000s. This is an extremely interesting adjective, directly pre-modifying *gay/homosexual* men in all but two of the 88 concordance lines arising from the collocation. In each of these two exceptions, *closeted* modifies a span of time, as in line 30 below, where entire *lives* are closeted, metaphorically indicating that preferring privacy regarding the details of one's sexuality leads to an entire lifetime spent in confinement and hiding. In further concordance lines, this adjective is listed alongside other (negative) attributes, as in line 31, where the identity of fundamentalist minister by day is contrasted with that of the closeted, club-cruising gay man by night (who commits suicide after an AIDS diagnosis), or in line 32, where repressed sexuality is taken out upon subordinates. This state of being *closeted* is explicitly described as a form of dishonesty (line 33), which links to Sedgwick's (1990) assertion of the Western obsession with the truth and disclosure. However, it is also linked to depression and self-harm in a call for the veil of secrecy to be lifted in the same concordance line.

30. Accurate statistics on substance abuse in the lesbian and gay community are difficult to obtain because many lesbians and **gay** men live closeted lives, said Fraelean Curtis, president of the National Association of Lesbian and Gay Alcoholism Professionals. (*Boston Globe*, 8 October 1990)
31. The actress [Anne Heche] says that her father, who killed himself in 1983 after an AIDS diagnosis, was a fundamentalist minister while a closeted, club-cruising gay man, and that "I won't repeat his mistakes." (*New Orleans Times-Picayune*, 25 April 1997)
32. It's closeted gay men in positions of power and influence, using wives and children as a front, who act out their repressed sexuality on their subordinate staff and aides. (*St. Louis Post-Dispatch*, 2 January 1994)
33. "If people voluntarily come out of the closet that can go a long way towards saving young people's lives," says Kellogg, who resigned from his post last week to write books. "Many (closeted) gay people either live...dishonestly unhappy lives, kill themselves or do something close to killing themselves because of their depression. If someone they already honor comes out, and comes out voluntarily, it illustrates that homosexuality is not something to be apologized for. But if somebody is discovered to be gay, it just underscores the (perceived) need for people to be secret." (*Chicago Sun-Times*, 1 April 1990)

34. His experience raises the question of whether openly gay soldiers are as great a threat to unit cohesion as *closeted homosexuals*, whose enforced secrecy gradually distances them from their fellow soldiers. (*New York Times*, 28 June 1998)

**Figure 9-11: Sample concordance lines where *outing* collocates with the node words**

The identity is directly contrasted with *openly gay* rather than just ‘gay’ (unmarked) in line 34—showing that homosexuality is something that must be either shown or hidden, but never just a passing, mundane fact. For more on this discourse of showing and hiding, see *openly* in the discussion on c-collocates, section 9.3.1.6.

### **9.3.1.5 USAS A category collocates unique to the 2000s**

In keeping with patterns in the 1990s data, USAS A category collocates unique to the 2000s (*oppose* and *mainstream*) show enhanced discourse explicitly negotiating the terms of exclusion and inclusion of gay people into various roles of society; this would indicate a moral panic continually associated with this social identity group into modernity.

Each of the 123 concordance lines resulting from the collocation between *oppose* and the gay set has been manually analysed to ascertain the object of opposition; frequencies are listed below in Table 9-7. The most controversial subject in the concordance lines by a vast margin is *gay marriage*, with a frequency of 80, followed distantly by *gay rights/equality* (18), and *gay people in the military* (7). Notably, every one of these broad topics is represented by at least one collocate in a semantically significant category, and therefore will be covered in-depth in this chapter. However, this indicates that the discourse of the data overall is one representing (or at least incorporating with extremely high frequency) anti-gay voices of opposition, perhaps in higher frequency than pro-gay voices of support. This tension is further reflected in concordance lines of the next collocate—*mainstream*.

Being opposed	Frequency
gay marriage	80
gay rights/equality	18
gays in the military	7
gay members of the clergy	5
discrimination against gay people	3
gay adoption	3
gay practices/gay lifestyle	2
gay unions	1
the gay agenda	1
"mentioning gays as one of the groups at highest risk for AIDS"	1
"allowing a gay contingent to march in the St. Patrick's Day parade"	1

Table 9-7: Concepts/n-grams co-occurring in context with the collocate *oppose*

The use of *mainstream* is accompanied by judgments of characteristics and/or behaviours against cultural norms and values. The concordance lines are somewhat evenly split as regards who is *mainstream*; of the 51 total cases, 29 describe gay people as being in the *mainstream* or *mainstreamed*, in contrast to 22 concordance lines assigning this attribute to broader (presumably heteronormative) identity groups, e.g. *mainstream* America and the *mainstream* media. When homosexuality has been *mainstreamed*, as in lines 35, 36, and 37 below, this may be normalized or rejected. Whereas belonging to the *mainstream* is aspirational, and *mainstream* behaviours are accepted, praised, and considered appropriate across an entire culture, this transition from the 'margin' to the 'mainstream' is sometimes resisted or greeted with derision. Though one line (35) demonstrates that acceptance by *mainstream* America can open up more residential options for gay people, the very acceptance itself is still questioned, and in the late 2000s, "gays could be seen as another group of outsiders, especially in...the prolonged era of AIDS" (line 36). Twenty-five years after the initial outbreak of AIDS, this is still being indicated as a reason that gay people are not truly part of the *mainstream*, "despite signs of...acceptance" (line 36).

35. Thornell, an East Atlanta resident since moving here eight years ago, sees neighborhoods like East Point and Ormewood Park as fertile ground for gays looking to branch out of Midtown. Another incentive, suggests Thornell, is the growing acceptance of **homosexuals** by *mainstream* America. (*Atlanta Journal-Constitution*, 24 July 2003)
36. Despite signs of mainstream acceptance, **gays** could be seen as another group of outsiders, especially in today's political climate and the prolonged era of AIDS. (*New York Times*, 6 March 2005)
37. "The downside of mainstream acceptance of **gays** is that gays themselves become more mainstream, more conformist," John explains later. (*San Francisco*, 13 February 2006)

Figure 9-12: Sample concordance lines where *mainstream* collocates with the node words

A counter-discourse to this process of “*mainstream acceptance*” is apparent in sample concordance line 37 above, where this is said to lead to conformism in gay people, a negatively appraised attribute. This collocate shows membership in the traditional in- and out-groups being questioned and problematized in the 2000s, while conversion from one to the other are simultaneously being *opposed*.

We now move on to consider USAS A category collocates co-occurring with the gay set in the 1980s, 1990s, and 2000s subcorpora, to investigate whether similar patterns of construal are echoed across the diachrony.

### **9.3.1.6 USAS A category c-collocates**

Only three lemmata collocate consistently with the gay set across the 1980s, 1990s, and 2000s subcorpora, and these are *openly*, *movement*, and *liberation*. In combination, these three items construct a positive semantic prosody with a domain of empowerment and acceptance. However (as always), concordance analysis offers a more comprehensive look at the data.

Previously, Baker (2005, pp. 75–81) found collocates such as *openly*, *publicly* and *obviously* to contribute to a discourse of shamelessness associated with gay people. In contrast to *closeted* people, for instance, who “are ashamed of their sexuality and conceal it, a related discourse exists – one in which another set of gay people have *no shame*” (Baker, 2005, p. 79). In his study, this was exponentially realized by additional collocates such as *flamboyantly* and *hideously*, which do not appear in my corpus of American broadsheet news. However, an interesting discursive pattern does arise here. *Openly gay* people are most likely to be those in high-profile civic positions, working in politics, the clergy, and civil service. To be *openly gay* in the 1980s is considered ‘daring’ (line 38), and the political influence does appear to grow over the course of the corpus. Though enumeration does occur, the *topos* of numbers relating to threat does not seem as much a factor; in line 39, the number of *openly gay* elected officials is not negatively appraised, but rather presented as an argument against Republican lack of representation of this identity group. The inclusion of gay and lesbian police officers in the force is positively appraised in line 40, and the ordination of a gay female bishop is commemorated in line 41. Openness in public service is considered very highly indeed, and consistently throughout the corpus.

38. The diversity and tolerance of New York have attracted many homosexuals to the city. Many have dared to live *openly* as **homosexuals**, becoming vocal about their interests and politically influential. (*New York Times*, 24 January 1986)
39. Nationally, there are 20 *openly gay elected officials*. Only two of them are Republicans - Ald. Jim McFarland of Madison, Wis., and Robert Ebersole, town clerk and treasurer of Lunenburg, Mass. (*Chicago Sun-Times*, 24 April 1988)
40. "Here we are commemorating this interaction between the police and lesbian and gay community , and now we have *openly gay sergeants, lieutenants and cops* helping the participants; it's remarkable," said the lieutenant, who has been working overtime lately with a lesbian officer, Detective Vanessa Ferro, in the special events office . (*New York Times*, 17 June 1994)
41. She is in line to become the second openly gay bishop in the Episcopal Church, after the Right Rev. V. Gene Robinson, who took office in New Hampshire in 2004. (*Wall Street Journal*, 7 December 2009)

Figure 9-13: Sample concordance lines where *mainstream* collocates with the node words

The next items initially appeared to hold the same positive promise. *Liberation* and *movement* are the last collocates from the USAS A broad category to collocate in the 1980s, 1990s, and 2000s, and these often work in tandem (as evidenced by line 42 below). Though there are a good number of concordance lines to review to see examples of these items in use, there are no indications of their change over time. The 'gay liberation movement' referred to seems to have taken place in the 1960s and 1970s—before the corpus collection start date. Even in the 1980s (as in line 42 below), the "early years" of the *gay liberation movement* are referred to parochially. By the 1990s, *gay liberation* is described as the source of change from "effeminate" stereotyping to one "defiant of the heterosexual, homophobic majority" through changes in sexual practice (line 43). By the 2000s (line 44), *gay liberation* is so old that it can be "renewed", though this is the only reference to the rebirth of the *movement*.

42. Unlike the early years of the **gay liberation movement**, when wedding ceremonies were often unavailable or shunned as a cheap imitation of heterosexual behavior, gay people are now accepting the institution and restructuring it to suit their own needs for ritual commitment. (*San Francisco Chronicle*, 24 June 1989)
43. Mr. Rotello and Mr. Signorile point out that at least through the 50's, gay men were stereotyped as effeminate and often obtained sex by fellating the so-called trade, nominally straight men looking for satisfaction. With the advent of **gay liberation**, they turned to having sex with one another and many made of multipartnered anal sex a militant outlaw culture, defiant of the heterosexual, homophobic majority. (*New York Times*, 25 May 1997)

44. There was talk of second-generation AIDS plays (like Doug Holsclaw's farce "The Baddest of Boys"), a post-AIDS sensibility, a renewed assertion of **gay** sexual *liberation*, a backlash against so-called "victim art" and a call to look at AIDS in broader contexts. (*San Francisco Chronicle*, 7 June 2006)

**Figure 9-14: Sample concordance lines where *liberation* collocates with the node words**

As *gay liberation* is thus construed as pre-dating the corpus, I cannot detail much more of it here in the corpus-based analysis, though can assert that this is taken as for-granted in the social context of the data. However, this does run in contrast to the discussion of homosexual *activities* as deviant risk behaviours explored in the 1980s above, and the concept of *closeted* and *closet* that occur through the 1990s and 2000s. Therefore, I believe that the data shows the common acceptance of the past success of the *gay liberation* still has room to be questioned. Similar motifs (including *activism* and *advocacy*) appear in the next USAS broad category, G: government and public.

### **9.3.2 Category G: Government and public**

Collocates from the USAS G broad category comprise 22.0% of all de-duplicated collocates of *gay/gays/homosexual/homosexuals*. However, as illustrated in Figure 9-16 below, this category became more salient over time. Only one collocate (*discriminate*) appears in the 1980s, and this is shared with the 1990s. By contrast, texts from the 2000s contain six unique collocates, in addition to two collocates shared with the 1990s. In the following section, I explore the changing construal of gay people through the scope of their varying association with items from the semantic category relating to politics and government. As this semantic category does not relate directly to the research question of construal of risk and negotiation of in-/out-groups, discussion will be necessarily brief.

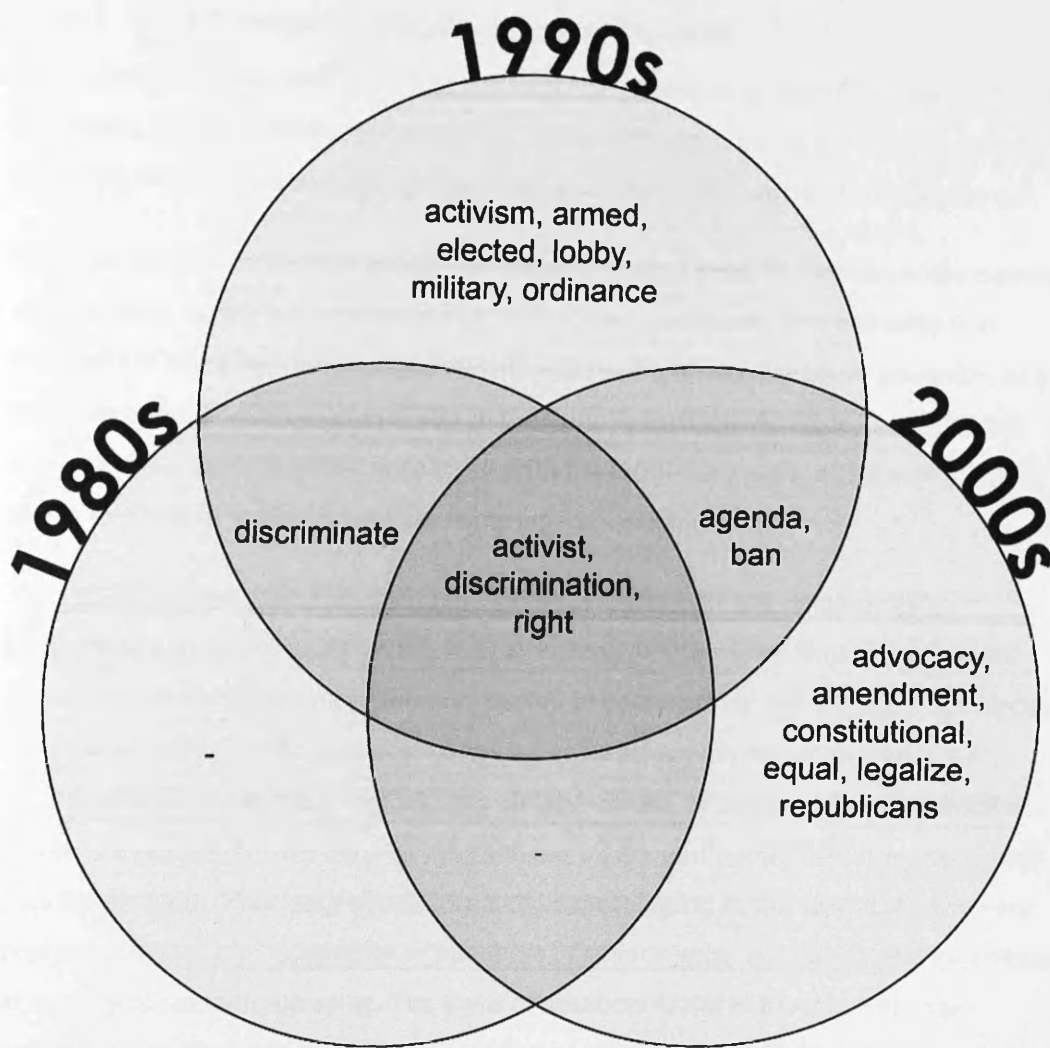


Figure 9-15: Venn diagram of USAS Category G collocates for *gay/gays/homosexual/homosexuals* depicted over time

### 9.3.2.1 USAS G category collocates shared between the 1980s and 1990s

There are no USAS G category collocates uniquely belonging to the 1980s, and only one collocating in both the 1980s and the 1990s: *discriminate*. I have already discussed the implications of collocation with *discriminate* in the context of PWAS in section 8.4.3, and the discourses arising are extremely similar. In the 1980s, there are many accusations of employers, housing officials, and insurance companies *discriminating* against gay people, following by discourses of a more legal nature in the 1990s regarding passage of laws banning such practices. Given the high overlap of these patterns, I refer to reader back to the previous section, and ahead to the c-collocates of this USAS semantic category (section 9.3.2.5), which includes the broader abstract concept, *discrimination*.

### 9.3.2.2 USAS G category collocates unique to the 1990s

Collocates of this category unique to the 1990s fall under three broad groups: activism (*activism*), governmental process (*elected, lobby, ordinances*), and the military (*armed, military*). This demonstrates a continued focus on legislation and inclusion in the era.

The abstract idea of *activism* collocates uniquely in the 1990s, in contrast to the human actor *activist*, which is a c-collocate that will be discussed later. This indicates that discourse of the 1990s was concerned with discussion of *activism* more generally, in a way that no longer held salience later in the sampling period. *Activism* as an abstract concept is less enduring than in contrast with the c-collocate *activist* (section 9.3.2.5), which lends itself to smaller stories featuring human agents.

Two additional collocates unique to the 1990s construct the gay community's participation in government, rather than as victims or citizens operating beneath its power (e.g. in discussion on *ordinances* passed to protect their rights); these are *elected* and *lobby*. Each of the 52 instances of *elected* collocating with the node words is a modification of the social group, e.g. 'gay elected official' as opposed to a statement of whom the gay community had *elected*. Of these, 22 are additionally modified by *openly* (see lines 45 and 46 below), which seems to be tautological in this instance—they are representing the gay community as members of government, and have made no attempt at secrecy about their sexuality. The *topos* of numbers features heavily with this collocation; in 22 concordance lines, *gay elected officials* are enumerated in some way or another, rather as a point of fact in a job title (e.g. in line 47) or in reporting of statistics (e.g. line 45). Line 46 is interesting in its mention of a *gay bloc*, signalling that the power of the oft-called 'gay vote' extends beyond candidates and into volunteer positions, before even considering voters. Some concordance lines also occur as a result of one story, represented in part by concordance line 47, wherein Secret Service guards greeted "a delegation of gay elected officials" wearing rubber gloves, in the hopes of protecting themselves against AIDS, "should any of the visitors be infected". This ill-informed or prejudiced behaviour is very heavily criticised by the Press. In the illustrative example below, they are said to have "shamed their service and embarrassed the White House" by behaving badly.

45. Coyle, one of Minnesota's three openly gay elected officials, said that, by the 1992 election, "as many as 100,000 people with AIDS may have died in this country." Lesbians and gay men "must demand a cure for AIDS." (*The Boston Globe*, 19 November 1990)



46. There are 70 openly gay elected officials nationwide, and 40 are on ballots in 18 states and Washington, D.C., this year. But the gay bloc transcends gay candidates. About 10,000 homosexuals have volunteered for Clinton. (*USA Today*, 29 October 1992)
47. Those uniformed Secret Service guards who donned rubber gloves to greet a delegation of gay elected officials shamed their service and embarrassed the White House they are supposed to protect. The guards presumably thought they were protecting themselves against the AIDS virus should any of the visitors be infected. (*The New York Times*, 16 June 1995)

**Figure 9-16: Sample concordance lines where elected collocates with the node words**

A unique semantic field constructed by collocates unique to the 1990s is that of civil service; both *armed* and *military* appear here and in neither of the other decades. Both collocates relate to the previously mentioned 1990s debate, headed by Bill Clinton, about the possible repeal of the 'Don't Ask, Don't Tell' policy prohibiting outed gay people from serving in the military. The popularity of this debate in this era alone, despite a ban against homosexuals in military service being enforced from World War II and only lifted in 2011, is emblematic of its time (Belkin, 2013). The 1990s was a cultural turning point for gay rights, where equality in contribution to society (including as civil servants) was being openly debated (Belkin, 2008).

### **9.3.2.3 USAS G category collocates shared between the 1990s and 2000s**

With the increased discourse on political participation by gay people in the 1990s came accompanying discourse opposing said integration. The two collocates shared between the 1990s and 2000s (*agenda* and *ban*) are symptomatic of this backlash.

The concept of 'the homosexual-' or 'gay *agenda*' was introduced into public discourse in America in 1992, when conservative Christian groups published inflammatory, anti-gay materials (e.g. Family Research Council, 1992) describing 'hidden' political agendas for organized opposition of heteronormativity in society. As evidenced by data in the corpus, the meaning has since shifted, and is sometimes in use now to describe LGBT political objectives progressively. The majority of occurrences, however, maintain the original (negative) prosody, with strong suggestions of moral panic, though these appear attributed to far right groups, and are most often distanced from the journalist through the use of scare quotes. The concept of the 'gay agenda' is also eschewed by gay people in the corpus, as in line 48. In the 2000s, however, the discursive construction of the "gay agenda" as being threatening to children recurs (line 58), and Christian groups are seen to still be actively opposing its advancement.

48. Eventually, Moyers interviews a representative of the gay rights organization Ground Zero and visits the homes of two gay couples, one male and one female. They dispute the notion of a “gay agenda” and list the traditional family values – “love, respect, compassion, support, integrity...” - that they personally honor. (*Denver Post*, 19 November 1993)
49. For groups that endorse self-described traditional values when it comes to education, such as Liberty Counsel, which has worked closely with CRC in its fight against the Montgomery public school system, and the Alliance Defense Fund, the mention of homosexuality invokes charges that advocacy groups are using the schoolhouse to push a “homosexual agenda” on children. (*Washington Post*, 9 December 2005)

**Figure 9-17: Sample concordance lines where the gay set collocates with *agenda***

*Ban* features as a collocate of the node words in both the 1990s and 2000s, with mutual information scores of 4.087 and 4.516. However, unlike *oppose* (a similar term, section 9.3.2.5) which shows both a semantic preference of *opposing* acts such as discrimination against homosexual people as well as *opposing* acts of homosexual people (e.g. *marriage*), *ban* is always prohibitive; its only change is in nature. In the 1990s, the most discussed topic of *ban* against homosexuals was that of the military restriction, whereas in the 2000s, this shifted to the *ban* on gay marriage (and to a lesser extent: adoption). This links in to discourses of legality associated with collocates unique to the 2000s: *legalize* and *equal*.

#### **9.3.2.4 USAS G category collocates unique to the 2000s**

A single topic links each of the USAS G category collocates unique to the 2000s (*legalize, equal, constitutional, amendment, advocacy, Republican*): the legalization of same sex marriage. This was extremely salient in reportage of the era. In 2000, Vermont became the first state in the country to grant full legal recognition to civil unions between same-sex couples; state legislature passed in 2009 has since legalized same-sex marriage there. By 2009 (the last year of data in the AIDS/HIV corpus), same-sex marriages were also legal in Massachusetts, Connecticut, Iowa, and Vermont, with several additional states issuing same-sex civil unions or recognizing same-sex marriages performed in other states. Though progress continues, it is slow. As of November 2013, same-sex marriage licenses are issued in only 14 of the country’s 50 states. Relating the plodding pace to the frame of a moral panic, Adam observes that:

Marriage has long been implicated in a politics of exclusion. It has been commonplace for ecclesiastical authorities in many places and eras to attempt to maintain religious endogamy by ruling nonbelievers outside the pale. Nation-building rhetoric employing analogies of the nation to the family (and thus marriage) inevitably manufactures a series of “others” thrown out of the national

family and uses marriage laws as a tool to mark that exclusion. In the American context, Nancy Cott<sup>21</sup> points to a lengthy record of marriage prohibitions based on race intended to prevent legal recognition of relationships between whites and Asian or African Americans. Lesbians and gay men find themselves unavoidably placed at the nexus of similar forces intent on capturing national identity, social privilege, or the moral high ground. (2003: 274-275)

The high frequency of this topic in the 2000s (as indicated by dominance of the collocates in this section) demonstrates the shift of moral panic from gay people as sexual deviants and (as members of the risk society) transferors of disease, to corruptors of sacred institutions such as family, as evidenced by the concepts and n-grams co-occurring in context with the collocates *oppose* and *legalize*, shown in Tables 9-8 and 9-9 below, respectively.

equal...	Frequency
rights	34
treatment	6
access to benefits/healthcare/jobs	6
marriage	3
citizenship	1
"status in community life"	1

**Table 9-8: Concepts/n-grams co-occurring in context with the collocate *oppose***

legalize...	Frequency
gay marriage	51
gay civil unions	7
gay sexual conduct/activities	2
gay relationships	1
"carrying 'legalize gay' signs"	1

**Table 9-9: Concepts/n-grams co-occurring in context with the collocate *legalize***

Many of the concordance lines arising from the collocation between the gay set and both *constitutional* and *amendments* are shared; these relate to the same topic of reporting, *constitutional amendments* on the allowance of marriage between same-sex couples, often related to the work of *gay advocacy* groups, lawyers, and organizations. The great majority of these instances relates state support of federal mandate (as in lines 50 and 51 below), and therefore does not fall within the remit of this thesis. It is notable, however, that once more social actors are removed from reporting. In only one of the 50 of *constitutional's* concordance lines do humans appear as being directly affected by *constitutional* rights—in line 52 below, the rights of "gay couples" are being upheld. The odds of including social actors in concordance lines of *amendment* are slightly higher,

<sup>21</sup> Cott, N. (2000). *Public Vows*. Harvard University Press: Cambridge, Massachusetts.

with three instances in 61 total concordance lines: see line 53 for an example. Again, however, it is important to note this dehumanizing and alienating discursive strategy of discussing the topics and issues regarding people while removing these people from the discourse entirely.

50. Still, Bush has made gestures that appear to be aimed at his conservative supporters rather than moderates: a *constitutional amendment* banning **gay marriage**, the recess appointments of conservative judges and ending the expiration of his tax cuts. (*Washington Post*, 4 April 2004)
51. A highly charged presidential election that saw 11 new states pass *constitutional amendments* banning **gay marriage** may boost attendance at this year's conference for activists and leaders in the lesbian, gay, bisexual and transgendered political effort. (*St. Louis Post-Dispatch*, 10 November 2004)
52. West has maintained he was upholding the **gay couples'** *constitutional* rights to equal protection -- and thus his oath of office -- by allowing them to wed. (*Washington Post*, 29 October 2005)
53. The *amendment* denies **gay** and lesbian **couples** rights afforded to heterosexuals. (*St. Petersburg Times*, 17 March 2006)

**Figure 9-18: Sample concordance lines where *amendment* and *constitutional* collocate with the node words**

One group of social actors is explicitly marked in the collocate list: the apparently hypocritical participation of *gay Republicans* in party politics (traditionally opposed to same-sex marriage, and firmly legislatively so throughout the 2000s) is problematized in concordance lines of this time. Nearly every one of the 101 instances of the collocation between *republican* and *gay/gays/homosexual/homosexuals* describes a dual identity; those showing distinct or conflicting identities (as in line 54 below) detail the perceived *republican* distaste for *gay* marriage and perception of this as a corruption of jurisprudence and (natural) tradition. The double identity is construed as duplicitous, and is often purported to result in rifts between politician and party.

54. In the platform they approved last month, *Republicans* associated **gay unions** with the distasteful recognition of "other living arrangements as equivalent to marriage." They proposed a Constitutional amendment "protecting" current marriage laws, which they said were grounded in "more than two centuries of American jurisprudence, and millennia of human experience." But this is a fiction. As Chauncey and Wolfson demonstrate, the rules of marriage have changed constantly. (*New York Times*, 26 September 2004)

**Figure 9-19: Sample concordance lines where *republican* collocates with the node words**

These localized cases of problematic individual behaviour notwithstanding, by the 2000s, reporting contained in the corpus largely problematizes moral panic discourses. Most of the collocate frequency of *legalize* and *equal* are contained in contexts which

report the stance of various politicians or parties on *legalizing equal marriage*, but some stand-outs incorporate additional extreme right anti-gay voices to reframe the argument. In line 55 below, several writers' assertions linking homosexuality with everything from children abuse to adultery to homelessness (all while echoing illness metaphors) are lampooned by the journalist, who acerbically asks whether we might next link terrorism to gay marriage. Line 56 problematizes the rhetoric on the website of the influential anti-gay-marriage organization Yes2Marriage, marking the journalist's disbelief with "Come on" before retorting that they have made a "shameful leap in logic" by concluding that schools would be unduly indoctrinating children.

55. Then there's the commentary written by a Times staffer who wrote that **gay marriage will jeopardize everything from Social Security to Mother's and Father's days**. And somehow, the homeless, jobless, and hungry will be ignored if gay marriage is allowed. Hundreds "or maybe 1,000" state laws will have to be changed if gays are allowed to marry, the writer said. How about terrorism? Can we somehow connect that with gay marriage, too? (*Boston Globe*, 4 April 2004)
56. However, the Yes2Marriage.org site states the law could easily be overturned, allowing the court to "impose same-sex marriage on the rest of society. This will require schools to begin teaching that **homosexual marriages are equal to marriages between a man and a woman**. Public schools will be forced to teach young children that two men being intimate are just the same as a husband and wife." Really? Come on, even a repeal of the law - unlikely given the recent conservative state Supreme Court appointments by Republican Gov. Charlie Crist - is not going to result in gay marriage becoming part of the school curriculum. That's a shameful leap in logic. (*St. Petersburg Times*, 12 October 2008)

**Figure 9-20: Sample concordance lines where *legalize* and *equal* collocate with the node words**

Interestingly, collocates from the 2000s tend to have a more negative overall prosody than c-collocates, which take into account the early ages of equal rights legislation for PWAs and gay people.

### **9.3.2.5 USAS G category c-collocates**

Three collocates from the USAS G broad category consistently occur throughout the 1980s, 1990s, and 2000s subcorpora: *activist*, *discrimination*, and *right*. This creates a stable discourse of political interaction (and intervention) in association with the social group under investigation.

*Activists* have unequivocally played a crucial role through the late twentieth and early twenty-first centuries in fighting to improve circumstances for minority groups (including gay people) and enhance social equality (Miller, 1995). This activism has

taken place “across a broad spectrum of sexual issues, from lesbian, bisexual, gay and transgender rights, to HIV treatment funding, and the freedom to practice SM [sodomasochism], and has resulted in many of the rights sexual minorities enjoy today” (Evans, in Eadie, 2004, p. 2). In the corpus, *activists* are construed as powerful and important, with secondary collocates in the 1980s including *right*; in the 1990s, *outspoken, well-known, prominent, and radical*; in the 2000s, *longtime, prominent, and political*. As there is no accurate and effective way to identify whether each *activist* in question is also *gay*, so that I might remain firm to my research aims in this chapter, I will not dwell too long exploring the construal of other social actors. Rather, I note here that the important role of *activists* are prominently acknowledged in the corpus, and move on.

The final two c-collocates in the USAS G category are closely related: *discrimination* and *right* both arise as a result of the civil rights debates and discourses surrounding this social group over the course of the corpus. The twinned depictions of a journey for equal *rights* and an end to *discrimination* are both indicative of moral panic and othering progressing even beyond immediate mention of HIV/AIDS. *Discrimination* in the workplace, housing market, and financial institutions (e.g. obtaining credit, banking, and mortgages) is a steady theme in the construal of gay people, and this is criticized throughout the corpus. As was the case in section 8.4.3 (when I compared various naming strategies for PWAs), locations of discrimination are named much more frequently than [human] sources of discrimination, removing agents at fault and making the *discrimination* more difficult to confront and cease. *Right* collocates secondarily with *groups, movement, and bill*, so this is not precisely within the purview of this thesis. However, various specific rights (to marry, adopt, etc.) have been discussed at length throughout this chapter, and one diachronic difference can be spotted in analysis of the concordance lines.

Node	1980s	1990s	2000s
<i>gay</i>	1014 (71.26%)	2303 (84.17%)	1300 (89.22%)
<i>homosexual</i>	281 (19.75%)	155 (5.67%)	29 (1.99%)
<i>homosexuals</i>	98 (6.89%)	144 (5.26%)	21 (1.44%)
<i>gays</i>	30 (2.11%)	134 (4.90%)	107 (7.34%)

**Table 9-10: Frequency breakdown of the collocation between *right* and the search nodes, expressed in raw hits and percentage of total collocation of search string per decade**

This major change in discourse over time can be viewed above, in the naming strategy/attribution directly preceding instances of *right*, as evidenced by the frequency breakdown of the collocation in Table 9-10 above. Over time, *gay rights* has increased

greatly in proportional frequency, accounting for approximately 70% of cases in the 1980s, but nearly 90% of cases in the 2000s. Conversely, **homosexual rights** has fallen very deeply out of fashion, from nearly 20% in the 1980s to less than 2% in the 2000s. A similar pattern is witnessed in constructions such as “*rights for homosexuals*” and “*rights for gays*”. This shows that in conjunction with rights discourse, the naming strategy/attribution marker favouring identity rather than behaviour has come into heavier usage. I am hopeful that this could signpost a heightened discourse of equality and *rights* focused on social actors as a whole rather than specifically relying upon behaviours previously leading to moral panics, in the future. Analysis of construal of social actors through the scope of USAS Category S collocates may be able to shed some more light on the veracity of this pattern.

### **9.3.3 Category S: Social actions, states and processes**

The final USAS broad category to be explored in this chapter is the most prominent. Category S (social actions, states and processes) consistently contains the highest proportion of collocates in each decade, accounting for 41.0% of the 1980s collocates, 37.7% of collocates from the 1990s, and 41.9% of all collocates in the 2000s. De-duplicated, Category S contains 36.6% of all collocates of the search string. A Venn diagram showing dispersion of collocates is in Figure 9-21 below.

As demonstrated there, collocates are quite evenly dispersed diachronically, but do shift over time. The number of items contained within this category, in combination with the nature of the group (social processes) makes its analysis extremely relevant to the aims of this thesis.

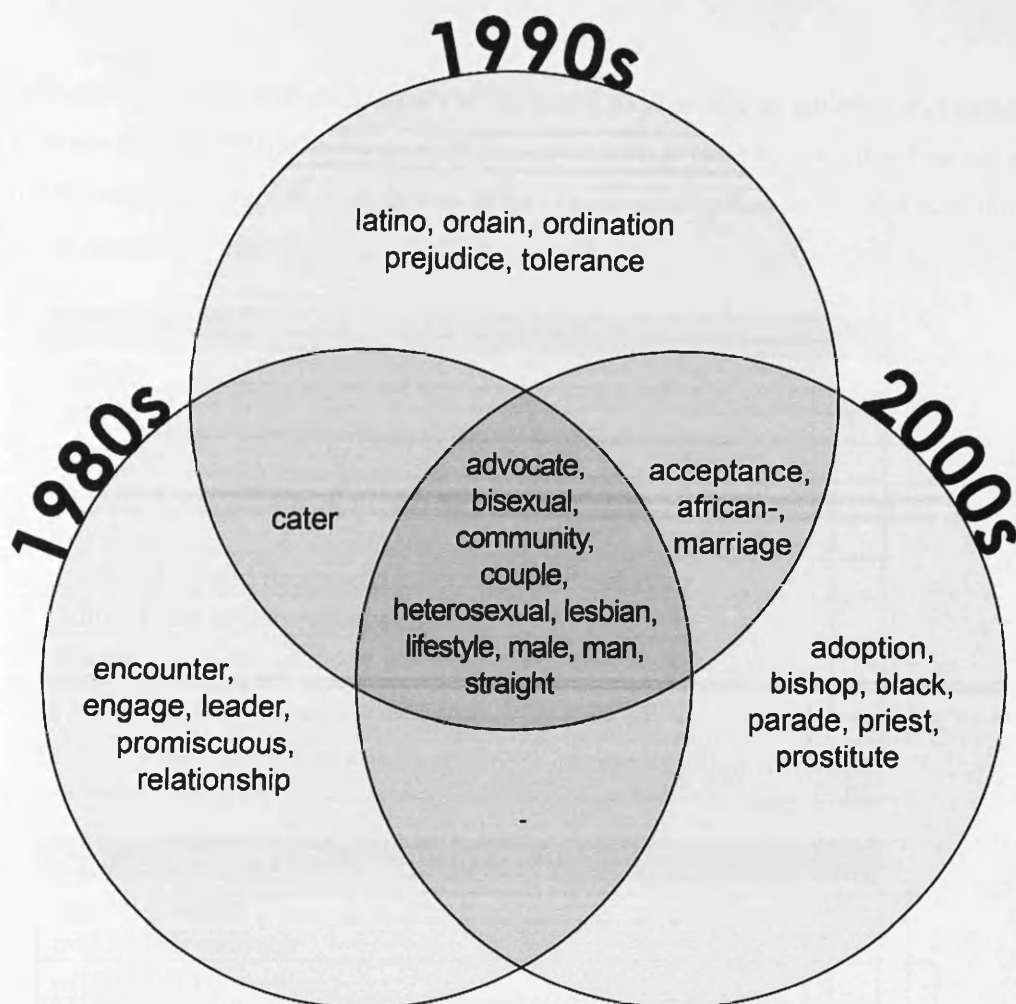


Figure 9-21: Venn diagram of USAS Category S collocates for *gay/gays/homosexual/homosexuals* depicted over time

### 9.3.3.1 USAS S category collocates unique to the 1980s

In the 1980s data, there are a total of 23 lemmas collocating with the gay set, including nine that are unique to this decade. This percentage of unique collocates (39.13%) out of the overall figure is the highest of the three decades under analysis, indicating that SOCIAL ACTIONS, STATES AND PROCESSES were expressed somewhat uniquely during this early time period. These 1980s-specific S-category collocates are as follows: *encounter, engage, leader, promiscuous, and relationship*. A semantic prosody of sexual activity presents itself clearly in reviewing these results.

Two collocates unique to the 1980s—*engage* and *encounter*—are tightly tied to the behavioural construal of homosexuality. In table 9-11 below, the acts *engaged* in by the social group under analysis are listed in order of frequency, and categorized by semantic meaning. Here we see that euphemistic expressions of sexual acts are the most frequent, occurring in 58 concordance lines, as compared to 41 concordance lines where homosexual activity is explicitly referred to as a sexual behaviour. *Encounter*, another



collocate occurring with a frequency of 58, could additionally be coded in this category. Overuse of euphemistic expressions indicates a trend in the 1980s of avoiding use of overt sexual language, though this could be a factor arising from the use of national, broadsheet, reputable papers.

<b>Euphemistic expressions of sexual behaviours</b>	
activity	30
acts	14
behavior	6
homosexual or bisexual practices	2
conduct	1
homosexual and promiscuous heterosexual behavior	1
homosexual or heterosexual activity	1
life style	1
practices	1
relations	1
<i>Total:</i>	58

<b>Non-euphemistic expressions of sexual behaviours</b>	
sex/intercourse	16
oral and/or anal sex	5
private sexual conduct	1
sexual activities	1
<i>Total:</i>	21

<b>Non-euphemistic expressions of 'weird' or 'immoral' sexual behaviours</b>	
high-risk sex	8
public sex, oral sex, masturbation or exhibitionism	2
sodomy	2
[would not engage in] sexual conduct with children	1
extra-marital affairs	1
anal homosexual congress	1
casual sex	1
prostitution	1
sex acts with animals	1
<i>Total:</i>	20

Total non-euphemistic freq.: 41

<b>Other</b>	
"blood terrorism"	2
<i>Total:</i>	2

Table 9-11: Behaviours *engaged in*, when *engage* collocates with the gay set, expressed by semantic category and frequency of occurrence in concordance lines

When considering the rarer, non-euphemistic set of behaviours being *engaged in*, a further pattern arises: approximately half of the expressions are unmarked or neutral—including *homosexual sex* or *gay sexual activities*—but the other half indicate ‘weird’ or immoral sexual behaviours, including *public sex*, *extra-marital affairs*, and *sex with animals*. The most frequent in this category—*high-risk sex*—indicates unsafe or somehow ‘strange’ behaviours left up to the reader to imagine. One behaviour here (‘blood terrorism’) stands apart, and relates to a highly criticised statement by a conservative Illinois state legislator: “Pullen has also said that some homosexuals engage in ‘blood terrorism’ by deliberately donating infected blood and encouraging others to do so.” (“Views of 4 U.S. AIDS Panelists Hit”, *Washington Post*, 26 August 1987). This very obviously relates behaviours of gay people not directly related to sexual activity to threats to greater society.

Only one collocate unique to the 1980s construes homosexuality as more than a collection of sexual behaviours and/or attitudes: *relationship*. *Relationship* collocates with the gay set 174 times in the 1980s subcorpus, and might be seen as a precursor to the collocate *marriage* shared between the 1990s and 2000s, when this was legislatively topical. While this word is superficially much more positive than those concerned entirely with sexual behaviour, its use in context often leans toward euphemism and negative semantic prosody. Of 174 instances, *homosexual relationships* were sometimes described as secretive undertakings (16) or specifically extramarital affairs (14), or even leading directly to AIDS (8), but the vast majority of instances of this collocate construe *gay relationships* as stable, loving, and needful of wider societal acceptance.

The collocate *promiscuous* relates to the category discussed above, of ‘weird’ or immoral sexual activity. The cultural trope of the promiscuous gay man appears to have been an emerging discourse or stereotype of risk behaviour through the 1980s (lines 57 and 58), which had fallen out of fashion in the collocates from the 1990s and 2000s. This collocate in particular construes gay sexual relationships as “transitory and unimportant” and is suggestive of “a large sexual appetite” (Baker, 2005, p. 82), both othered qualities.

57. At present, the disease is largely confined to two population groups: promiscuous homosexual men and needle-using drug addicts. But health officials expect it to spread into the non-addict heterosexual population, increasing from 1,100 cases now to 7,000 cases within five years. (Newsday, 30 October 1986)

58. The bad news is that this modest progress, by isolating *promiscuous gays* and black and Hispanic intravenous drug abusers as the groups chiefly at risk of infection, threatens to reduce public concern-and perhaps public funding-for AIDS research. (*Washington Post*, 28 October 1987)

**Figure 9-22: Sample concordance line where *promiscuous* collocates with the node words**

I would like to problematize a strictly negative analysis of the presence of this collocate, however. Given the context of the data (reportage on AIDS/HIV), I believe that the presence of the pre-modifying qualifier (*promiscuous*) does add an important dimension in categorization of risk: it is not just gay men who are exposed, but gay men who engage in certain risk behaviours – namely, promiscuous (unprotected) sex. While these lines could be read to generalize promiscuity across the social group, I suggest instead that the marking is critical in strategic social sequestering.

Though USAS S category collocates unique to the 1980s are dominated by a semantic preference for sexuality, there is one (positive) indicator of growing social power. *Leader* is a very commanding and constructive collocate of the gay set unique to the 1980s. Given the extreme paucity of naming strategies and concordance lines construing either PAKs or PWAs as positively appraised agents, *leader* stands alone as a naming strategy in its depiction of power on behalf of social actors. Though it is unclear whether these *leaders* are themselves *gay/homosexual*, position within the group can be assumed as more directly related than, for instance, *activists* or *advocates*. Unlike other naming strategies, secondary collocates of the gay set and *leader* include adjectives of positive evaluation (*prominent, longtime*) as well as verbs of emotion (*fear, worry*) and material agency (*meet*). *Leaders* voice the emotion concerns of the groups or communities (see lines 59 below), and are depicted as being influential politically, meeting with legislators and with the aim of enacting change (line 60).

59. *Leaders of homosexual* rights groups say they fear that irrational public beliefs about casual transmission of the virus - exemplified in recent boycotts of schools allowing AIDS victims to attend and in calls for the quarantine of AIDS patients - could lead to blacklists of those found to carry the virus, creating a class of modern-day lepers. (*New York Times*, 30 September 1985)
60. The sit-in at the governor's office coincided with a lobbying effort by other **homosexual** *leaders* who met with legislators during the morning seeking support for bills backed by the gays and lesbians during the current session of the Legislature. (*San Francisco Chronicle*, 10 May 1988)

**Figure 9-23: Sample concordance line where *leader* collocates with the node words**

Overall, collocates unique to the 1980s contribute to othering by typifying gay people through (oft 'weird') sexual practices. The single exception to this is *leader*, which is a

gem not only in this subcorpus, but in the entire AIDS/HIV corpus, as well as the Katrina corpus. In the sections below, I explore construal through the 1990s and 2000s for additional traces of burgeoning social empowerment.

### 9.3.3.2 USAS S category collocates shared between the 1980s and 1990s

Only one S-category collocate—*cater*—is shared between the 1980s and 1990s, and gay people are always the recipients of this action. In the 1980s, a large number of businesses across the country *catering* to gay people were closed as a result of moral panic and risk society. Establishments such as “bathhouses, bars, theaters and pornography shops” (line 61) were seen as a threat, associated both with the spread of AIDS, and with the spreading threat to public safety and decency, driving up instances of “offensive touching” and “public lewdness” (line 62).

61. Diane McGrath, the Republican mayoral candidate, proposed yesterday that bathhouses, bars, theaters and pornography shops catering to homosexuals be closed to avoid the spread of AIDS. (*New York Times*, 2 October 1985)
62. “Public safety is becoming a problem!” bellowed an irate citizen after another irate citizen suggested Rehoboth was turning into “bar city.” “These bars are *catering* to **homosexuals**! We have a police force that’s incapable of handling this. On the beach and the boardwalk, there is offensive touching! Public lewdness!” (*Washington Post*, 1989)
63. He said the town has many services and events *catering* to **gays**, including a weeklong carnival with a cross-dressing parade, a well-funded AIDS assistance program and several school programs promoting cultural diversity. (*Boston Globe*, 15 September 1991)
64. Tempers blazed at town meetings. “These bars are *catering* to **homosexuals**,” some said in heated tones. (*Washington Post*, 29 August 1999)

Figure 9-24: Sample concordance line where *cater* collocates with the node words

By the 1990s, *catering* to gay people is construed in slightly more positive terms, with towns and businesses boasting their wares. The very things being incorporated to *cater* to gay people indicates a shift in ideology: in line 63, a weeklong carnival, “well-funded AIDS assistance program” and “several school programs promoting cultural diversity” have usurped the title from bars and clubs. This is not to say that moral panic dies down, entirely, however; in line 64, we see an instance where *catering* to gay clientele causes tempers to flare.

### 9.3.3.3 USAS S category collocates unique to the 1990s

Collocates unique to the 1990s and belonging to the USAS S category are quite mixed. Evident is a continued interest in increasing *tolerance* and fighting *prejudice* (as

illustrated in concordance line 65 below), but conversely, deviancy doubling occurs in collocation between the gay set and *latino*, and exclusion on the basis of moral panic is apparent in concordance lines of *ordain* and *ordination*, a matter discussed at length further below.

65. Stable government protects their minority rights and their right to question and dissent. Socially, **homosexuals** in the county wish only to have dignity. It is incumbent upon Westchester's political leaders to articulate the need for social *tolerance* of **homosexuals** and to nurture an atmosphere defused of **gay hatred**, **gay prejudice** and **gay fear**. (New York Times, 7 October 1990)

Figure 9-25: Sample concordance line where *prejudice* and *tolerance* collocate with the node words

In the collocation between *latino* and the gay set, we encounter somewhat of a return to form; this collocation shows a preference for deviancy doubling, as described in sections 4.3.5.1 and 8.4.1.2, among others. Fifty-seven concordance lines result from this collocation, of which nine have been excluded from this portion of analysis due to referring to non-human agents, e.g. “gay Latino nightclub”). In 38 out of remaining 48 total cases, social actors are described as being both *Latino* and gay—in some instances, this is additionally modified by further out-groups, such as being a PWA/H (two instances) or a woman (one instance, reflected below in sample concordance line 81). This double deviancy group—along with black gay men, linked conjunctively nine times in these concordance lines—is associated with risky sexual practices and heightened risk of HIV infection, symptoms of moral panic (see lines 66 and 67 below). The double deviancy leads one such construed individual to report that this leads others to believe that “Latino gays are twice the scum of the city” (line 68). This is not the only pattern occurring, of course; in an additional 11 concordance lines of the total 57, the social actor is either *gay* or *Latino*, though lack of overlap is problematized explicitly, for instance in line 69 below. This difficulty in construing social actors as occupying more than one attribute per identity is then reflected back, as in line 70, where the woman draws strength from both “her gay and Latino communities”, which are distinct and discrete entities.

66. Both are triple jeopardy fighters: **homosexual**, *Latino*, the one a woman, the other HIV positive. (Seattle Post-Intelligencer, 28 September 1991)
67. Surveys by the San Francisco Department of Public Health show that black and *Latino* **gays** and bisexuals have altered risky sex practices only slightly, leading to charges that AIDS organizations have failed to educate those groups. (San Francisco Chronicle, 2 April 1990)
68. “There are a lot of people who think that Latino gays are twice the scum of the city. This has created a new awareness. It acknowledges you exist.” (New York Times, 11 November 1991)

69. The unspoken rule is that you can exist only as one thing at a time -- a Latino or a gay man -- with no recognition of reality's complexity. Add an H.I.V. diagnosis to this mix and it gets complicated. (*New York Times*, 15 May 1993)
70. She draws strength from both her gay and Latino communities, and particularly from her mother, a 63-year-old widow and retired health-care worker who at first opposed Morales' involvement in the lawsuit. (*San Francisco Chronicle*, 4 March 1990)

**Figure 9-26: Sample concordance lines where *latino* collocates with the node words**

The final two S category collocates unique to the 1990s (*ordain* and *ordination*) are both collocates at this time due to relatively frequent discussion surrounding the election and appointment of gay people into the clergy. This indicates a continued preoccupation with the inclusion of gay people in previously restricted social (religious) areas. However, as these focus on the process of *ordination* rather than the social actors involved, I defer analysis and discussion of this discourse to a later section 9.3.3.5, collocates of the 2000s. There, I describe the collocation of *priest* and *bishop* with *gay/gays/homosexual/homosexuals*, as it holds a greater relevance to my research interest of the construal of social actors.

#### **9.3.3.4 USAS S category collocates shared between the 1990s and 2000s**

Three USAS S category collocates are shared between the 1990s and 2000s: *acceptance*, *marriage*, and *African-*. While these seem to indicate quite a variety of discourses, we will be able to dispatch with each of them quite quickly in turn.

Firstly, the appearance of *African-* is not due to rising rates of infection in *gay African-Americans* (as was my first impression), but in fact occurs as a result of an oft-repeated tagline appearing in the *New Orleans Times-Picayune*: "SATURDAY CITYWIDE Man Talk, an informal discussion group focusing on gay African-American issues: healthy identity, spirituality, HIV." Eighty-six of the 87 concordance lines of the 1990s, and 50 of the 55 concordance lines of the 2000s, are recurrences of this phrase. As a result, this has been excluded from analysis.

The discourse around *acceptance* is, in many ways, the foil of discourse around *oppose*, explored in Section 9.3.1.5. The frequency of concordance lines indicating or questioning *acceptance* of various concepts and n-grams in the 1990s and 2000s appears in Table 9-12 below. The most frequent object of *acceptance* is gay people generally in both subcorpora, with dramatically less popular themes being acceptance of gay people in the church/clergy, in the military/workplace, as being granted the right to marry/adopt, and

of general concepts related to homosexuality, e.g. 'gay lifestyle' and 'homosexual behaviour'.

Acceptance of...	Frequency 1990s	Frequency 2000s	Total
gay people generally	52	48	100
in the church/clergy	6	18	24
in the military/workplace	6	4	10
lifestyles/culture	5	4	9
behavior/acts	5	4	9
families/marriages	2	7	9
as performers	4	1	5
gay rights/issues	1	4	5
'the gay vote'/in politics	2	3	5
as scouts	1	0	1
as citizens	1	0	1
N/A	3	2	5

Table 9-12: Concepts/n-grams co-occurring in context with the collocate *acceptance*

All of these concepts have been previously discussed at length, along with *civil unions/marriage*. I pause here only to discuss the ideological implications of a highly frequent prosody of *acceptance* occurring around social actors themselves, rather than of specific behaviours or roles. In the 1990s, *acceptance* is a goal that is construed as far-off (line 71) or impossible, given misconceptions of implications for greater society (line 72). In the 2000s, *acceptance* is still in the distance, though generations (and readers) are urged toward this (lines 73 and 74).

71. Many feel isolated because of society's still-limited *acceptance* of **gays**. (*New York Times*, 26 June 1990)
72. Do people honestly believe that the *acceptance* of **homosexuals** would turn them loose, sparking a mass sexual revolution in the effort to convert heterosexuals into their ranks? How are they immoral? (*St. Louis Post-Dispatch*, 17 July 1993)
73. Also to the report's credit, it urges *acceptance* of **gays** and lesbians, acknowledging that there is no scientific evidence that they can or should be converted to heterosexuality. (*St. Louis Post-Dispatch*, 24 July 2001)
74. We are watching attitudes change, one generation replacing another, in the direction of full *acceptance* of **gay** Americans. We're not there yet. (*Boston Globe*, 4 March 2004)

Figure 9-27: Sample concordance lines where *acceptance* collocates with the node words

Arguments against *acceptance* that are alluded to in these concordance lines and others like them are that enveloping the out-group into the in-group could disrupt the social (and sexual) order of the *status quo*. This is indicative of a prevalent moral panic continuing through the 2000s, and no longer associated with certain actions seen to be

imposing upon 'straight society' (e.g. marriage, parenthood, participation in religious services and political process), but rather a deep moral panic associated with the people themselves, independent of all factors save social identity attribute.

### 9.3.3.5 USAS S category collocates unique to the 2000s

As a decade, the 2000s contains the highest number of unique USAS S category collocates. Four out of these six collocates mark aspects of identity, related either to race (*black*) or career (*bishop, priest, prostitute*)—note that two of these (*black* and *prostitute*) tend to be othered identities. Two nominalized processes (*adoption, parade*) round out this section's grouping.

While some of the discourses indicated by collocates unique to the 2000s are quite familiar (for instance, the collocate *black* functions in exactly the same way as the collocation with *latino*, as described in section 9.3.3.3), most are indicative of changing social climate and attitudes toward the social group under investigation. One collocate in particular—*parade*—marks the heightened espousal of pro-social activities, eschewing the 'closet' and centred around pride. Though this doesn't describe human social actors (and is therefore out of the scope of the current research), I mention it in passing as it signposts evolution of public sentiment and personal acceptance (see, for instance, line 75 below).

75. Thirty-five years after simply marching in a gay pride parade was seen as an act of courage, Chicago now has an openly gay alderman, a gay representative in Springfield and six openly gay judges in the Cook County Circuit Court system. (*Chicago Sun-Times*, 27 June 2004)

**Figure 9-28: Sample concordance lines where *parade* collocates with the node words**

Another collocate unique to the 2000s truly indicative of the discourse of the time is *adoption*. The co-occurrence between *adoption* and *gay* or *homosexual* reflects many states' (ongoing) evolution of legislation to either include or disallow *adoption* by same-sex couples or single gay parents. Despite the media coverage in the corpus being quite superficially balanced—e.g. including various sources quoted as both being for or against adoption by gay parents, or largely objectively reporting on the passage of bills—the phraseology is quite telling. In 60 concordance lines, *gay modifies adoptions*, as in sample concordance lines 76 and 79 below. The addition of this attributional adjective construes adoption by gay parents as a subdivision of adoption that is entirely unlike 'traditional' adoption (by heterosexual parents). This phraseology also has the effect of removing human agents from reportage; readers are not interacting discursively with



potential parents, but with an ideological concept of *gay adoption*. Use of *adoption* where humans are involved is much less common, occurring in only 37 concordance lines. These 37 results can be further subdivided: 20 use *gay* or *homosexual* in the attributional fashion, describing adoption by “gay men” or “gay couples” (e.g. line 77), whereas 17 lines use *gay* or *homosexual* in the essentialising fashion, describing “adoption by gays” (e.g. line 78). Though it is preferable (in my opinion) to include social actors in the construal of a process that very centrally comprises people, the essentialising use is still damaging as compared to the attributional use. A clear pattern in preference for phraseology does not emerge in contrasting concordance lines showing disagreement with the concept (line 76) or reporting decreased resistance to it (line 77), though many of the concordance lines using *gay adoption* are more likely to use additionally harmful language, as in the essentialising use in line 79. The majority of social actor deletion during reportage on this event, however, is extremely negative. Further, the fact that the frequencies of these uses (17 essentialising, 20 attributional) are so close is disheartening, as the reportage identifies prospective adoptive parents solely on the basis of their sexuality nearly as often as they are presented as potentially more complex characters.

76. A federal judge upheld Florida's ban on **gay adoptions** yesterday, accepting the state's long-standing policy that allowing only married couples to adopt is in the best interest of the state 's children. (*Washington Post*, 31 August 2001)
77. Judge James Lawrence King upheld a 1977 Florida law that bans adoptions by **gay men and lesbians**. “It is arguable that placing children in married homes is in the best interest of Florida's children,” Judge King wrote. (*New York Times*, 8 September 2001)
78. In 1977, fewer than 15 percent of Americans supported adoption by gays. In 2001, 52.5 percent approve. Attitudes about sexual orientation, including legal same-sex relationships or workplace rights, enjoy similar percentage gains because people, fortunately, learn. (*Newsweek*, 28 November 2001)
79. Resistance to **gay adoption** has waned as studies show that children raised by gays look a lot like those raised by straights--and are no more or less likely to be gay. (*Newsweek*, 7 July 2003)

**Figure 9-29: Sample concordance lines where *adoption* collocates with the node words**

One popular method of identifying *gay* people in the 2000s was with nomination strategies based upon job title, both upstanding (*bishop, priest*) and deviant (*prostitute*). These contradictory identities require further explanation.

*Bishop* and *priest* both collocate due to a series of high-profile cases of *gay bishops* and *gay priests* being elected to post during this time period. Many familiar modifiers co-occur with this collocation: in 85 total concordance lines where *bishop* collocates with

the gay set, 45 of these are additionally modified by *openly*, and two are modified by *actively*; in 74 total concordance lines where *priest* and the gay set collocate, 10 occurrences are modified by *openly*, and one additional instance is modified by *practicing*. This is an interesting aspect of the continuation of shamelessness and oversexualisation in discourses around gay people; see, for example 'disclose' and 'closeted' in line 80 below. More interestingly, while each text contains at least one mention of AIDS or HIV, the immediate contexts that *gay/gays/homosexual/homosexuals* are mentioned in here do not relate to the disease, but do nonetheless contribute to another moral panic. In line 81, an established in-group meets to discuss a split from their church following the election of an "openly gay bishop"; in line 82, the group is enumerated, both in terms of the number having this sexuality in the clergy (which is already established to be threatening to the moral order, as in the line before), and in terms of those who are sexually active (which is itself a 'deviant' behaviour, and additionally against the rule of their post, and therefore the alleged will of God). More rarely, but very importantly, a counter-discourse also arises in these concordance lines indicating that it is this broad othering that can directly lead to antisocial, true deviant behaviour (e.g. alcoholism and sexual abuse) in a minority (see line 83). Moreover, othering of this group is indicated as allowing for greater breaches of social order in the in-group, as in a double-bind on sexual behaviours of any kind in the clergy, including extramarital affairs and sexual abuse of minors (see line 84 below).

80. The only **gay bishop** to disclose his sexuality before now is retired Bishop Otis Charles of Utah, who sent a letter to the church's House of Bishops in 1993 sharing his experience as a closeted gay churchman watching the bishops wrestle with the issue. (*New York Times*, 8 June 2003)
81. Nearly 3,000 Episcopalians opposed to the church's election of an openly gay bishop last year are gathering in northern Virginia this weekend to discuss a possible split in the national Episcopal Church. (*New York Times*, 11 January 2004)
82. Based on case studies of 2,700 priests, Sipe thinks 30 percent of priests are gay and that half of the 30 percent are sexually active. (*New Orleans Times-Picayune*, 13 February 2000)
83. Some priests and behavioral experts say that imposing strict rules on sexuality can actually make matters worse. "With gay priests, if we continue to say their sexuality is disordered and we continue to tell them to hide and deny who they are, then we're going to have problems, because inner conflict like that can lead to alcoholism, it can lead to sexual abuse, it can lead to violations of celibacy," said the Rev. Robert Nugent, a Baltimore priest and co-founder of New Ways Ministry, a national gay center. (*New Orleans Times-Picayune*, 6 February 2000)

84. The protection of the aura of celibacy demands the coverup of a whole range of activities with the common denominator of tarnishing the aura of celibacy. This puts priests in a situation of mutual blackmail. A **gay bishop** who is innocent of pedophilia may be hesitant to push for punishment of a fellow priest if that could expose him and his own partners to unwanted scrutiny. (*Boston Globe*, 24 March 2002)

**Figure 9-30: Sample concordance lines where *adoption* collocates with the node words**

Semantic preference for (hyper-) sexuality is also reflected in the gay set's collocation with *prostitute*, occurring 56 times. In only two cases out of this total is the compound identity of a *gay prostitute* being described; the rest occur in list structures. When the structures of the lists are broken down, some patterns do emerge; the last item in the list is most frequently intravenous drug users, *prostitutes* are much more likely to occur just before the last position (last -1, or L-1 below), whereas *gay* men favour the second-to-last position. Other social groups appear with far less frequency. I hypothesize that this is symptomatic of a hierarchy of social sequestering, where some 'bads' are ranked worse than others. In the 2000s, intravenous drug use was the most socially deviant activity, as evidenced by its position as the rhetorical 'anchor'. However, this would bear deeper analysis, and is a clear point for further study.

#### **9.3.3.6 USAS S category c-collocates**

The final analysis section in this chapter is also the most diverse. USAS broad category S contains 10 collocates, relating broadly to gender (*male, man*), sexuality (*bisexual, heterosexual, lesbian, straight*), relationships (*couple, community*), socio-political action (*advocate*) and behaviour (*lifestyle*).

While there is a good number of varied collocates in this section, very few lend new insight into the data. For instance, *male* and *man* are both c-collocates of the gay set over the course of the corpus, but analysis of these relationships exposed no new patterns in the data. Secondary collocates of both of these primary collocations are a familiar bunch, including *bisexual, lesbian, primarily, high-risk, black, openly, discriminate*, and so forth. Frequency of these collocations is certainly a factor: *male* collocates with the gay set 1,632 times over the course of the corpus, and *man* collocates 15,352 times. The node terms *gay* and *homosexual* pre-modify *man* and *male* so often that this can be considered a core identity and naming strategy within itself, and all collocations discussed so far reflect back upon these collocates in some fashion. Similarly, secondary collocates of *couple* reflect primary collocates of a given time period—in the 1980s, *allow, including*; in the 1990s, *openly, adopt, marry*; in the 2000s, *marry, adoption, openly, marriage, union*,

*civil, legal, benefit*. *Couple* functions as the head noun in noun phrases with adjectival pre-modifiers *gay* and *homosexual* (i.e. where *gays* and *homosexuals* are not unmodified plural common head nouns of the phrases). As was the case with *activists* in Section 9.3.2.5, it is unclear whether the referents of the naming strategy *advocates* are in fact PWAs, and therefore this has been discounted for deeper analysis. It is noteworthy, however, that the core group of PWAs is characterized by the presence of *activists* throughout the corpus, indicating both an enduring powerlessness (having social actors act on their behalf) and social import (being a prominent enough group to warrant such advocacy). However, as no new discourses were exposed through analysis of primary or secondary collocates of these items, I proceed now to less expected groups of collocates, beginning with groups of differing sexualities.

The search nodes *gay/gays/homosexual/homosexuals* collocate consistently throughout the 1980s, 1990s, and 2000s with quite a wide variety of other markers of identity related to sexuality, namely: *lesbian, bisexual, heterosexual, and straight*. Each of these relationships conveys a different meaning in context—these will be discussed in turn. Illustrative concordance lines appear below in groups of three, clustered by collocate and ordered by time period.

*Lesbian* is the naming strategy most closely aligned with *gay/gays/homosexual/homosexuals*; in many concordance lines, these groups appear on opposite sides of a conjoining conjunction. Concordance line 85 demonstrates the relationship between the *gay* set and *lesbian*. When this collocation is embedded within a text, the discourse tends to be of a political or radical nature, particularly in the 1980s and 1990s (e.g. line 85). *Gay* and *lesbian* rights and legislation appear to be a central point of discussion in the news; this group continues to gain power through the corpus, and are construed as wielding a good deal of political clout by the close of the data set. It is only during the last third of the corpus that additional groups appear alongside this collocation, as in “transgender rights” in concordance line 86 below. This, in combination with the c-collocation of *bisexual*, indicates that while the media shows awareness of other identities defined in part by sexuality, it was still the core *gay* and *lesbian* group that formed the basis of interest until the last decade.

85. Anti-bias violence bills have passed and are now law in California and Wisconsin. The bills in New York, California and Wisconsin protect all minorities from bias-related violence, including *lesbians* and **gay men**. (*New York Times*, 7 August 1988)

86. But in a field where most of the half-dozen or so candidates have strong records on issues of importance to gay men and lesbians, including gay marriage and transgender rights, other factors besides sexual orientation are informing the gay vote. (*New York Times*, 15 July 2001)

**Figure 9-31: Sample concordance lines where *lesbian* collocates with the node words**

By contrast, collocation with *bisexual* has a greater possibility of denoting risk, particularly in the 1980s (see line 87). In the 1990s, the stereotypical risk groups of MSM was being questioned and criticized (as reflected in line 88 below), and rejection of affiliation with certain risk groups (here: drug addicts, bisexuals, homosexuals, haemophiliacs) juxtaposes the concept of evolving risk in the new risk society: monogamous heterosexuals who might have had one bad blood transfusion, for instance. By the 2000s, the continued ideological assumption that the classic MSM group should be maintained in Federal Drug Administration (FDA) questionnaires—thereby reproducing stereotypes and opening up blood banks to potential harm—is sharply critiqued as a “sweeping generalization” (line 89), illustrating enhanced fossilization of new risk society conceptions.

87. Dr. Jeffrey P. Koplan, an assistant director of the Centers for Disease Control, said 71 percent of the AIDS cases had occurred among homosexual or bisexual men. (*New York Times*, 25 May 1983)
88. My husband died a little over a year ago, and he died from acquired immune deficiency syndrome. He wasn't a drug addict, bisexual, homosexual or a hemophiliac, and he wasn't famous. (*New York Times*, 25 June 1991)
89. Here's a suggestion: Before people roll up their sleeves to donate blood, ask them, “How often do you have unprotected sex of any sort?” By changing the question, the FDA would make the national blood supply even safer - and would do away with its sweeping generalizations about gay and bisexual men. (*Boston Globe*, 27 November 2005)

**Figure 9-32: Sample concordance lines where *bisexual* collocates with the node words**

Surprisingly, the collocation between the gay set and *heterosexual* has a highly similar function. In the 1980s, risk related to male-male sex is foregrounded, though “promiscuous heterosexual behavior” (line 90) is additionally marked as a risk activity. It is interesting that promiscuity is indicated in the heterosexual aspect of this concordance line but omitted from the homosexual side of the conjunction, indicating an underlying ideology that all homosexual sex is a) more risky, and b) more promiscuous. By the 1990s (line 91), we see a balancing out of risk, with “homosexuals and heterosexuals” being presented alike. Again, this is indicative of risk society (large-scale exposure to uncontrollable, catastrophic possibility) rather than moral panic (related to isolated, othered groups). However, the 2000s are demarcated overtures of a return to

moral panic discourse. Despite the end result (integration of “homosexual and heterosexual communities”) being positive, the conceptualization of AIDS as one of many “gay issues” being ‘forced’ “to the forefront of society” is negative and heteronormative (line 92). I will further address the binary distinction between ‘homosexual communities’ and ‘heterosexual communities’ below, shortly.

90. But Phillips, reflecting the views on AIDS espoused by some influential conservatives, said he believes there ought to be criminal penalties for people who engage in **homosexual** and promiscuous heterosexual behavior. (*Boston Globe*, 15 February 1987)
91. The Centers for Disease Control does not know where the virus originated, but it infected both **homosexuals** and heterosexuals alike. (*St. Louis Post-Dispatch*, 15 October 1995)
92. Ironically, the despair of AIDS eventually led to further integration of America’s homosexual and heterosexual communities by forcing gay issues to the forefront of society, activists said. (*Atlanta Journal-Constitution*, 29 June 2003)

**Figure 9-33: Sample concordance lines where *heterosexual* collocates with the node words**

Alongside *heterosexual*, another marker of this identity—*straight*—additionally collocates with the gay set. The co-occurrence of these groups creates familiar discourses. In the 1980s (line 93), we see the *straight* and *gay* people belonging to separate ‘worlds’ (both of which ostracize and discriminate against PWAs). By the 1990s (line 94), identity groups remain opposed, but are linked by a common goal of acceptance and diversity. However, the same double-edged prosody recurs in the 2000s: while it is positive that young gay kids are “less psychologically wounded, more self-confident [and] less isolated” than their predecessors of previous generations, the assertion that it is “increasingly difficult to distinguish between gay and straight teens today” (line 95) is problematic and could feed into a moral panic (of the concealed other) in some circumstances.

93. I was luckier than many guys I knew who had been fired, evicted and ostracized by the *straight* and **gay worlds** for having AIDS. (*San Francisco Chronicle*, 21 September 1985)
94. Yes, straight or gay, rich or poor, healthy or sick, we can all learn to accept one another. (*New York Times*, 30 May 1994)
95. That’s why it has become increasingly difficult to distinguish between gay and straight teens today -- or even young gay and straight adults. Less psychologically wounded, more self-confident, less isolated, young gay kids look and sound increasingly like young straight kids. (*Chicago Sun-Times*, 4 December 2005)

**Figure 9-34: Sample concordance lines where *straight* collocates with the node words**

Having explored the relationship between various naming strategies in conjunction with the gay set, I now move to description of this core group *en masse*: the *gay community*. While the sense of belonging and attachment that is associated with the *community* metaphor has been powerful for health policy, political mobilization, and individual psychological welfare, the core concept of a *gay community* is essentially that, defined entirely by “shared experiences of a transgressive sexuality” (Irvine, 1994: 238), gay people “inhabit communities...completely set off from the rest of society, that they are members of an altogether different culture and even a different species” (Stein and Plummer, 1994: 179). This conceptualization, though still in active usage (as evidenced by its status as a c-collocate), may be out-dated, as “collectivized and sexualized notions of community ignore the diversity within the gay scene and new research on post traditional communities and the personal sources of solidarity” (Wilkson, Bittman, Holt, Rawstorne, Kippax and Worth, 2012: 1162).

My first point of investigation was to acquire a frequency breakdown for the collocation between *community* and the constituent nodes contained within the gay set. The result is shown in Table 9-13 below. Two trends are notable: the steady increase of preference for *gay community* and the steady decline of *homosexual community*. (Note: *community + gays* and *community + homosexuals* are always minority percentages, but these undergo the same increase/decrease relationship.) Why might this be?

Node	1980s	1990s	2000s
<i>gay</i>	1581 (81.16%)	2346 (93.39%)	1352 (96.92%)
<i>homosexual</i>	352 (18.07%)	142 (5.65%)	32 (2.29%)
<i>homosexuals</i>	8 (0.41%)	7 (0.28%)	0 (0.00%)
<i>gays</i>	7 (0.36%)	17 (0.68%)	11 (0.79%)

**Table 9-13: Frequency breakdown of the collocation between *community* and the search nodes, expressed in raw hits and percentage of total collocation of search string per decade**

Analysis of the concordance lines shows that *homosexual + community* conveys risk associated with sexual practice (as in line 96); sexual activity, as I have previously demonstrated, is a defining feature of the sense of *homosexual*, particularly in contrast to the more identity-centric social sense of *gay*. This use fell out of fashion as perceived risk decreased. For the most part, secondary collocates of the dominant *gay + community* pattern from the 1980s indicate a largely political semantic preference for this relationship: the gay community is sizable, organized, galvanized and politically active (see line 97 below).

96. Homosexual men still represent 75 percent of the disease's victims, and the specter of AIDS haunts every member of the homosexual community, especially in the cities where it is most prevalent (they are, in descending order, New York, San Francisco, Los Angeles, Miami, Newark, Houston, Chicago, Boston, Washington and Philadelphia). (*New York Times*, 6 February 1983)
97. The city's official indifference to the AIDS threat is perplexing, given the fact that Houston is home to a highly visible and politically powerful gay community. (*San Francisco Chronicle*, 30 July 1987)
98. The idea of a "post-gay era" now being discussed with regard to the progress gays have made in politics and public policy starts from the conviction that we are already members of society and, therefore, inherently possess all the rights enjoyed by the majority. Though we are not at that point yet, I believe we are close. Once this insight is shared by more than some of us, it may be that a single community of gays and lesbians, urban and suburban, wisely will grow together into one; and that the distinctions of urban and suburban gay communities, and the differences that such terms connote, will be obliterated once and for all. (*Newsday*, 3 May 1999)
99. "It really signifies support on the part of the organized Jewish community for the gay and lesbian Jewish community," said Rabbi Suzanne Griffel of Congregation Or Chadash, a Reform synagogue in Lake View that has catered to gay and lesbian Jews for 25 years. (*Chicago Sun-Times*, 18 June 2001)

Figure 9-35: Sample concordance lines where the gay set collocates with *community*

By contrast, the 1990s concordance lines are more concerned with definition of the gay community, with secondary collocates including *predominantly, within, specifically, primarily, active, suburban, and metropolitan*. However, some mixing of messages is evident in these lines, illustrated in example 98 above. The writer posits that gay people are at once "already members of society" while simultaneously calling for "a single community of gays and lesbians" to "grow together" without distinctions—disregarding the distinction that excludes non-gay, non-lesbian members of society. This is indicative of the climate of the time; "[a]ssumptions of solidarity...developed in the seventies political movement and through the AIDS crises of the eighties, [were] increasingly difficult to sustain" (Ridge, Miniciello and Plummer, 1997: 146) as more individual communities emerged. By the 2000s, secondary collocates show splintering and hybridization with other *communities*, including *latino, urban, black, and jewish* (see line 99 above). Discussing a slightly different cultural context, Fraser (2008: 261) sums this up by stating that:

It is also important to note that along with the advent of queer, other related factors have also contributed to current thinking on the notion of gay community. For example, the increasing acceptance of gays and lesbians within (parts of) mainstream Australian culture and society has taken some of the emphasis off gay community by offering other options for sociality. And, as



suggested earlier, political organizing in the absence of the “glue” of the AIDS crisis has revealed the extent to which difference infuses and cuts across the category of gay. (2008: 261)

The *gay community* is in part defined by a group of people taking part in the same *lifestyle*—another c-collocate of the *gay set*. If one is uncertain as to what the *gay lifestyle* entails, it is not difficult to unpick this in even the most passing review of the context. In the 1980s, the *gay lifestyle* was “fun”, centred around bathhouses (an additional collocate not analysed due to its semantic categorisation) – and gay nightclubs (line 100), or (described from a 2000s perspective) defined by ‘sexual liberation’—or promiscuity (line 102).

100. Business in gay nightclubs and bath houses has plummeted -- partly from unfounded fears of contamination. Once-popular dark rooms, where clients could have sex with unseen partners, have acquired a grim nickname - - corridors of death. “The **gay lifestyle is supposed to be fun; people hate to be serious**. When we make announcements about AIDS at a dance, everyone boos,” said Roberto Silva, owner of a gay club called Acropolis. (*Boston Globe*, 26 February 1987)
101. He asked “Did Louganis knowingly flirt with his health by engaging in a gay lifestyle?” What is the gay lifestyle, and does Miklasz flirt with his health by engaging in the heterosexual lifestyle? As a gay man, I don’t believe that my lifestyle is any way different from Miklasz’s! I pay taxes. I have a full-time job. I have bills that need paying. I have a life partner, and I have my family and friends. Is that the gay lifestyle? It sounds an awful lot like what could be anyone’s lifestyle. HIV/AIDS is not a gay disease. HIV/AIDS doesn’t know boundaries. (*St. Louis Post-Dispatch*, 2 March 1995)
102. Some gay activists who are critical of the marriage movement also see it as part of the fallout from the AIDS epidemic, which quashed the sexually liberated lifestyle of gay men as they tried to fight the perception that they were promiscuous carriers of a plague by becoming more like everybody else. (*New York Times*, 30 July 2006)

Figure 9-36: Sample concordance lines where the *gay set* collocates with *lifestyle*

The conceptualization of a single *gay lifestyle* (just like a ‘gay community’) that is maintained by all gay people is overgeneralized, harmful and prejudiced. It is questioned in the corpus very rarely, but one lengthy criticism appears in line 101 above, from the 1990s. Here, the moral panic (othered sexuality) is challenged bluntly, and the risk (“HIV/AIDS doesn’t know boundaries”) is instead visibly foregrounded.

## 9.4 Summary

In this chapter, the construal of a single social group was analysed over time for traces of changing discourses. This was accomplished by ascertaining search nodes of high frequency, and targeting collocates from the most proportionally populous USAS broad

semantic categories. Evidence of continuing moral panic and elements of risk society discourse was apparent through the 1980s, 1990s, and 2000s, though themes and ways of othering did shift over time. Results related to identity construal can be found in 9.4.1 below. Finally, the method used here has been refined, both since the sister Katrina chapter 6, and over the course of the thesis. Perceived efficacy is discussed in section 9.4.2.

### **9.4.1 Of results**

Each collocate from three broad USAS semantic categories (A, G, and S) was analysed in this chapter.

Collocates from the USAS A broad semantic category (general and abstract terms) show a clear shift in discourse over time. The context surrounding the construction of gay people in the 1980s is primarily concerned with definition and negotiation of AIDS-related risk (e.g. *active, activity*), with traces of social sequestering evidenced through the isolation of risk to *mainly* and *mostly* gay men and people who use intravenous drugs. From the 1980s to 1990s, the discourses changed to wider discussion of sexuality, for instance, whether a social actor is *exclusively* homosexual or in the *closet*. By the 2000s, opposition to gay people accepting certain roles or integrating into the *mainstream* prevailed. Over time, collocates are more and more disconnected from the concepts of AIDS/HIV specifically and of risk generally. C-collocates from the USAS A broad category (*liberation, movement, openly*) indicate a lasting concern with obtaining social equality, but also a preoccupation with secrecy and the concept of shamelessness associated with gay people.

Degree of change is not so evident in analysis of collocates from USAS category G: government and public. Broadly, collocates from the 1990s show more focus on political and legal exclusion (*discriminate, military*) compared to collocates from the 2000s (e.g. *amendment, legalize*), which show both progress and heated debate in legislation aimed to rectify inequality. The presence of c-collocates *activist* and *discrimination* signpost enduring poor treatment past the locus of AIDS discourse, and the necessity of others in wielding power on behalf of gay people.

Shifting discourses were the most apparent in analysis of collocates from USAS category S: social actions, states and processes. Risk of contracting and spreading HIV is dominant in the early data, as are traces of discriminatory discourses and tropes, assigning

negative out-group qualities to gay people's sexual behaviour, e.g. *promiscuous*, *encounter*. Again, the 1990s and 2000s show a shift from risk society discourse (spreading an epidemic) to moral panic indicators not related to AIDS/HIV, but rather to other attributive and behavioural associations with gay people. This is due in part to claiming of legal/political power by the othered group: previous research and concordance line analysis of collocates from the 1990s indicate that gay people and activists played a large role in reversing discriminatory discourses in this era. The discovery of therapies in the 1990s that prolong lives, and in the 2000s which made AIDS a chronic illness for many (who had access to this) also shifted the risk discourse. Rather, the right of gay people to occupy various roles in society, particularly to belong to a church, and to marry or adopt, are protested throughout the later portion of the corpus.

Overall, whereas the 1980s discourse contained a mixture of meanings displaying both risk society and moral panic discourses, collocates from the 1990s and 2000s indicate a greater focus on moral panic. This was not to do with risk associated with AIDS/HIV, but integration into a heteronormative society. One method of maintaining out-group status was social sequestering in the 1980s, and deviancy doubling in the 1990s and 2000s, where gay people are aligned with others, such as drug users, child molesters, and so forth. Though the risk society cycle seems to have moved on, the moral panic cycles continue to proliferate.

#### **9.4.2 Of method**

In this chapter, proportional distribution of collocates was considered both on the micro level (per decade) and on the macro level (de-duplicated and tallied across the corpus). This has allowed for a large-scale diachronic analysis of one social group.

The minimum frequency of collocation was raised to account for differences in corpus size and frequency of nodes, as compared to the Katrina chapters and previous AIDS/HIV search terms. This has resulted in a much more manageable number of collocates, still of high interest. Though some noteworthy collocates of a lower frequency will have necessarily been excluded, the rewards have been greater than the sacrifices. Secondary collocation was more feasible in this chapter than ever before, allowing an even more robust view of the discourses created.

Viewing collocates as proportional distributions of broad semantic categories has been an effective method of gaining an overview of general change in discourse over time. More nuanced differences (e.g. shifting use of a collocate appearing in more than one decade) were exposed in qualitative analysis, making this a more than sufficient integration between corpus-based broad analysis and discourse analytical close reading. The incorporation of Venn diagrams to visualize the evolution of discourse was also extremely helpful in revealing patterns of topic change over time.

A final evaluation of this method as a whole can be found in the next chapter, the conclusion.

## **Chapter 10: Conclusion**

### **10.1 Introduction**

In the conclusion of this thesis, I will begin by discussing my main findings, to include both those relating to the construal of social actors in American 'Act of God' reporting (section 10.2.1) and the process of developing and applying the method of proportional semantic collocation to two corpora (section 10.2.2). In section 10.3 I discuss the perceived efficacy of the methods utilised throughout the corpus, and in section 10.4 I reflect critically upon my work, presenting a discussion of some of the limitations of the study, before responding in part to these with suggestions of further work in section 10.5. The thesis concludes with section 10.6, where I consider how the research journey has impacted me as a researcher, and offer some final thoughts.

### **10.2 Discussion of findings**

In chapter 1, I began this study with the overarching research question:

1. How does American media discourse negotiate the concept of risk in 'Acts of God' reporting on Hurricane Katrina and the AIDS/HIV epidemic, particularly in relation to attributes of identity such as ethnicity, class, and sexuality?

In the course of exploring this broad question, I have discovered various points of distinction in the corpora. Notable features related to general concepts (moral panic and risk society in 10.2.1, deviancy doubling and social sequestering in 10.2.2, and lack of agency in 10.2.3) are discussed below, before I move on to discussion of the actionable research question in section 10.2.4.

#### **10.2.1 Moral Panic versus Risk Society**

Moral panic and risk society discourses frequently overlap and interchange in both the Katrina corpus and the AIDS/HIV corpus. Othered attributes (e.g. being poor, black, homeless, gay) and behaviours (e.g. promiscuous, unprotected, and especially male homosexual sex, intravenous drug use, sex work) are assigned to social actors in the corpus, and this was found to contribute heavily to the symbolic removal of power and creation of distance between those most affected by 'acts of God' and the 'average American'. Some collocates did implicitly indicate that blame is attributed to victims in 'act of God' reporting. For instance, *innocent* PAKs are contrasted to other inhabitants of

the Astrodome who are labelled ‘murderers and rapists’ in section 4.3.1. PAKs are also blamed for endangering disaster support staff by refusing to (or being unable to) evacuate in time, or aggravating post-disaster conditions through water, disease, and filth metaphors. They are generally construed as a sullyng agent when returning to the New Orleans area too soon, described in terms of *topoi* of threat and numbers in reference to the greater ‘us’ population. In early texts of the AIDS/HIV corpus, ‘deviant’ behaviours are obvious indicators of othering leading to moral panic about the collapse of social order and spread of disease. However, as the disease spread, risk populations became more diverse, and minority groups countered bigoted and othering discourse levelled at them, this shifted.

Over the course of both corpora, blame is increasingly assigned to institutions, particularly the local and federal governments, healthcare service providers, legal bodies, and social service organisations. The continual risk associated with acts of God, and the inability to compartmentalise these risks into deviant groups, is construed as a failure and/or betrayal on the part of the institution. Noticeably, risk society discourse does traditionally look to the press as an agent of protection and catalyst for change, but in these corpora, the press have omitted accusations of wilful misrepresentation (e.g. figures on looting and other crimes in Katrina) and have failed to digitally reproduce the most inflammatory texts, whether for time or virtue, we may never know. In this way, the press maintains its own power while re-negotiating blame in risk society discourse, amplifying previous moral panics around minorities and supporting indictments of governmental, medical, and other social figures perceived as not having done their part—all while perpetuating a disempowering discourse of victimhood.

### ***10.2.2 Deviancy Doubling and Social Sequestering***

Throughout analysis on both corpora, I noticed a pattern of linking othered groups, or deviancy doubling. When two groups are repeatedly linked in powerful, institutionalized discourse, this reflects a linkage between respective social actors that is presumed to be widely comprehensible. In the case of deviancy doubling, groups that have little in common other than heightened distribution of ‘bads’ are construed as sharing equal vulnerabilities, and collapsed into one large othered group, e.g. *elderly, infirm, poor, and black*. This has two effects. First, the very different circumstances of vulnerability and experiences of risk groups are obscured, dramatically reducing the possibility of opening dedicated dialogue on the complex issues surrounding racial inequality, an aging population, and the inadequate public healthcare system in America. Further, listing

othered groups has a cumulative effect. By linking various othered groups together (sometimes quite tenuously), deviancy doubling allows greater distance between readers and 'us' groups and the reported upon 'them' groups. If it is difficult to relate to a social actor who is rarely given voice or agency in a given text, who is construed as 'deviant' due to attribute or behaviour, it is exponentially more difficult to empathise with an entire list of these. Deviancy doubling indicates much about the ideology of a text's producers, and is also one discursive method allowing consumers to 'opt out' from problematizing in-/out-group membership when considering a text.

An additional pattern noted throughout the thesis has been identified as 'social sequestering', or a segregation of certain groups into risk or moral panic categories, typified by discourses creating distance or dissuading reader identification with these groups. An early indicator of this pattern collocation with exclusivizers and/or particularizers (e.g. *mainly, mostly, predominantly, disproportionately*). In concordance analysis, unalienable mitigating factors leading to negative social repercussions (such as contracting a terminal illness, or losing one's home to a hurricane) might be explicitly or implicitly portrayed, allowing or encouraging readers to make judgment calls on deservedness of such repercussions. There is also an element of 'geographical distancing' involved in social sequestering; groups further away from publication centres (e.g. in the south, from the north, or in Africa, from America) are construed as having greater risks and less advanced methods of managing these. I theorise that these patterns allow risk to be compartmentalised and deemed inapplicable to the general public, assuaging fears arising from living within the risk society.

### **10.2.3 (Lack of) Agency**

Considering the contents of the entire thesis, the finding that has surprised me the most is lack of agency in the construal of othered social actors. Though previous research has documented lack of agentive force as a known feature of relatively powerless groups, this is often offset by prevalent discourses construing others as antisocial, negative agents (e.g. committing acts of crime, taking part in violence or terrorism), contributing to early stages or iterations of moral panics.

In both the Katrina and AIDS/HIV corpora, people affected by Katrina and people with AIDS are nearly entirely passivated. This is a somewhat positive outcome, as I did not locate highly frequent reference to 'looting' or 'sinning', for instance. But on balance, I consider this to be an extremely harmful and negative method of social actor construal

that is all but endemic across publications, time period, and source event. There are some very striking omissions. While PAKs are helped, they are not described as *helping* one another, despite reports that such pro-social behaviour was widespread. Likewise, PWAs are treated, but do not ever seek treatment, placing them in the powerless periphery even in terms of their own illness.

Linguistic deletion of power in this way indicates the symbolic removal of power that was (and, in many cases, continues to be) experienced by PAKs and PWAs. As patients and beneficiaries, these social actors are included in reportage as the context in which other, more powerful people act. Though I searched specifically for the ‘victims’ of acts of God to gain a better understanding of their actions and experiences, the results framed other participants entirely. Reportage was found to focus on social actors such as doctors, lawyers, politicians, engineers, and scientists, presenting their experiences most dynamically.

This offers a clear view of in- and out-group membership. Those who are passive, wounded, recipients of charity are othered by deletion of power to alter their circumstances (for better or worse). Those who are active, healthy, bearers of beneficence are praised and elevated to hero status. I believe that this is another form of social sequestering, whereby the suffering of others—particularly those who might not be so distant from the in-group in many qualities—is framed out in favour of focus upon the success of the unaffected.

#### ***10.2.4 Research questions revisited***

**In Chapters 4 and 7, I responded to the first actionable research question:**

2. Which identities are associated with the highest frequency of in- and out-group qualities in the Hurricane Katrina and AIDS/HIV corpora?
  - a. To what extent do these identities correlate to existing ‘folk devil’ categories? What does negative/positive construal of identities indicate about the underlying ideologies of the publications represented in the corpus?

In both of the first analysis chapters, *people* was the most frequent naming strategy. Frequent types of out-group attribution were common across both corpora. Enumeration and the *topos* of numbers typified the construal of both groups, depicting



them as unmanageable threats to social order. Attributes of helplessness (such as *disabled, sick, ill*) were frequent in the Katrina corpus, and verbal collocates in the AIDS/HIV corpus positioned PWAs as the beneficiaries of help. In the Katrina corpus particularly, race was (superfluously) indexed, though this was overtly indicated as a major factor between successful and failed evacuation/rescue. Notably, (lack of) income was hugely salient in the construal of both PAKs and PWAs; to be *poor* in America is marked, and to be a *poor* survivor of a natural disaster or victim of a terminal illness is doubly deviant. *Poor, low-income, working-class* people are consistently construed as socially feeble and potentially dangerous, certainly to the moral order, as they drain resources (for recovery, housing, treatment, scientific experimentation) from the perceived national 'fund'. Duration of othered status also emerged as important in the Katrina corpus, where the temporary, storm-related *hungry* and *homeless* were compared to the long-term *hungry* and *homeless* as somehow less othered, and more deserving of aid. This has implications for America's abiding handling of social welfare issues, whereby enduring social problems are not perceived as newsworthy, critical, or actionable, especially in comparison to sudden catastrophic events rending the partition between in- and out-groups.

Conversely, the presence of positive adjectival collocates (e.g. *wonderful, smart, talented, generous*) do not relate to the out-group of PAKs, but rather to the in-group of 'hero' figures from other locales, who contributed to the charitable salvation of people displaced by the hurricane. In the texts, the relationship between risk, race, and income is overt: upper-class white people are construed as having left early and lost little. This in-group also maintains agency throughout both corpora, *treating, rescuing, and assisting* PAKs and PWAs, as well as being represented intertextually in the form of long directly and indirectly attributed quotations.

It seems that to be part of the in-group in these texts, one should be in the middle- or upper-classes, white, employed, able-bodied, and completely self-reliant, depending not upon institutional assistance, but rather in a position to provide aid to the less fortunate. In this way, the in-group is construed as being removed from the hazards inherent in risk society discourse. Conversely, the out-group is the recipient of all of the 'bads' in the texts; they lack financial resources, health, agency, and therefore: power. The out-group is therefore most reliant on institutional and individual assistance, but in these texts, they are repeatedly framed as having little or no access to governmental/medical/charitable aid.

### Chapters 5 and 8 dealt with the second actionable sub-question:

3. What are the most frequent naming strategies employed in Hurricane Katrina and AIDS/HIV reporting?
  - a. What can corpus-based discourse analysis of associated attributes, transitivity, and cognitive metaphors tell us about the in- and/or out-group qualities attached to the people that the naming strategies represent?

Below *people*, the most frequent naming strategies in the Katrina corpus were *resident*, *evacuee*, *victim*, and *survivor*; in the AIDS/HIV corpus, they were *patient*, *victim*, *sufferer*, and *carrier*. Each of these naming strategies displayed its own particular semantic preference, prosody, and profile. While some similar patterns occurred with these as in the construal of *people* in Chapters 4 and 7, there was a far greater emphasis on distinguishing categories of risk, need, and deservedness. For instance, use of terms such as *victims* and *survivors* in the Katrina corpus indicate that successful rescues will be reported, as opposed to *residents* and *evacuees*, who are subjects of failed or absentee rescue attempts, and are disempowered through attribution and lack of agency in assisting themselves or one another. *Evacuees* were frequently construed via water metaphors, with 'waves' 'pouring' out from the Gulf Coast and being 'absorbed' by new cities. *Residents* in the process of evacuating are described as dirty water, *clogging* and *jammimg* outbound lanes. However, small numbers of returning (and in minority cases, escaping) evacuees and residents are more positively conveyed through the use of water metaphors.

The delimiters of in- and -out-group membership are defined by the collocates of common PWA naming strategies; the social repercussions of *quarantining*, *isolating*, *shunning*, and *discriminating* against *patients*, *victims*, *sufferers*, and *carriers* form a central debate running alongside construal of each of these nomination strategies. The reason for this sort of othering behaviour is clear: attributes associated with all naming strategies show skew toward negative semantic prosody. In the AIDS/HIV corpus, markers of (homo-)sexuality are significant in relationship to *victims* and *sufferers*, linking desolation (victimhood and suffering) to perceived sexual deviance. The social actors in this corpus are associated with other 'fringe' social actors: homeless people, handicapped people, and foreigners. The discourse surrounding them is borrowed from terrorism or crime discourse, with discussion of *known* PWAs. Shortage of treatment and

limits on care available to *patients* and *victims* is an object of some alarm in the AIDS/HIV corpus, especially as related to institutional denial of services. But of greater concern is the perceived overuse of such resources (e.g. hospital beds, healthcare benefits) at the supposed detriment of non-PWAs, which furthers the *topos* of threat/strain contributing to the moral panic.

These discourses construct PWAs and PAKs not as members of a shared social circumstance (drawing upon the same medical and social services as other American residents), but as somehow receiving more than their due. This constitutes both moral panic discourse (surrounding the attributes or behaviours contributing to the conditions of reporting) and risk society discourse (surrounding the apparent inability of the government to assist all citizens to the highest level).

**Finally, Chapters 6 and 9 were concerned with the third actionable sub-question:**

4. What evidence does corpus-based discourse analysis of the two corpora give of changing or resistant discourses surrounding othered social groups over time?
  - a. What kind of contextual or cultural changes result in shifting preference for terms or interest in discourses surrounding certain identities?

In Chapter 6, I analysed the changing use of *refugee* as it applied to people internally displaced by Katrina. This naming strategy was selected manually through intuition and my own interaction with news reporting at the time of the storm. Initially in the corpus, UNCONTESTED uses of *refugee* (where the naming strategy was applied seemingly congruently, without any intimation of problematisation) were extremely frequent, constituting the most common usage by a far margin. However, collocation and manual concordance analysis revealed that UNCONTESTED *refugees* are distinguished by lack of agency, performing little actions at all other than (often restricted or disallowed) movement. Water metaphors are employed once more, construing this group as boundary and order threats. Negative associations with the term gave rise to a CONTESTED discourse around *refugee*, wherein use of the naming strategy was problematized in the media. DEFINITIONS of *refugee* (some of which included geography of displacement) were employed to argue both for and against the term's application to this particular group. However, PARALLELISMS between 'Katrina's *refugees*' and *refugees* in other contexts, as well as a proclivity for referring to PAK triage areas as *refugee CAMPS*, draw relationships between PAKs and (often poor, non-white, somehow

'feral') others. Quite quickly in the corpus, CONTESTED instances overtook UNCONTESTED instances in frequency, and the word's meaning has been more sharply defined as a result. The press equating homelessness, destitution, ill health, and other manner of personal tragedy, with statelessness was ultimately deemed unacceptable by news creators and consumers. This indicates a hierarchy of othering, in which Americans (at the greatest risk, or having encountered the toughest circumstances) are placed at a level closer to 'us' than the 'distant other', or the foreigner.

In chapter 9, I studied the changing construal of gay people over time by analysing (in-)consistencies in semantic collocation over nearly three decades. The corpus was collected on the basis of inclusion of AIDS/HIV nodes rather than use of seed words related to gay people. This feature allowed me to track changes in the immediate associations of a given social group, as these did or did not directly relate to mentions of AIDS/HIV elsewhere in the article. Though I expected a gradual decline in othering discourses over time, the 1990s were more progressive as related to moral panics than the 2000s. Whereas the 1980s texts could be typified by discussion of risk, the 1990s were more dominated by discourses of political and social progress and problematisation of (individual and institutionalised) othering. By the 2000s, however, tropes had returned, with deviancy doubling and moral panics recreating social distance between the 'us' and 'them' group. This indicates that sexuality is very much still a point upon which moral panics turn, though specific risks—spreading of a disease, endangering the morality of children by adopting—may change.

### ***10.3 Discussion of methods***

In this section, I will offer some concluding thoughts on methodological choices and developments made throughout the thesis, including the use of: dual corpora and matched studies (10.3.1), a mixed methodological corpus-based discourse analytical approach (10.3.2), and proportional semantic categorization (10.3.3).

#### ***10.3.1 Dual corpora and matched studies***

From a qualitative, discourse analytical perspective, the implications of selecting dual corpora has brought both drawbacks and benefits. The main drawback is, of course, that the inclusion of two data sets has significantly increased the amount of time devoted to menial tasks (e.g. corpus design, data collection, cleansing, and indexing) and reduced the amount of space that might be dedicated to in-depth analysis of a single data set.

However, the aim of the thesis (from a discourse analytical perspective) was not to finely detail a finite number of features of social actor construal, but rather to discover whether shared discourse strategies could be found in generalizable patterns across more varied texts.

From a corpus linguistic perspective, that the corpora are so dissimilar in terms of size has made it necessary for me to be critical of cut-off points, particularly those related to minimum frequency thresholds and confidence statistics. Differences in diachrony between the two data sets have also led to creativity in viewing the data different ways, e.g. categorisation by part-of-speech versus semantic tag, incorporation of Venn diagrams in chapter 9 to indicate c-collocates, etc.

Had I just worked with a single data set, evolution of the method would not have been possible. Furthermore, comparison in patterns arising in both corpora (for instance, lack of agency, focus on the experiences of the in-group, water metaphors, perceived failure of institutional care) through matched studies would have been impossible. Though data collection and methods development was extremely challenging at times, I believe that the use of two corpora and attempts to loosely correspond analysis across chapters has ultimately been a strength of this thesis.

### ***10.3.2 Corpus-based critical discourse analysis***

The analytical sections of this thesis could not have been effectively accomplished without adoption of a mixed methodological approach. The corpora collected for this study tend toward the larger end of the spectrum, and would have been impossible to review, annotate, or markup manually. Chapters 4, 5, 7, and 8 adopt a more corpus-driven approach, with frequency figures identifying naming strategies for analysis. Chapters 6 and 9 are more corpus-based, with social actors selected on the basis of personal intuition. Investigation into the construal of naming strategies in all six analysis chapters relied upon collocation (first classified by part-of-speech tag, then by category of semantic tag). Semantic domains of collocates served as indicators for areas of further interest, and guided manual exploration.

Techniques from critical discourse analysis have guided and structured interpretation of collocation and manual concordance analysis. Of particular relevance have been as *topoi*, attribution, predication, transitivity analysis, “victim-agent reversal” and the “scapegoat

strategy”, and linguistic generalizations, analogies, and/or cross-mapped representation through metaphors (Matouschek et al., 1995).

Without the use of corpus linguistics tools and techniques, I would not have been able to make generalisations about frequency (or infrequency) of certain methods of conveying risk, in- and out-group qualities, or to determine whether patterns discussed had statistical significance, or were typical of general newspaper reportage of the era.

### ***10.3.3 Proportional semantic collocation***

The use of dual corpora and matched studies, and incorporation of a mixed methodological corpus-based critical discourse analytical approach, has enabled me to test and develop a method of downsampling collocates that has proven effective on corpora of differing sizes, spans of diachrony, and topics within the broad category of act of God reporting. Other popular methods of downsampling tend to consider sets of collocates as a whole, delimiting these with arbitrary cut-offs to do with rank, frequency, or statistical measure. However, this does not take into account a critical approach with perspective for frequent discourses. The Hurricane Katrina corpus was small enough—and the resulting collocates of each naming strategy relatively infrequent enough—to allow me to categorise collocates first by part of speech. This enabled in-depth analyses of transitivity, attribution, and predication strategies. The larger AIDS/HIV corpus was better considered through the primary scope of broad semantic category, where collocates could be viewed as proportions of a field of semantic preference.

Disregarding collocates with infrequent or uninteresting parts-of-speech or semantic categories allows a given researcher to devote a greater amount of his or her time to analysing only the most salient patterns of collocation, feeding directly onward to manual concordance analysis. If I had chosen, for instance, to view only the ‘top 50’ collocates of a given naming strategy, collocates contributing to a very important field but occurring relatively infrequently would have been disregarded in favour of collocates that were perhaps not as critical to construal of a given social group, when considered in the context of distribution across all of the semantic domains.

It was found that USAS chooses the correct candidate tag for the first position approximately 90% of the time. In 5% of cases, the most appropriate tag is found elsewhere in the string of candidates, and in only 5% of cases is the correct tag not found in the string or at all (i.e. the Z99: Unmatched semtag is assigned). Though I have chosen

to manually disambiguate tags to be absolutely certain of the categorisation and proportionality of semantic fields, this precision rate would be perfectly acceptable to many researchers dealing with large-scale corpora or unwieldy collocate lists. I recommend it for establishing semantic preference of search terms, and for contrasting semantic preference of one term to another, or investigating changes in semantic preference in a single term over time.

#### **10.4 Reflecting on the study**

In selecting the data for my corpora, I attempted to gather large samples that were representative of quality American newspaper reportage over time periods most salient for topical reporting. This involved utilisation of the pre-packaged Factiva 'Major U.S. print newspapers' set of publications. On the one hand, this has removed the chance of bias in researcher hand-selection of publications for inclusion; on the other, I have effectively eliminated local publications with much smaller circulations, and the enhanced variety of (perhaps more discriminatory) discourse that this could have achieved. Though I am satisfied with having collected data from newspapers with the highest circulation in the United States, I acknowledge that inclusion of *all* American newspapers, or a proportional split of data (e.g. of northern/southern publications), or even digital companions to print publications, would have likely brought a whole new dimension to the analysis.

Another shortcoming of using a news aggregator service for collection of diachronic corpora is imbalanced data population. Very early reporting on AIDS/HIV is sparsely represented in *Factiva*, either due to lack of digitisation efforts, or lack of interest in making these texts (perhaps parochial or offensive in their datedness) widely available. As such, the foundational discourses of the epidemic—and much of the most obliquely discriminatory discourses surrounding AIDS/HIV—are not represented in my corpus. These would have been an excellent addition to the data, particularly in the analysis of the evolution of othering. However, the blatant prejudice exhibited in early reporting has been the focus of many studies longer and larger than my own. In the absence of this data, I have focussed instead on more surreptitious traces of othering, for instance deviancy doubling and reframing through the powerful agents' experiences.

As texts were gathered from the *Factiva* database, multimodal data was all but lost; no information regarding the original font scheme, layout, accompanying pictures, etc., has been retained. The metadata indicating section and page number is erratically included

and oft unreliable. As such, this study has been purely textual, without reference to additional meanings created in visual imagery, and with all texts from various subgenres of reporting (e.g. editorial, obituary) being treated equally. Differences in genre could certainly have contributed to observed differences in the texts, though manually checking/amending metadata in the 208,540 texts included in the two corpora would have been prohibitive. It is also my opinion that ideologies in newspapers are built up incrementally across the publication as a whole rather than in cognitively partitioned sections, which has allowed me to justify considering them as unsegregated units.

There is no conformity in the use of collocation, the most fundamental corpus tool—researchers have not reached consensus on “whether or not collocations should be identified across sentence boundaries; whether collocation should be calculated on the basis of wordforms, lemmas or multi-word units; what the minimum frequency cut-off should be below which a co-occurring word will not be considered as a potential collocate; and whether collocates should be explored within one or many genres” (McEnery and Hardie, 2012, p. 129). Baker (2004), McCarthy and Carter (2002), Biber et al. (2004), and Mautner (2004), have all likewise noted the arbitrariness of cut-off points as limitations in their contemporary studies. Differentiation in researcher preference and practice in the application of these methodological tools can have drastic effects on the potential to compare multiple studies (McEnery and Hardie, 2012) or to triangulate results.

Relatedly, enforcing cut-offs of minimum collocation frequency has additionally affected my results. I have adjusted minimum collocation frequency in this thesis in accordance with the size of the corpus, the frequency of search words and terms, and the feasibility of analysing the full output list of collocates. I have, as such, sacrificed the maintenance of fully stringent thresholds standards across each chapter, in favour of total coverage of every single collocate of the highest significance and frequency in proportionally represented parts of speech or semantic domains. One guideline that I do recommend after having completed this analysis is experimentation with the interplay of log likelihood values with minimum collocation frequency. Maintaining a firm threshold of mutual information equal to or exceeding 3, I was able to determine when the minimum collocation frequency was high enough when each collocate returned was automatically also above the minimum log likelihood value of 10.83. This indicated that the frequency ‘floor’ was high enough to guarantee confidence while MI allowed an indication of strength.



Finally, the implementation of proportional semantic collocation as a way of downsampling results has necessarily eliminated some potentially interesting findings from inclusion in my analysis. Disregarding collocates from underrepresented semantic categories is a primary limitation that I intend to address in further research.

### **10.5 Further research**

The main focus of this research has been description of the most prominent features: the highest-frequency naming strategies, most proportionally represented semantic categories, most oft-repeated collocates. The description of such a variety of social actors construed in two separate corpora, and the tandem development of a method of dealing closely with large sets of data, has made views of the highs the most amenable ones to consider. However, I believe that this method could be very easily adapted to assist in the analysis of low-frequency features or rare discourses. By establishing a cut-off of the proportionally underrepresented semantic categories (e.g. those containing less than 3% of all collocates of a given item), researchers would be presented with a swift view of atypical, uncommon, and creative discourses surrounding their given concept. This would be of particular use for: 1) those working in fields with well-developed, current literature to avoid repetition of established results; 2) researchers interested in linguistic creativity or developing metaphors/idioms; 3) investigating infrequent uses or semantic senses of a given word.

Another avenue in which semantic collocation could be useful is corpus-based critical discourse analysis of the in-text construal of 'newsworthiness'. It is relatively straightforward, especially when using a part-of-speech tagged corpus (as I do, with CLAWS), to search for lexical items or grammatical classes indicating newsworthiness, e.g. adverbs such as *suddenly*. For instance: degree adverbs (e.g. *very, more, most*), comparative adverbs (*better, longer*), superlative adverbs (e.g. *best, longest*), and quasi-nominal adverbs of time (e.g. *now, tomorrow*) climb in frequency between August 2005, September 2005, and October 2005 in the Hurricane Katrina corpus, but drop off suddenly and steadily after this. This shows evidence of lexical indication of newsworthiness patterns corresponding to timeliness. However, enhancing this with semantic collocation, researchers could also search within diachronic subcorpora for markers of evaluation (USAS semtag A5), time (T1), or emotions including fear/bravery/shock (E5) or worry/concern/confidence (E6) for semantic clusters indicating changing conceptualisations of newsworthiness over time.

This is one method that I intend to experiment with in another area of further study: comparison of social actor construal between the Katrina corpus and other disaster narratives. During my PhD (in 2012), Hurricane Sandy formed and became the second-most destructive hurricane in the history of the United States (after Hurricane Katrina). Critically, the storm affected 24 states, with the most severe damage being done in New Jersey and New York, states that are much more northern, white, and affluent than Louisiana. Comparing the construal of social actors in the wake of such similar storms with such different groups of affected people would certainly provide interesting results to inform research on moral panics and risk society discourse.

At the outset of my PhD, my intention was to compare the construal of multiple 'risk groups' in AIDS/HIV over the full diachrony of the collected corpus. As the study developed and the findings in Chapter 9 expanded, it became clear that the scope of this project could not include such extensions. However, my immediate goal will be to extend this analysis to a contrastive study, to include a companion analysis of the changing construal of people who use intravenous drugs, as well as sex workers, haemophiliacs, and 'the distant other' (namely PWAs in Africa). It would be extremely interesting to trace the evolution of discourses of moral panic and risk society in the event of a shift in risk activities and period of reporting (e.g. haemophiliacs were mainly subjects in the 1980s, versus African people in the 2000s). Other research objectives included in the conception stage, but removed from the analysis portions, include cross-cultural comparisons (e.g. social actor representation in US versus UK texts) and enhanced analysis of publication preferences (particularly on northern/southern publication centre divisions).

Finally, Kanye West's effect on the discourse of Hurricane Katrina, and his (perhaps short-lived) popularization of the accusation of institutional uncaring, has struck a chord with me. A dedicated study could be carried out on this phrase alone, with the aim of discovering how often it is used with the rhetorical strategy of reproaching an institutional body while all but negating the possibility of retort. Texts dealing with lack of delivery on any commercial or social services would prove very illuminating.

## ***10.6 Concluding comments***

The research aims for this thesis were conceived of my own interest in critically detailing and problematizing discriminatory discourses in Hurricane Katrina and AIDS/HIV reporting. However, the aspects of othering that I hypothesised early in this

process to be the most salient (e.g. attribution, predication, explicit assignment of blame) were not the most surprising or influential in this sample of discourse.

I found myself surprised by the near total lack of outright bigotry in these texts, but also discomfited by the insidiousness of othering discourses present in its place. Where I expected to find accusations of sinning, looting, and shooting, instead I found deviancy doubling and social sequestering: ways by which readers and society members are enabled (and even encouraged) to disregard empathy and to consider themselves somehow above moral panic, or outside of risk. Having set out specifically to analyse the othering of social actors in the wake of acts of God, I have instead been confronted with majority discourses presenting in-group perspectives from the inside, glimpsing only rarely out. At the start of the study, I had perhaps not fully grasped the true domination of in-group experience as conveyed in institutional media. I believed that I would be able to study the depictions of behaviour of any social actors named in the corpus, but time and again, this has been stymied by reframing of the social events that were Katrina and AIDS as affecting non-PAKs or non-PWAs—members of the in-group.

The challenges of using two corpora and developing a method of downsampling have been worthwhile, in my opinion. I have been enabled to consider features discovered as potentially being part of a larger pattern rather than isolated instances relating to a single event. I believe that there is also a lack of critical discourse analytical work based entirely on 'broadsheet', reputable, or purportedly more 'objective' sources. Much more work is still being done on tabloid newspapers, political propaganda, and advertising, all of which take a much more explicit stance. Though patterns might not have been as visible to me as if I had chosen to compare broadsheets to tabloids, my findings are representative of 'trusted' press, and this is very worrying indeed.

Alongside depictions of culpability, risk, and mortality, something was very subtly being conveyed in the texts: a deep distance between those being reported on, and those being reported to. By failing to represent the people most affected by Katrina and AIDS as complex, important, dynamic figures who are central in their own stories, the press has continually squandered the opportunity to close the gulf between 'us' and 'them', and failed to provide the most newsy of the news. Instead, the press was seen to engender risk society discourse emphasising the risks of the few and disregarding perils of the many. Animosity and suspicion are just as disruptive to social order as epidemics and hurricanes; more risky than any act of God is indifference.

## References

- Adendorff, R., & de Klerk, V. (2009). The role of appraisal resources in discussing AIDS. In V. de Klerk (Ed.), *Corpus linguistics and world Englishes: An analysis of Xhosa English* (pp. 70–88). London: Continuum International Publishing Group.
- Adger, W. N. (2006). Vulnerability. *Global Environmental Change*, 16(3), 268–281.
- Adolphs, S., Knight, D., & Carter, R. (2011). Capturing Context for Heterogeneous Corpus Analysis: Some First Steps. *International Journal of Corpus Linguistics*, 16, 305–324.
- Archer, D., Wilson, A., & Rayson, P. (2002). Introduction to the USAS Category System. Retrieved from [http://ucrel.lancs.ac.uk/usas/usas\\_guide.pdf](http://ucrel.lancs.ac.uk/usas/usas_guide.pdf)
- Atkinson, J. M., & Heritage, J. (1984). *Structures of social action*. Cambridge: Cambridge University Press.
- Baker, P. (2004). Querying Keywords: Questions of Difference, Frequency, and Sense in Keywords Analysis. *Journal of English Linguistics*, 32(4), 346–359.
- Baker, P. (2005). *Public discourses of gay men*. London: Routledge.
- Baker, P. (2006). *Using corpora in discourse analysis*. London: Continuum International Publishing Group.
- Baker, P. (2010). Will Ms Ever be as Frequent as Mr?: A Corpus-based Comparison of Gendered Terms Across Four Diachronic Corpora of British English. *Gender and Language*, 4(1), 125–149.
- Baker, P., Gabrielatos, C., Khosravini, M., McEnery, T., & Wodak, R. (2008). A Useful Methodological Synergy? Combining Critical Discourse Analysis and Corpus Linguistics to Examine Discourses of Refugees and Asylum Seekers in the UK Press. *Discourse & Society*, 19(3), 273–306.
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse analysis and media attitudes: The representation of Islam in the British press*. Cambridge: Cambridge University Press.
- Baker, P., Hardie, A., & McEnery, T. (2006). *A glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Barnbrook, G. (1996). *Language and computers*. Edinburgh: Edinburgh University Press.
- Barnett, A., Hodgetts, D., Nikora, L., Chamberlain, K., & Karapu, R. (2007). Child Poverty, News Media Framing and the Symbolic Disempowerment of Families in Need. *Journal of Community and Applied Social Psychology*, 17, 296–312.
- Barnshaw, J., & Trainor, J. (2010). Race, class, and capital amidst the Hurricane Katrina diaspora. In D. L. Brunson, D. Overfelt, & J. S. Picou (Eds.), *Sociology of Katrina* (pp. 91–105). Plymouth, UK: Rowman & Littlefield.

- Bauman, R., & Briggs, C. (1990). Poetics and performance as critical perspectives on language and social life. *Annual Review of Anthropology*, 19, 59–88.
- Beck, U. (1992). *Risk society: Towards a new modernity*. (M. Ritter, Ed.) *Nation* (Vol. 2, p. 260). Sage.
- Bednarek, M., & Caple, H. (2012). "Value Added": Language, Image and News Values. *Discourse, Context & Media*, 1(2-3), 103–113.
- Belkin, A. (2008). 'Don't Ask, Don't Tell': Does the Gay Ban Undermine the Military's Reputation? *Armed Forces and Society*, 34(2), 276–291.
- Belkin, A. (2013). Combat Exclusion RIP. Will Patriarchy's Demise Follow? *Critical Studies on Security*, 1(2), 249–250.
- Bell, A. (1991). *The language of news media*. Oxford: Basil Blackwell.
- Bell, A., & Garrett, P. (Eds.). (1998). *Approaches to media discourse*. New York: Wiley-Blackwell.
- Berger, P., & Luckman, T. (1967). *The social construction of reality*. Harmondsworth: Penguin.
- Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8(4), 243–257.
- Biber, D., & Burges, J. (2001). Historical shifts in the language of women and men: Gender differences in dramatic dialogue. In D. Biber & S. Conrad (Eds.), *Variation in English: Multi-dimensional studies* (pp. 157–170). London: Longman.
- Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic Prose. In H. Hasselgård, S. Johansson, & S. Oksefjell (Eds.), *Out of corpora: Studies in honour of Stig Johansson* (pp. 181–190). Amsterdam: Rodopi.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Bishaw, A., & Iceland, J. (2003). *Poverty: 1999 (Census 2000 Brief)*. Washington, D.C.
- Blaikie, P., Cannon, T., Davis, I., & Wisner, B. (1994). *At risk: Natural hazards, people's vulnerability and disasters*. London: Routledge.
- Blommaert, J. (2005). *Discourse: A critical introduction*. Cambridge: Cambridge University Press.
- Bondi, M. (2007). Key-words and emotions: A case study of the Bloody Sunday Enquiry. In N. Fairclough, G. Cortese, & P. Ardizzone (Eds.), *Discourse and contemporary social change* (pp. 407–432). Oxford: Peter Lang.
- Bourdieu, P. (1989). Social Space and Symbolic Power. *Sociological Theory*, 7, 14–25.

- Brunsmas, D L, Overfelt, D., & Picou, J. S. (2007). *The sociology of Katrina: Perspectives on a modern catastrophe*. New York: Rowman & Littlefield Publishers.
- Burr, V. (1995). *An introduction to social constructionism*. *The British Journal of Sociology* (Vol. 47, p. 198).
- Calavita, K. (1984). *U.S. immigration law and the control of labor, 1820-1924*. London: Academic Press.
- Caldas-Coulthard, C. (1996). Women who pay for sex. And enjoy it. In C. R. Caldas-Coulthard & M. Coulthard (Eds.), *Practices: Readings in Critical Discourse Analysis* (pp. 250–270). London: Routledge.
- Caldas-Coulthard, C. R. (1993). From discourse analysis to Critical Discourse Analysis: The differential re-representation of women and men speaking in written news. In J. Sinclair, M. Hoey, & G. Fox (Eds.), *Techniques of description: Spoken and written discourse* (pp. 196–208). London: Routledge.
- Caldas-Coulthard, C. R. (1995). Man in the news: The misrepresentation of women speaking in news as narrative discourse. In S. Mills (Ed.), *Language and gender: Interdisciplinary perspectives* (pp. 226–239). London: Longman.
- Caldas-Coulthard, C. R., & Moon, R. (2010). Curvy, Hunky, Kinky: Using Corpora as Tools in Critical Analysis. *Discourse & Society*, 21(2), 1–35.
- Center for Disease Control. (1982). *Morbidity and Mortality Weekly Report: Current Trends Update on Acquired Immune Deficiency Syndrome (AIDS) -- United States*. *Morbidity and Mortality Weekly Report*. Atlanta, GA. Retrieved from <http://www.cdc.gov/mmwr/preview/mmwrhtml/00001163.htm>
- Center for Disease Control. (2008). *HIV Transmission Rates in the United States, 1977–2006* (pp. 1–3). Atlanta, GA. Retrieved from [http://www.cdc.gov/nchhstp/newsroom/docs/CDC\\_TransmissionRates\\_FactSheet.pdf](http://www.cdc.gov/nchhstp/newsroom/docs/CDC_TransmissionRates_FactSheet.pdf)
- Center for Disease Control. (2011). *Diagnoses of HIV Infection in the United States and Dependent Areas*. Atlanta, GA. Retrieved from [http://www.cdc.gov/hiv/library/reports/surveillance/2011/surveillance\\_Report\\_vol\\_23.html](http://www.cdc.gov/hiv/library/reports/surveillance/2011/surveillance_Report_vol_23.html)
- Charteris-Black, J. (2004). *Corpus approaches to Critical Metaphor Analysis*. New York: Palgrave MacMillan.
- Charteris-Black, J. (2005). *Politicians and rhetoric: The persuasive power of metaphor*. London: Palgrave-MacMillan.
- Charteris-Black, J. (2006). Britain as a container: Immigration metaphors in the 2005 election campaign. *Discourse & Society*, 17(5), 563–581.

- Charteris-Black, J. (2009). Metaphor and gender in British parliamentary debates. In K. Ahrens (Ed.), *Politics, gender and conceptual metaphor* (pp. 139–165). Basingstoke, Hampshire: Palgrave MacMillan.
- Charteris-Black, J. (2012). Shattering the Bell Jar: Metaphor, Gender and Depression. *Metaphor and Symbol, 27*(3), 199–216.
- Chibnall, S. (1977). *Law and order news: An analysis of crime reporting in the British press*. London: Tavistock.
- Chilton, P. (2005). Missing links in mainstream CDA: Modules, blends and the critical instinct. In R. Wodak & P. Chilton (Eds.), *A new agenda in (Critical) Discourse Analysis: Theory, methodology and interdisciplinarity* (pp. 19–51). Amsterdam: John Benjamins Publishing.
- Chouliaraki, L. (2008). The Media as Moral Education: Mediation and Action. *Media, Culture & Society, 30*(6), 831–852.
- Cohen, S. (1972). *Folk devils and moral panics: The creation of the Mods and Rockers*. London: Routledge.
- Collins Cobuild Dictionary Online. (2011). Retrieved November 28, 2011, from <http://www.collinsdictionary.com/>
- Conrad, P. (1986). The Social Meaning of AIDS. *Social Policy, 16*(1), 51–56.
- Costello, R. B. (Ed.). (2000). *Random House Webster's college dictionary* (2nd ed., p. 1536). New York: Random House Reference.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly, 34*(2), 213–238.
- Coxhead, A. (2008). Phraseology and English for academic purposes: Challenges and opportunities. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 149–161). Amsterdam: John Benjamins Publishing Company.
- Coxhead, A., & Byrd, P. (2007). Preparing Writing Teachers to Teach the Vocabulary and Grammar of Academic Prose. *Journal of Second Language Writing, 16*(3), 129–147.
- Curran, D. (2013). Risk Society and the Distribution of Bads: Theorizing Class in the Risk Society. *The British Journal of Sociology, 64*(1), 44–62.
- Curtius, R. E. (1990). *European literature and the Latin Middle Ages*. Princeton, NJ: Princeton University Press.
- Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social Vulnerability to Environmental Hazards. *Social Science Quarterly, 84*(2), 242–261.
- Davies, M. (2009). The 385+ Million Word Corpus of Contemporary American English (1990–2008+): Design, Architecture, and Linguistic Insights. *International Journal of Corpus Linguistics, 14*(2), 159–190.

- De Fina, A., Schiffrin, D., & Bamberg, M. (2006). Introduction. In *Discourse and identity* (pp. 1–23). Cambridge: Cambridge University Press.
- Delanty, G., Wodak, R., & Jones, P. (Eds.). (2011). *Migration, identity and belonging*. Liverpool: Liverpool University Press.
- Dow Jones Factiva. (2009). Inside-Out: Complete End-user Reference Guide for Academic Dow Jones Factiva.com. *Dow Jones Factiva Inside-Out Reference Guide*. Retrieved from [http://cesdoc.univ-paris1.fr/fileadmin/CESDOC/BDD/Guides\\_BDD/factiva\\_user\\_guide.pdf](http://cesdoc.univ-paris1.fr/fileadmin/CESDOC/BDD/Guides_BDD/factiva_user_guide.pdf)
- Duguid, A. (2007). Men at work: How those at Number 10 construct their working identity. In G. Garzone & S. Sarangi (Eds.), *Discourse, ideology and specialized communication* (pp. 453–484). Bern: Peter Lang.
- Eadie, J. (2004). *Sexuality: The essential glossary*. Bloomsbury USA.
- Eggs, S. (2004). *An introduction to systemic functional linguistics*. (M A K Halliday & J. J. Webster, Eds.) *Journal of SocioLinguistics* (Vol. 10, p. 404). Continuum International Publishing Group.
- Fairchild, H. P. (1926). *Immigration: A world movement and its American significance*. New York: Macmillan.
- Fairclough, N. (1995). *Media discourse*. London: Bloomsbury Academic.
- Fairclough, N. (2000). *New Labour, new language?* London: Routledge.
- Fee, E., & Krieger, N. (1993). Thinking and Rethinking AIDS: Implications for Health Policy. *International Journal of Health Services*, 23(2), 323–46.
- Firth, J. R. (1957). *Papers in linguistics, 1934-1951*. Oxford: Oxford University Press.
- Fiske, J., & Hartley, J. (1978). *Reading television*. London: Routledge.
- Fothergill, A. (2003). The Stigma of Charity: Gender, Class, and Disaster Assistance. *The Sociological Quarterly*, 44(4), 659–680.
- Foucault, M. (1980). *Language, counter-memory, practice: Selected essays and interviews*. (D. F. Bouchard, Ed.). Cornell, NY: Cornell University Press.
- Fowler, R. (1991). *Language in the news: Discourse and ideology in the British press*. London: Routledge.
- Fowler, R., Hodge, B., Kress, G., & Trew, T. (1979). *Language and control*. London: Routledge.
- Fries, P. H., & Gregory, M. (1995). *Discourse in society: Systemic Functional perspectives*. New York: Ablex.



- Fussell, E., Sastry, N., & VanLandingham, M. (2010). Race, Socioeconomic Status, and Return Migration to New Orleans After Hurricane Katrina. *Population and Environment*, 31(1-3), 20–42.
- Gabrielatos, C., & Baker, P. (2008). Fleeing, Sneaking, Flooding: A Corpus Analysis of Discursive Constructions of Refugees and Asylum Seekers in the UK Press, 1996-2005. *Journal of English Linguistics*, 36(1), 5–38.
- Galtung, J., & Ruge, M. H. (1965). The Structure of Foreign News: The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. *Journal of Peace Research*, 2(1), 64–90.
- Galtung, J., & Ruge, M. H. (1973). Structuring and selecting news. In S. Cohen & J. Young (Eds.), *The manufacture of news: A reader* (pp. 52–63). London: SAGE Publications.
- Goldthorpe, J. H. (2002). Globalisation and Social Class. *West European Politics*, 25(3), 1–28.
- Goode, E., & Ben-Yehuda, N. (2010). *Moral panics: The social construction of deviance*. Hoboken, NJ: Wiley Blackwell.
- Greatbatch, D. (1986). Aspects of Topical Organization in News Interviews: The Use of Agenda-shifting Procedures by Interviewees. *Media Culture & Society*, 8(4), 441–455.
- Greenslade, R. (2005). *Seeking Scapegoats: The Coverage of Asylum in the UK Press*. London. Retrieved from [www.ippr.org.uk/ecomm/files/wp5\\_scapegoats.pdf](http://www.ippr.org.uk/ecomm/files/wp5_scapegoats.pdf)
- Grieco, E. M., & Cassidy, R. C. (2001). *Overview of Race and Hispanic Origin (Census 2000 Brief)* (pp. 1–11). Washington, D.C.
- Grmek, M. D., Maulitz, R. C., & Duffin, J. (1990). *History of AIDS: Emergence and origin of a modern pandemic*. Princeton, NJ: Princeton University Press.
- Grue, J. (2009). Critical Discourse Analysis, Topoi and Mystification: Disability Policy Documents from a Norwegian NGO. *Discourse Studies*, 11(3), 305–328.
- Haarman, L., & Lombardo, L. (2009). Introduction. In L. Haarman & L. Lombardo (Eds.), *Evaluation and stance in war news: A linguistic analysis of American, British, and Italian television news reporting of the 2003 Iraqi War* (pp. 1–26). London: Continuum International Publishing Group.
- Habermas, J. (1971). Knowledge and human interests: A general perspective. In *Knowledge and human interests* (pp. 301–317).
- Hall, H. I., Song, R., Rhodes, P., Prejean, J., An, Q., Lee, L. M., ... Janssen, R. S. (2008). Estimation of HIV Incidence in the United States. *Journal of the American Medical Association*, 300(5), 520–529.
- Hall, J. K. (1996). Who needs “identity”? In S. Hall & P. du Gay (Eds.), *Questions of cultural identity* (pp. 1–17). London: SAGE Publications.

- Hall, Stuart. (1997). *Representation: Cultural representations and signifying practices*. (Stuart Hall, Ed.). London: SAGE Publications Ltd.
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Maryland: University Park Press.
- Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Edward Arnold.
- Hardie, A. (2012). CQPweb: Combining Power, Flexibility and Usability in a Corpus Analysis Tool. *International Journal of Corpus Linguistics*, 17(3), 380–409.
- Hartley, J. (1982). *Understanding news*. London: Routledge.
- Herek, G. M., Capitanio, J. P., & Widaman, K. F. (2002). HIV-related Stigma and Knowledge in the United States: Prevalence and Trends, 1991-1999. *American Journal of Public Health*, 92(3), 371–377.
- Heritage, J. (1985). Analysing news interviews: Aspects of the production of talk for an “overhearing” audience. In T. A. van Dijk (Ed.), *Handbook of Discourse Analysis, vol. 3: discourse and dialogue*. London: Academic Press.
- Hodge, R., & Kress, G. (1979). *Language as ideology*. London: Routledge.
- Hodge, R., & Kress, G. (1988). *Social semiotics*. Cambridge: Polity Press.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hunston, S., & Francis, G. (2000). *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins Publishing Company.
- Hurley, S. F., Jolley, D. J., & Kaldor, J. M. (1997). Effectiveness of Needle-exchange Programmes for Prevention of HIV Infection. *The Lancet*, 349(9068), 1797–1800.
- Hutchby, I. (1996). Power in Discourse: The Case of Arguments on a British Talk Radio Show. *Discourse & Society*, 7(4), 481–497.
- Hymes, K. B., Greene, J. B., Marcus, A., William, D. C., Cheung, T., Prose, N. S., ... Laubenstein, L. J. (1981). Kaposi's Sarcoma in Homosexual Men: A Report of Eight Cases. *The Lancet*, 318(8247), 598–600.
- Ibarrán, M. E., & Ruth, M. (2009). Climate change and natural disasters: Economic and distributional impacts. In M. Ruth & M. E. Ibarrán (Eds.), *Distributional impacts of climate change and disasters: Concepts and cases*. Cheltenham: Edward Elgar.
- Ivie, R. L. (1980). Images of Savagery in American Justifications of War. *Communication Monographs*, 47, 279–294.

- Johnson, K. a., Dolan, M. K., Johnson, L., Sonnett, J., & Reppen, R. (2010). Interjournalistic Discourse about African Americans in Television News Coverage of Hurricane Katrina. *Discourse & Communication*, 4(3), 243–261.
- Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.
- Kennedy, G. (1991). Between and through: The company they keep and the functions they serve. In K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics: Studies in honour of Jan Svartvik* (pp. 95–110). London: Longman.
- KhosraviNik, M. (2010). The Representation of Refugees, Asylum Seekers and Immigrants in British Newspapers: A Critical Discourse Analysis. *Journal of Language and Politics*, 9(1), 1–28.
- Kienpointner, M. (1992). *Alltagslogik: Struktur und Funktion von Argumentationsmustern*. Stuttgart-Bad Cannstatt: Frommann-Holzboog.
- Knabb, R. D., Rhome, J. R., & Brown, D. P. (2005). *Tropical Cyclone Report: Hurricane Katrina*. National Hurricane Center. Miami, FL.
- Koller, V. (2004). *Metaphor and gender in business media discourse: A critical cognitive study*. Basingstoke: Palgrave MacMillan.
- Krishnamurthy, R. (1996). Ethnic, racial and tribal: The language of racism? In C. Caldas-Coulthard & M. Coulthard (Eds.), *Texts and practices: Readings in Critical Discourse Analysis* (pp. 129–149). London: Routledge.
- Kroskrity, P. (2000). Identity. *Journal of Linguistic Anthropology*, 9(1-2), 111–114.
- Krzyzanowski, M., Triandafyllidou, A., & Wodak, R. (2009). Introduction. In A. Triandafyllidou, M. Krzyzanowski, & R. Wodak (Eds.), *The European public sphere and the media: Europe in crisis* (pp. 1–12). Basingstoke: Palgrave MacMillan.
- Krzyzanowski, M., & Wodak, R. (2009). *The politics of exclusion: Debating migration in Austria*. Piscataway, NJ: Transaction Publishers.
- L'Hote, E. (2010). New Labour and Globalization: Globalist Discourse with a Twist? *Discourse & Society*, 21(4), 355–376.
- Lacy, M. G., & Haspel, K. C. (2011). Apocalypse: The media's framing of black looters, shooters, and brutes in Hurricane Katrina's aftermath. In M. G. Lacy (Ed.), *Critical rhetorics of race* (pp. 21–46). New York: New York University Press.
- Lakoff, G. (1979). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.). Cambridge: Cambridge University Press.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago: University of Chicago Press.
- Leech, G. (1992). Corpora and theories of linguistic performance. In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991* (pp. 105–122). Berlin: Mouton de Gruyter.
- Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4: The Tagging of the British National Corpus. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING 94)* (pp. 622–628). Kyoto, Japan.
- Livingstone, S. M. (1998). *Making sense of television: The psychology of audience interpretation. International series in social psychology*. London.
- Lloyd, A. (1995). *Doubly deviant, doubly damned: Society's treatment of violent women*. London: Penguin.
- Logan, J. R. (2006). *The Impact of Katrina: Race and Class in Storm-damaged Neighborhoods*. New Orleans, LA.
- Longman Dictionary of Contemporary English Online. (2011). Retrieved November 28, 2011, from <http://www.ldoceonline.com/>
- Louisiana Department of Health and Hospitals. (2006). *Reports of Missing and Deceased*. Retrieved from <http://www.dhh.louisiana.gov/offices/page.asp?ID=192&Detail=5248>
- Louw, W. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 157–176). Amsterdam: John Benjamins.
- Lupton, D. (1993). AIDS Risk and Heterosexuality in the Australian Press. *Discourse & Society*, 4(3), 307–328.
- Malkki, L. H. (1995). *Purity and exile: Violence, memory, and national cosmology among hutu refugees in tanzania*. Chicago: University of Chicago Press.
- Marchi, A., & Venuti, M. (2009). Mark-up and the narrative structure of television news. In L. Haarman & L. Lombardo (Eds.), *Evaluation and stance in war news: A linguistic analysis of American, British, and Italian television news reporting of the 2003 Iraqi War* (pp. 27–47). London: Continuum International Publishing Group.
- Mårdh, I. (1980). *Headlines: On the grammar of English front page headlines*. Lund: Liberläromedel/Gleerup.
- Masquelier, A. (2006). Why Katrina's Victims Aren't Refugees: Musings on a "Dirty" Word. *American Anthropologist*, 108(4), 735–743.

- Matouschek, B., Januschek, F., & Wodak, R. (1995). *Notwendige Massnahmen gegen Fremde? [Necessary measures against foreigners?]*. Vienna: Passagen Verlag.
- Mautner, G. (1995). Only Connect: Critical Discourse Analysis and Corpus Linguistics. In *UCREL Technical Paper 6*. Lancaster.
- Mautner, G. (2007). Mining Large Corpora for Social Information: The Case of Elderly. *Language in Society*, 36(1), 51–72.
- Mautner, G. (2009). Checks and balances: How Corpus Linguistics can contribute to CDA. In R. Wodak & M. Meyer (Eds.), *Methods of Critical Discourse Analysis* (2nd ed., pp. 122–143). London: SAGE Publications.
- Mautner, G., & Koller, V. (2004). Computer applications in Critical Discourse Analysis. In C. Coffin, A. Hewings, & K. O'Halloran (Eds.), *Applying English grammar: Functional and corpus approaches* (pp. 216–228). London: Hodder and Stoughton.
- McArthur, T. (1981). *Longman lexicon of contemporary English*. London: Longman.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics* (2nd ed.). Edinburgh: Edinburgh University Press.
- McEnery, T., Xiao, R., & Tono, Y. (2005). *Corpus-based language studies: An advanced resource book*. London: Routledge.
- Menz, F., Wodak, R., Gruber, H., & Lutz, B. (1988). Power Struggles in the Media: A Case Study. *Folia Linguistica*, 23, 439–457.
- Merriam Webster Dictionary Online. (2013). Retrieved from <http://www.merriam-webster.com/>
- Meyer, M. (2001). Between theory, method, and politics: Positioning of the approaches to CDA. In R. Wodak & M. Meyer (Eds.), *Methods of Critical Discourse Analysis* (pp. 14–32). London: SAGE Publications.
- Moscovici, S. (1982). The coming era of representations. In J. Codol & J. Leyens (Eds.), *Cognitive analysis of social behaviour*. The Hague: Martinus Nijhoff.
- Mythen, G. (2005a). Employment, Individualization and Insecurity: Rethinking the Risk Society Perspective. *The Sociological Review*, 53(1), 129–149.
- Mythen, G. (2005b). From “Goods” to “Bads”? Revisiting the Political Economy of Risk. *Sociological Research Online*, 10(3).
- O'Brien, G. V. (2003). Indigestible Food, Conquering Hordes, and Waste Materials: Metaphors of Immigrants and the Early Immigration Restriction Debate in the United States. *Metaphor and Symbol*, 18(1), 33–47.

- Office of the United Nations High Commissioner for Human Rights. (1967). *Status and Protocol Relating to the Status of Refugees*. Retrieved from <http://legal.un.org/avl/ha/prsr/prsr.html>
- Oxford English Dictionary Online. (2011). Retrieved November 28, 2011, from <http://www.oed.com/>
- Palmié, S. (2007). Genomics, Divination, "Racecraft." *American Ethnologist*, 34, 205–222.
- Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching*. Amsterdam: John Benjamins Publishing Company.
- Partington, A. (2003). *The linguistics of political argument: The spin-doctor and the wolf pack at the White House*. London: Routledge.
- Pearce, L., & Wodak, R. (2010). The Construction of Regional Identity in Europe and Beyond. *European Journal of Cultural Studies*, 13(1), 1–7.
- Pearce, M. (2008). Investigating the Collocational Behaviour of MAN and WOMAN in the British National Corpus Using Sketch Engine. *Corpora*, 3(1).
- Pearsall, J., & Hanks, P. (Eds.). (1998). *The New Oxford Dictionary of English*. Oxford: Oxford University Press.
- Perry, I. (2011). *More beautiful and more terrible: The embrace and transcendence of racial inequality in the United States*. New York University Press.
- Plyer, A. (2011). *What Census 2010 Reveals about Population and Housing in New Orleans and the Metro Area*. New Orleans, LA. Retrieved from [https://gnocdc.s3.amazonaws.com/reports/GNOCDC\\_Census2010PopulationAndHousing.pdf](https://gnocdc.s3.amazonaws.com/reports/GNOCDC_Census2010PopulationAndHousing.pdf)
- Prilleltensky, I. (2003). Poverty and power: Suffering and wellness in collective, relational, and personal domains. In T. Sloan & S. Carr (Eds.), *Community approaches to poverty* (pp. 19–44). Dordrecht: Kluwer/Plenum.
- Rayson, P. (2003). *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*. Lancaster University.
- Rayson, P. (2008). From Key Words to Key Semantic Domains. *International Journal of Corpus Linguistics*, 13(4), 519–549.
- Rayson, P., Archer, D., Piao, S. T., & McEnery, T. (2004). The UCREL Semantic Analysis System. In *Beyond Named Entity Recognition Semantic labelling for NLP tasks in association with 4th International Conference on Language Resources and Evaluation (LREC 2004)* (pp. 7–12). Lisbon, Portugal.
- Rayson, P., Berridge, D., & Francis, B. (2004). Extending the Cochran Rule or the Comparison of Word Frequencies Between Corpora. In G. Purnelle, C. Fairon, & A. Dister (Eds.), *Le Poids des Mots: Proceedings of the 7th International Conference on*

- Statistical Analysis of Textual Data (JADT 2004)* (pp. 926–936). Louvain-la-Neuve, Belgium: Presses universitaires de Louvain.
- Redefining Progress. (2004). *African Americans and Climate Change: An Unequal Burden*. Oakland, CA. Retrieved from [http://www.sustainlex.org/BlackCaucusfullCBCF\\_REPORT\\_F.pdf](http://www.sustainlex.org/BlackCaucusfullCBCF_REPORT_F.pdf)
- Reisigl, M., & Wodak, R. (2001). *Discourse and discrimination: Rhetorics of racism and antisemitism*. Routledge.
- Reisigl, M., & Wodak, R. (2009). The Discourse-Historical Approach (DHA). In R. Wodak & M. Meyer (Eds.), *Methods for critical discourse analysis* (2nd ed., pp. 87–121). London: SAGE Publications.
- Richardson, J. (2004). *(Mis)representing Islam: The racism and rhetoric of British broadsheet newspapers*. Amsterdam: John Benjamins Publishing Company.
- Richardson, J., & Wodak, R. (2009). The Impact of Visual Racism: Visual Arguments in Political Leaflets of Austrian and British Far-right Parties. *Controversia*, 6(2), 45–77.
- Rodriguez, H. (2006). Rising to the Challenges of a Catastrophe: The Emergent and Prosocial Behavior following Hurricane Katrina. *The ANNALS of the American Academy of Political and Social Science*, 604(1), 82–101.
- Scarborough Research. (2009). *Newspaper Readership Report*. Scarborough, NY.
- Scott, M. (2002). Picturing the key words of a very large corpus and their lexical upshots – or getting at the Guardian’s view of the world. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 43–50). Amsterdam: Rodopi.
- Seale, C., & Charteris-Black, J. (2008). The Interaction of Age and Gender in Illness Narratives. *Ageing and Society*, 28(7), 1025–1045.
- Sedgwick, E. K. (1990). *Epistemology of the closet*. Berkeley, CA: University of California Press.
- Sigley, R., & Holmes, J. (2002). Girl-watching in Corpora of English. *Journal of English Linguistics*, 30(2), 138–157.
- Sinclair, J. M. (1991). *Corpus, concordance, collocation: Describing English language*. Oxford: Oxford University Press.
- Sinclair, J. M. (1996). *EAGLES: Preliminary recommendations on Corpus Typology*. Retrieved from <http://www.ilc.cnr.it/EAGLES/corpusyp/corpusyp.html>
- Sinclair, J. M. (1998). The lexical item. In Edda Weigand (Ed.), *Contrastive Lexical Semantics* (pp. 1–24). Amsterdam: John Benjamins.
- Siplon, P. D. (2002). *AIDS and the policy struggle in the United States*. Washington, D.C.: Georgetown University Press.

- Skeggs, B. (2004). *Class, self, culture*. London: Routledge.
- Smith, R. A. (1998). *Encyclopedia of AIDS: A social, political, cultural and scientific record of the HIV epidemic*. New York: Taylor & Francis.
- Sontag, S. (1989). *AIDS and its metaphors*. New York: Farrar, Straus and Giroux.
- Stubbs, M. (1996). *Text and corpus analysis*. Oxford: Blackwell Publishing.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. London: Blackwell Publishing.
- Stubbs, M., & Gerbig, A. (1993). Human and inhuman geography: On the computer-assisted analysis of long texts. In M. Hoey (Ed.), *Data, Description, Discourse: Papers on the English Language in honour of John Sinclair on his Sixtieth Birthday* (pp. 64–85). London: Harper Collins.
- Summers, D. (Ed.). (2003). *Longman dictionary of contemporary English: The living dictionary* (4th ed.). Essex, U.K.: Longman.
- Sutton, M. Y., Jones, R. L., Wolitski, R. J., Cleveland, J. C., Dean, H. D., & Fenton, K. A. (2009). A Review of the Centers for Disease Control and Prevention's Response to the HIV/AIDS Crisis Among Blacks in the United States, 1981-2009. *American Journal of Public Health*, 99(S2), S351–S359.
- Svartvik, J. (1992). Corpus Linguistics comes of age. In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991* (pp. 7–14). Berlin: Walter de Gruyter.
- Talbot, M. (2007). *Media discourse: Representation and interaction*. Edinburgh: Edinburgh University Press.
- Teo, P. (2000). Racism in the news: A Critical Discourse Analysis of news reporting in two Australian newspapers. *Discourse & Society*, 11(1), 7–49.
- The Pew Research Center. (2012). *Trends in News Consumption: 1991-2012*. Washington, D.C. Retrieved from [http://www.people-press.org/files/legacy-pdf/2012 News Consumption Report.pdf](http://www.people-press.org/files/legacy-pdf/2012%20News%20Consumption%20Report.pdf)
- Tierney, K., Bevc, C., & Kuligowski, E. (2006). Metaphors Matter: Disaster Myths, Media Frames, and Their Consequences in Hurricane Katrina. *The ANNALS of the American Academy of Political and Social Science*, 604(1), 57–81.
- Tierney, Kathleen. (2008). Hurricane Katrina: Catastrophic impacts and alarming lessons. In J. M. Quigley & L. A. Rosenthal (Eds.), *Risking house and home: Disasters, cities, public policy* (pp. 119–136). Berkeley, CA: Berkeley Public Policy Press.
- Titscher, S., Meyer, M., Wodak, R., & Vetter, E. (2000). *Methods of text and discourse analysis*. London: SAGE Publications.



- Tognini-Bonelli, E. (2001). *Corpus Linguistics at work*. Amsterdam: John Benjamins Publishing Company.
- Treichler, P. A. (1987). AIDS, Homophobia and Biomedical Discourse: An Epidemic of Signification. *Cultural Studies*, 1(3), 263–305.
- Tyrwhitt-Drake, H. C. (1999). Resisting the Discourse of Critical Discourse Analysis: Reopening a Hong Kong Case Study. *Journal of Pragmatics*, 31, 1081–1088.
- U.S. Census Bureau. (2004a). *2004 American Community Survey: New Orleans City Economic Characteristics*. Washington, D.C. Retrieved from <http://www.columbia.edu/itc/journalism/cases/katrina/FederalGovernment/CensusBureau/DemographicData/Data/NewOrleans/OrleansEconomicCharacteristics.pdf>
- U.S. Census Bureau. (2004b). *2004 American Community Survey: New Orleans Social Characteristics*. Washington, D.C. Retrieved from <http://www.columbia.edu/itc/journalism/cases/katrina/FederalGovernment/CensusBureau/DemographicData/Data/NewOrleans/OrleansSocialCharacteristics.pdf>
- U.S. Census Bureau. (2009). *New Orleans was Nation's Fastest-growing City in 2008: Population Getting Closer to Pre-Katrina Levels*. Washington, D.C. Retrieved from <http://www.census.gov/Press-Release/www/releases/archives/population/013960.html>
- U.S. Commission on Civil Rights. (2009). *Racial Categorization in the 2010 Census*, (March 2009). Retrieved from [http://www.usccr.gov/pubs/RC2010Web\\_Version.pdf](http://www.usccr.gov/pubs/RC2010Web_Version.pdf)
- Ungar, S. (2001). Moral Panic Versus the Risk Society: The Implications of the Changing Sites of Social Anxiety. *The British journal of sociology*, 52(2), 271–291.
- Ungerer, F. (2000). *English media texts – Past and present: Language and textual structure*. Amsterdam: John Benjamins Publishing Company.
- Van Dijk, T. A. (1988). *News as discourse*. London: Routledge.
- Van Dijk, T. A. (1991). *Racism and the press*. London: Routledge.
- Van Dijk, T. A. (1993). Principles of Critical Discourse Analysis. *Discourse & Society*, 4(2), 249–283.
- Van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. London: SAGE Publications.
- Van Dijk, T. A. (2005). *Discourse and racism in Spain and Latin America*. Amsterdam: John Benjamins Publishing Company.
- Van Dijk, T. A. (2006). *Discourse, context and cognition*. *Discourse Studies* (Vol. 8, pp. 159–177). London: SAGE Publications.

- Van Dijk, T. A. (2008). *Discourse and power*. New York: Palgrave Macmillan.
- Van Dijk, T. A. (2009). Critical Discourse Studies: A sociocognitive approach. In R. Wodak & M. Meyer (Eds.), *Methods of Critical Discourse Analysis* (pp. 62–86). London: SAGE Publications.
- Van Leeuwen, T. (2005). *Introducing Social Semiotics*. London: Routledge.
- Van Leeuwen, T. (2008). *Discourse and practice: New tools for critical analysis*. Oxford: Oxford University Press.
- VanLandingham, M. J. (2007). Murder Rates in New Orleans, LA, 2004–2006. *American Journal of Public Health, 97*(9), 1614–1616.
- Voorhees, C. C. W., Vick, J., & Perkins, D. D. (2007). “Came Hell and High Water”: The Intersection of Hurricane Katrina, the News Media, Race and Poverty. *Journal of Community & Applied Social Psychology, 17*, 415–429.
- White, P. (1997). Death, disruption and the moral order: the narrative impulse in mass-media hard news reporting. In F. Christie & J. R. Martin (Eds.), *Genres and institutions: Social processes in the workplace and school* (pp. 101–133). London: Cassell.
- Widdowson, H. G. (2000). On the Limitations of Linguistics Applied. *Applied Linguistics, 21*(1), 3–5.
- Widdowson, H. G. (2004). *Text, context, pretext: Critical issues in discourse analysis*. Oxford: Blackwell Publishing.
- Wijkman, A., & Timberlake, L. (1984). *Natural disasters: Acts of God or acts of man?* London: Earthscan.
- Wodak, R. (2001). The Discourse Historical Approach. In *Methods of Critical Discourse Analysis* (1st ed., pp. 63–93). London: SAGE Publications.
- Wodak, R. (2003). Populist Discourses: The Rhetoric of Exclusion in Written Genres. *Document Design, 4*(2), 132–148.
- Wodak, R. (2009). Language and politics. In J. Culpeper, F. Katamba, P. Kerswill, R. Wodak, & T. McEnery (Eds.), *English language: Description, variation and context* (pp. 576–593). Basingstoke: Palgrave MacMillan.
- Wodak, R. (2011). Politics as usual: Investigating political discourse in action. In J. Gee & M. Handford (Eds.), *The Routledge handbook of discourse analysis* (pp. 525–540). London: Routledge.
- Wodak, R., & Busch, B. (2004). Approaches to media texts. In J. Downing, D. McQuail, P. Schlesinger, & E. Wartella (Eds.), *The SAGE handbook of media studies* (pp. 105–123). Thousand Oaks, CA: SAGE Publications.

- Wodak, R., Cillia, R. De, Reisigl, M., & Liebhart, K. (2009). *The discursive construction of national identity* (2nd ed.). Edinburgh: Edinburgh University Press.
- Wodak, R., & Meyer, M. (Eds.). (2009). *Methods of Critical Discourse Analysis* (2nd ed., p. 216). London: SAGE Publications.
- Wodak, R., & Reisigl, M. (1999). Discourse and Racism: European Perspectives. *Annual Review of Anthropology*, 28(1), 175–199.
- Worster, D. (2004). *Dust Bowl: The Southern Plains in the 1930s*. New York: Oxford University Press USA.
- Wynne, M. (1996). A Post-Editor's Guide to CLAWS7 Tagging. *UCREL*. Retrieved from <http://www.natcorp.ox.ac.uk/docs/claws7.html>
- Žagar, I. (2010). Topoi in Critical Discourse Analysis. *Lodz Papers in Pragmatics*, 6(1), 3–27.
- Zimmerman, D. H., & Wieder, D. L. (1970). Ethnomethodology and the problem of order: Comment on Denzin. In J. D. Douglas (Ed.), *Understanding everyday life: Toward the reconstruction of sociological knowledge* (pp. 285–298). Chicago: Aldine Publishing.

## Appendices

### Appendix A: Full USAS semantic tagset

Tag	Description
A1	GENERAL AND ABSTRACT TERMS
A1.1.1	General actions, making etc.
A1.1.2	Damaging and destroying
A1.2	Suitability
A1.3	Caution
A1.4	Chance, luck
A1.5	Use
A1.5.1	Using
A1.5.2	Usefulness
A1.6	Physical/mental
A1.7	Constraint
A1.8	Inclusion/Exclusion
A1.9	Avoiding
A2	Affect
A2.1	Affect:- Modify, change
A2.2	Affect:- Cause/Connected
A3	Being
A4	Classification
A4.1	Generally kinds, groups, examples
A4.2	Particular/general; detail
A5	Evaluation
A5.1	Evaluation:- Good/bad
A5.2	Evaluation:- True/false
A5.3	Evaluation:- Accuracy
A5.4	Evaluation:- Authenticity
A6	Comparing
A6.1	Comparing:- Similar/different
A6.2	Comparing:- Usual/unusual
A6.3	Comparing:- Variety
A7	Definite (+ modals)
A8	Seem
A9	Getting and giving; possession
A10	Open/closed; Hiding/Hidden; Finding; Showing
A11	Importance
A11.1	Importance: Important
A11.2	Importance: Noticeability
A12	Easy/difficult
A13	Degree
A13.1	Degree: Non-specific
A13.2	Degree: Maximizers
A13.3	Degree: Boosters
A13.4	Degree: Approximators
A13.5	Degree: Compromisers

Tag	Description
A13.6	Degree: Diminishers
A13.7	Degree: Minimizers
A14	Exclusivizers/particularizers
A15	Safety/Danger
B1	Anatomy and physiology
B2	Health and disease
B3	medicines and medical treatment
B4	Cleaning and personal care
B5	Clothes and personal belongings
C1	Arts and crafts
E1	EMOTIONAL ACTIONS, STATES AND PROCESSES General
E2	Liking
E3	Calm/Violent/Angry
E4	Happy/sad
E4.1	Happy/sad: Happy
E4.2	Happy/sad: Contentment
E5	Fear/bravery/shock
E6	Worry, concern, confident
F1	Food
F2	Drinks
F3	Cigarettes and drugs
F4	Farming & Horticulture
G1	Government, Politics and elections
G1.1	Government etc.
G1.2	Politics
G2	Crime, law and order
G2.1	Crime, law and order: Law and order
G2.2	General ethics
G3	Warfare, defence and the army; weapons
H1	Architecture and kinds of houses and buildings
H2	Parts of buildings
H3	Areas around or near houses
H4	Residence
H5	Furniture and household fittings
I1	Money generally
I1.1	Money: Affluence
I1.2	Money: Debts
I1.3	Money: Price
I2	Business
I2.1	Business: Generally
I2.2	Business: Selling
I3	Work and employment
I3.1	Work and employment: Generally
I3.2	Work and employmeny: Professionalism
I4	Industry
K1	Entertainment generally
K2	Music and related activities
K3	Recorded sound etc.
K4	Drama, the theatre and showbusiness
K5	Sports and games generally

Tag	Description
K5.1	Sports
K5.2	Games
K6	Childrens games and toys
L1	Life and living things
L2	Living creatures generally
L3	Plants
M1	Moving, coming and going
M2	Putting, taking, pulling, pushing, transporting &c.
M3	Vehicles and transport on land
M4	Shipping, swimming etc.
M5	Aircraft and flying
M6	Location and direction
M7	Places
M8	Remaining/stationary
N1	Numbers
N2	Mathematics
N3	Measurement
N3.1	Measurement: General
N3.2	Measurement: Size
N3.3	Measurement: Distance
N3.4	Measurement: Volume
N3.5	Measurement: Weight
N3.6	Measurement: Area
N3.7	Measurement: Length & height
N3.8	Measurement: Speed
N4	Linear order
N5	Quantities
N5.1	Entirety; maximum
N5.2	Exceeding; waste
N6	Frequency etc.
O1	Substances and materials generally
O1.1	Substances and materials generally: Solid
O1.2	Substances and materials generally: Liquid
O1.3	Substances and materials generally: Gas
O2	Objects generally
O3	Electricity and electrical equipment
O4	Physical attributes
O4.1	General appearance and physical properties
O4.2	Judgement of appearance (pretty etc.)
O4.3	Colour and colour patterns
O4.4	Shape
O4.5	Texture
O4.6	Temperature
P1	Education in general
Q1	LINGUISTIC ACTIONS, STATES AND PROCESSES; COMMUNICATION
Q1.1	LINGUISTIC ACTIONS, STATES AND PROCESSES; COMMUNICATION
Q1.2	Paper documents and writing
Q1.3	Telecommunications
Q2	Speech acts
Q2.1	Speech etc:- Communicative

Tag	Description
Q2.2	Speech acts
Q3	Language, speech and grammar
Q4	The Media
Q4.1	The Media:- Books
Q4.2	The Media:- Newspapers etc.
Q4.3	The Media:- TV, Radio and Cinema
S1	SOCIAL ACTIONS, STATES AND PROCESSES
S1.1	SOCIAL ACTIONS, STATES AND PROCESSES
S1.1.1	SOCIAL ACTIONS, STATES AND PROCESSES
S1.1.2	Reciprocity
S1.1.3	Participation
S1.1.4	Deserve etc.
S1.2	Personality traits
S1.2.1	Approachability and Friendliness
S1.2.2	Avarice
S1.2.3	Egoism
S1.2.4	Politeness
S1.2.5	Toughness; strong/weak
S1.2.6	Sensible
S2	People
S2.1	People:- Female
S2.2	People:- Male
S3	Relationship
S3.1	Relationship: General
S3.2	Relationship: Intimate/sexual
S4	Kin
S5	Groups and affiliation
S6	Obligation and necessity
S7	Power relationship
S7.1	Power, organizing
S7.2	Respect
S7.3	Competition
S7.4	Permission
S8	Helping/hindering
S9	Religion and the supernatural
T1	Time
T1.1	Time: General
T1.1.1	Time: General: Past
T1.1.2	Time: General: Present; simultaneous
T1.1.3	Time: General: Future
T1.2	Time: Momentary
T1.3	Time: Period
T2	Time: Beginning and ending
T3	Time: Old, new and young; age
T4	Time: Early/late
W1	The universe
W2	Light
W3	Geographical terms
W4	Weather
W5	Green issues

Tag	Description
X1	PSYCHOLOGICAL ACTIONS, STATES AND PROCESSES
X2	Mental actions and processes
X2.1	Thought, belief
X2.2	Knowledge
X2.3	Learn
X2.4	Investigate, examine, test, search
X2.5	Understand
X2.6	Expect
X3	Sensory
X3.1	Sensory:- Taste
X3.2	Sensory:- Sound
X3.3	Sensory:- Touch
X3.4	Sensory:- Sight
X3.5	Sensory:- Smell
X4	Mental object
X4.1	Mental object:- Conceptual object
X4.2	Mental object:- Means, method
X5	Attention
X5.1	Attention
X5.2	Interest/boredom/excited/energetic
X6	Deciding
X7	Wanting; planning; choosing
X8	Trying
X9	Ability
X9.1	Ability:- Ability, intelligence
X9.2	Ability:- Success and failure
Y1	Science and technology in general
Y2	Information technology and computing
Z0	Unmatched proper noun
Z1	Personal names
Z2	Geographical names
Z3	Other proper names
Z4	Discourse Bin
Z5	Grammatical bin
Z6	Negative
Z7	If
Z8	Pronouns etc.
Z9	Trash can
Z99	Unmatched

*Reproduced from:*

Rayson, P. (n.d.) *UCREL Semantic Analysis System (USAS)*. Retrieved 10 December 2013.  
From <http://ucrel.lancs.ac.uk/usas/semtags.txt>.



**Appendix B: Full collocation table for people in the Katrina corpus (Chapter 4)**

	Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
38	<i>missing</i>	ADJ	1,445	4.95	133	126	4.75	631.12	A3-A10-A9-	A
9	<i>ordinary</i>	ADJ	541	1.85	87	77	5.55	513.67	A6.2+	A
66	<i>desperate</i>	ADJ	1,403	4.81	85	79	4.14	332.69	Q2.2 A9-	A
29	<i>qualified</i>	ADJ	446	1.53	45	36	4.88	221.91	A1.2+/I3.2+ A1.7+	A
98	<i>familiar</i>	ADJ	1,489	5.10	56	48	3.46	168.32	A6.2+ S1.2.1+	A
39	<i>innocent</i>	ADJ	306	1.05	28	28	4.74	132.52	G2.1+ X9.1- G2.2+	A
118	<i>wonderful</i>	ADJ	1,322	4.53	37	36	3.03	91.30	A5.1+	A
45	<i>disabled</i>	ADJ	896	3.07	78	68	4.67	361.33	B2-A1.1.2 B2-/S2c	B
95	<i>sick</i>	ADJ	1,152	3.95	45	42	3.51	138.43	B2-O4.2- X5.2- E3-	B
101	<i>ill</i>	ADJ	829	2.84	30	27	3.40	88.03	B2-A1.1.2 B2-/S2c	B
94	<i>hungry</i>	ADJ	611	2.09	24	23	3.52	74.07	F1-/B1 S1.2.2+ A1.5.1/N5+	F
53	<i>homeless</i>	ADJ	2,522	8.64	200	161	4.53	889.25	H4-	H
50	<i>poor</i>	ADJ	6,562	22.48	533	435	4.57	2396.69	I1.1-A5.1- N5- E4.1- X9.1-	I
30	<i>low-income</i>	ADJ	1,016	3.48	102	95	4.87	501.99	I1.1-	I
52	<i>needy</i>	ADJ	642	2.20	51	47	4.54	226.88	I1.1-	I
81	<i>wealthy</i>	ADJ	870	2.98	43	40	3.85	151.39	I1.1+	I
107	<i>rich</i>	ADJ	1,746	5.98	54	49	3.17	142.94	I1.1+ O4.1 N5+ X5.2+ E4.1+	I
63	<i>unemployed</i>	ADJ	374	1.28	24	22	4.23	96.63	I3.1-	I
72	<i>middle-class</i>	ADJ	364	1.25	21	18	4.07	80.19	S5+	I
115	<i>working</i>	ADJ	812	2.78	23	22	3.05	57.24	S5+c	I
104	<i>dead</i>	ADJ	3,718	12.74	123	122	3.27	340.70	L1- X5.2- A4.2+ A1.1.2 A1.1.1- E3+ B2- A13.3	L
74	<i>displaced</i>	ADJ	3,945	13.52	225	211	4.06	854.46	M1	M
60	<i>many</i>	ADJ	38,283	131.17	2554	2211	4.28	10524.12	N5+	N

	Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
88	<i>most</i>	ADJ	15,909	54.51	693	646	3.67	2276.50	Z99	N
35	<i>1000</i>	ADJ	2,036	6.98	191	188	4.78	913.79	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
6	<i>100000</i>	ADJ	816	2.80	139	131	5.64	837.91	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
83	<i>100</i>	ADJ	5,014	17.18	231	224	3.75	782.62	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
64	<i>200</i>	ADJ	2,617	8.97	163	158	4.18	646.89	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
26	<i>10000</i>	ADJ	1,144	3.92	126	123	5.01	644.04	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
3	<i>180000</i>	ADJ	99	0.34	20	18	5.88	612.25	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
7	<i>200000</i>	ADJ	569	1.95	93	89	5.58	552.33	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
16	<i>20000</i>	ADJ	696	2.39	95	91	5.32	552.33	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
1	<i>1300</i>	ADJ	326	1.12	79	79	6.14	537.56	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N

	Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
49	2500	ADJ	402	1.38	33	33	4.58	527.91	I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2 N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
2	500000	ADJ	335	1.15	73	72	5.99	479.33	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
79	500	ADJ	2,536	8.69	128	126	3.88	455.85	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
71	fewer	ADJ	1,940	6.65	112	107	4.08	427.88	N5--	N
34	3000	ADJ	913	3.13	86	84	4.78	412.09	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
69	300	ADJ	1,788	6.13	105	104	4.10	404.62	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
112	50	ADJ	5,303	18.17	154	150	3.08	390.27	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
62	400	ADJ	1,345	4.61	88	88	4.26	357.68	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
22	30000	ADJ	549	1.88	66	64	5.13	349.26	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
43	700	ADJ	720	2.47	63	62	4.68	292.48	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1	N

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
18	1200	ADJ	390	1.34	52	5.28	286.39	T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
13	150000	ADJ	278	0.95	41	5.43	283.34	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
24	25000	ADJ	461	1.58	54	5.10	282.85	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
47	1500	ADJ	679	2.33	58	4.64	266.46	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
17	300000	ADJ	323	1.11	44	5.31	244.30	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
57	250	ADJ	790	2.71	56	4.37	236.45	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
42	6000	ADJ	558	1.91	50	4.71	234.52	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
25	7000	ADJ	407	1.40	45	5.01	230.34	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
4	1800	ADJ	196	0.67	36	5.74	222.85	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N

	Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
									T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	
73	600	ADJ	988	3.39	57	57	4.07	217.67	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
61	5000	ADJ	782	2.68	52	51	4.28	213.00	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
54	4000	ADJ	602	2.06	47	47	4.51	207.46	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
27	40000	ADJ	394	1.35	41	39	4.93	204.77	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
21	12000	ADJ	307	1.05	37	37	5.14	195.98	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
31	8000	ADJ	410	1.41	40	40	4.83	194.51	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
19	400000	ADJ	274	0.94	35	35	5.22	189.62	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
8	14000	ADJ	191	0.65	31	31	5.57	183.63	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
11	70000	ADJ	192	0.66	30	30	5.51	175.28	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
51	50000	ADJ	482	1.65	39	4.56	174.93	T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
15	80000	ADJ	203	0.70	27	5.38	163.92	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
70	150	ADJ	1,274	4.37	73	4.08	152.71	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
28	2000	ADJ	1,192	4.08	120	4.91	148.97	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
14	600000	ADJ	177	0.61	26	5.42	148.50	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
32	60000	ADJ	308	1.06	30	4.83	145.78	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
5	75000	ADJ	121	0.42	22	5.73	135.70	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
10	18000	ADJ	164	0.56	26	5.53	127.97	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
41	1100	ADJ	277	0.95	24	4.72	117.62	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1	N

	Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
78	120	ADJ	600	2.06	32	32	3.96	117.33	T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
20	1700	ADJ	174	0.60	21	21	5.14	111.29	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
37	250000	ADJ	228	0.78	21	21	4.75	99.66	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
33	9000	ADJ	210	0.72	20	20	4.80	96.27	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
116	65	ADJ	989	3.39	28	26	3.05	69.66	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
23	15000	ADJ	459	1.57	54	53	5.10	45.35	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
67	black	ADJ	8,692	29.78	515	346	4.11	1994.96	O4.3 G2.2- A5.1- E4.1- S5+	S
91	american	ADJ	14,784	50.65	587	477	3.54	1825.69	Z2/S2mf	S
105	white	ADJ	5,348	18.32	169	125	3.21	454.07	O4.3 G2.2- A5.1- E4.1- S5+	S
44	stranded	ADJ	712	2.44	62	57	4.67	287.24	S5-	S
76	encouraging	ADJ	714	2.45	40	40	4.03	150.47	S8+	S
93	iraqi	ADJ	839	2.88	33	30	3.52	101.93	Z2 Z2/S2mf	S
58	resilient	ADJ	326	1.12	22	21	4.30	90.75	S1.2.5+ O4.1	S

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
109 <i>generous</i>	ADJ	840	2.88	25	25	3.12	64.44	S1.2.2- N5+	S
119 <i>gay</i>	ADJ	788	2.70	22	20	3.03	54.19	S3.2/B1 O4.2+ E4.1+	S
36 <i>young</i>	ADJ	7,223	24.75	670	506	4.76	3192.53	T3-	T
65 <i>elderly</i>	ADJ	2,110	7.23	129	118	4.16	507.24	T3+++ T3+/S2c	T
82 <i>talented</i>	ADJ	472	1.62	22	22	3.77	74.99	X9.1+	X
108 <i>smart</i>	ADJ	918	3.15	28	25	3.15	73.41	O4.2+ X9.1+ S1.2.6+	X
56 <i>these</i>	ADJ	17,552	60.14	1250	1102	4.378	5300.125	Z5 Z8	Z
106 <i>where</i>	ADV	12,902	44.21	404	383	3.192	1079.367	M6	M
55 <i>worldwide</i>	ADV	468	1.60	36	34	4.489	157.834	W3 M7 A1.8+	W
59 <i>who</i>	PRON	98,238	336.59	6627	5078	4.299	27651.66	Z8	Z
120 <i>whose</i>	PRON	8,359	28.64	231	225	3.012	565.043	Z8	Z
111 <i>dozen</i>	SUBST	4,124	14.13	120	119	3.086	304.511	N1[i3.2.1 T3[i3.2.1 T1.2[i3.2.1 T1.3[i3.2.1 N3.3[i3.2.1 N3.5[i3.2.1 I1[i3.2.1 N1 T1.2 T3 T1.3 N3.2	N
87 <i>abita</i>	SUBST	666	2.28	30	30	3.717	100.315	Z99	Z
77 <i>affect</i>	VERB	5,655	19.38	302	248	3.962	1108.647	A2.2/E1	A
46 <i>trap</i>	VERB	1,053	3.61	91	85	4.657	420.218	A1.7+ M3@fn	A
40 <i>pluck</i>	VERB	309	1.06	28	28	4.725	131.961	A9+ A1.1.1 F1	A
110 <i>line</i>	VERB	1,455	4.99	43	42	3.109	110.294	O4.4 A4.1 Q1.3 X4.2 Q3 Q1.2 Q2.1/K4 S4 S7.1 A1.7+ A6.1- I3.1 M4@c G3@ K5.1@	A
117 <i>suffer</i>	VERB	4,769	16.34	135	134	3.047	335.932	E4.1- B2- X3 A5.1- /A2.1 S7.4+	E
113 <i>laugh</i>	VERB	1,011	3.46	29	28	3.066	72.83	E4.1+ /X3.2	E
85 <i>die</i>	VERB	7,087	24.28	323	283	3.734	1087.897	L1- T2-	L
48 <i>displace</i>	VERB	3,242	11.11	269	257	4.598	1220.542	A2.1+ M2	M
100 <i>flee</i>	VERB	2,205	7.56	80	77	3.405	235.141	M1	M



	Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
86	<i>educate</i>	VERB	621	2.13	28	28	3.718	93.679	P1	P
97	<i>advise</i>	VERB	1,152	3.95	44	42	3.479	133.509	Q2.2/S8+ Q2.1/X2.2+	Q
68	<i>scream</i>	VERB	525	1.80	31	30	4.107	119.771	Q2.2 E4.1+	Q
92	<i>interview</i>	VERB	939	3.22	37	36	3.524	114.416	Q2.1	Q
99	<i>persuade</i>	VERB	848	2.91	31	31	3.415	91.541	Q2.2	Q
103	<i>convince</i>	VERB	806	2.76	27	27	3.289	75.39	A7+/Q2.2	Q
75	<i>rescue</i>	VERB	3,414	11.70	192	172	4.037	723.873	S8+	S
12	<i>strand</i>	VERB	692	2.37	104	96	5.455	599.069	Z2 S5-	S
84	<i>encourage</i>	VERB	2,865	9.82	131	128	3.738	441.855	S8+	S
102	<i>gather</i>	VERB	2,385	8.17	80	78	3.291	223.597	S5+ A9+ N5+/A2.1 X2.5+ A1.1.1%	S
90	<i>enable</i>	VERB	749	2.57	31	30	3.595	98.713	S8+	S
114	<i>deserve</i>	VERB	1,164	3.99	33	31	3.049	82.185	S1.1.4+	S
80	<i>discourage</i>	VERB	440	1.51	22	22	3.867	77.937	S8-	S
96	<i>urge</i>	VERB	2,938	10.07	114	112	3.501	349.312	X7+	X
89	<i>remind</i>	VERB	1,511	5.18	63	62	3.605	201.493	X4.1	X

**Appendix C: Full collocation table for resident in the Katrina corpus (Chapter 5)**

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
7	<i>determining</i>	ADJ	117	0.45	51	6.83	407.54	A2.2 A7+ X6+	A
196	<i>affected</i>	ADJ	1,061	4.076	33	3.10	89.58	A2.1+ A5.4-/A8	A
128	<i>disabled</i>	ADJ	896	3.442	40	3.61	134.67	B2- A1.1.2 B2-/S2c	B
209	<i>desperate</i>	ADJ	1,403	5.39	44	3.03	108.63	E4.1- X5.2+	E
64	<i>low-income</i>	ADJ	1,016	3.903	77	4.30	318.44	I1.1-	I
24	<i>displaced</i>	ADJ	3,945	15.156	654	5.43	696.08	A2.1+ M2	M
162	<i>remaining</i>	ADJ	1,489	5.721	60	3.39	175.49	T2+++ M8 N5.2+	M
29	<i>returning</i>	ADJ	212	0.815	27	5.05	140.04	M1 A9- N6	M
91	<i>low-lying</i>	ADJ	618	2.374	36	3.92	130.32	O4.1	M
167	<i>uptown</i>	ADJ	1,168	4.487	45	3.33	127.90	M6	M
38	<i>evacuated</i>	ADJ	173	0.665	20	4.91	99.73	N5.1- M2 B1	M
40	<i>60000</i>	ADJ	308	1.183	35	4.89	173.33	N1[i]1.2.1 T3[i]1.2.1 T	N
8	<i>460000</i>	ADJ	53	0.204	22	6.76	172.98	N1[i]1.2.2 T3[i]1.2.2 T	N
205	<i>nearby</i>	ADJ	2,055	7.895	66	3.06	165.71	N3.3-	N
99	<i>100000</i>	ADJ	816	3.135	46	3.88	163.67	N1[i]2.2.1 N1 T1.2 T3	N
58	<i>500000</i>	ADJ	335	1.287	28	4.44	121.18	N1[i]3.2.1 T3[i]3.2.1 T	N
138	<i>200000</i>	ADJ	569	2.186	25	3.52	77.12	N1 T1.2 T3 T1.3 N3.2	N
186	<i>20000</i>	ADJ	696	2.674	24	3.17	63.29	N1 T1.2 T3 T1.3 N3.2	N
39	<i>stranded</i>	ADJ	712	2.735	82	4.91	408.40	S5-	S
35	<i>frail</i>	ADJ	188	0.722	22	4.93	110.25	S1.2.5- O4.1	S
2	<i>life-long</i>	ADJ	43	0.165	23	7.12	2108.27	T1.3+	T
76	<i>elderly</i>	ADJ	2,110	8.106	146	4.17	577.74	T3+++ T3+/S2c	T
44	<i>longtime</i>	ADJ	1,398	5.371	144	4.75	196.53	T1.3+	T
202	<i>coastal</i>	ADJ	3,153	12.114	102	3.07	257.50	W3/M4	W
21	<i>hurricane-</i>	ADJ	167	0.642	32	5.64	193.76	W4 Z3	W

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
37	<i>orleans-area</i>	ADJ	181	0.695	21	4.92	104.84	Z99	Z
168	<i>pontchartrain</i>	ADJ	706	2.712	27	3.32	76.32	Z99	Z
3	<i>lifelong</i>	ADJ	464	1.783	247	7.11	67.88	Z99	Z
41	<i>eligibility</i>	SUBST	489	1.879	54	4.85	264.25	A1.2+ S4	A
55	<i>nursing</i>	SUBST	1,186	4.557	104	4.51	460.36	B3 A9+ S8+	B
163	<i>belongings</i>	SUBST	724	2.782	29	3.38	84.47	B5	B
184	<i>complex</i>	SUBST	2,506	9.628	88	3.19	235.17	A12-	H
183	<i>rooftop</i>	SUBST	870	3.342	31	3.21	83.62	H2	H
107	<i>influx</i>	SUBST	627	2.409	34	3.82	118.44	M1	M
135	<i>thousand</i>	SUBST	8,716	33.486	387	3.53	1202.64	N1 I1	N
25	<i>referral</i>	SUBST	612	2.351	93	5.31	517.00	Q2.2/S2mf Q2.2	Q
96	<i>input</i>	SUBST	382	1.468	22	3.91	79.19	O2 Q2.2 Y2 A1.1.1 S5+	Q
179	<i>native</i>	SUBST	2,401	9.224	88	3.25	242.03	M7 A6.2+	S
152	<i>bernard</i>	SUBST	7,148	27.462	296	3.43	881.18	Z1mf	Z
112	<i>algiers</i>	SUBST	2,822	10.842	150	3.79	516.93	Z2	Z
166	<i>mandeville</i>	SUBST	2,646	10.166	103	3.34	294.67	Z2	Z
194	<i>gretna</i>	SUBST	2,518	9.674	84	3.12	216.71	Z2	Z
27	<i>waggaman</i>	SUBST	242	0.93	33	5.15	175.89	Z99	Z
82	<i>abita</i>	SUBST	666	2.559	42	4.04	158.51	Z99	Z
108	<i>westwego</i>	SUBST	775	2.978	42	3.82	146.26	Z99	Z
49	<i>folsom</i>	SUBST	358	1.375	32	4.54	142.81	Z99	Z
79	<i>hano</i>	SUBST	466	1.79	31	4.11	120.25	Z99	Z
85	<i>terrytown</i>	SUBST	491	1.886	30	3.99	111.40	Z99	Z
165	<i>lafitte</i>	SUBST	744	2.858	29	3.34	83.02	Z99	Z
187	<i>lacombe</i>	SUBST	872	3.35	30	3.16	79.05	Z99	Z
189	<i>laplace</i>	SUBST	766	2.943	26	3.14	67.88	Z99	Z

	Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
146	<i>lakeview</i>	SUBST	1,683	6.466	72	64	3.48	63.06	Z99	Z
122	<i>return</i>	VERB	17,118	65.765	846	674	3.69	2802.19	M1 A9- S1.1.2+ N6 I1.	M
120	<i>evacuate</i>	VERB	7,335	28.18	364	319	3.69	1207.43	N5.1- M2 B1	M
118	<i>flee</i>	VERB	2,205	8.471	110	105	3.70	365.78	M1	M
134	<i>scatter</i>	VERB	1,328	5.102	59	59	3.53	183.24	M1 N3.6+/A2.1	M
139	<i>stay</i>	VERB	661	2.54	29	29	3.51	89.41	Z99	M
198	<i>transport</i>	VERB	828	3.181	27	24	3.09	68.48	M3 M2 M4 M5	M
88	<i>displace</i>	VERB	3,242	12.455	191	186	3.94	17.23	M1 N3.6+/A2.1 T2-	M
182	<i>complain</i>	VERB	1,891	7.265	68	64	3.23	184.64	Q2.2/E2-	Q
136	<i>advise</i>	VERB	1,152	4.426	51	49	3.53	158.06	Q2.2/S8+ Q2.1/X2.2+	Q
151	<i>encourage</i>	VERB	2,865	11.007	119	113	3.43	354.80	S8+	S
100	<i>strand</i>	VERB	692	2.659	39	38	3.88	138.76	Z2 S5-	S
52	<i>urge</i>	VERB	2,938	11.288	260	227	4.53	1155.95	X7+	X
53	<i>heed</i>	VERB	306	1.176	27	27	4.52	119.80	X5.1+	X
201	<i>resolve</i>	VERB	1,081	4.153	35	35	3.08	88.34	X6+ X9.2+	X

**Appendix D: Full collocation table for victim in the Katrina corpus (Chapter 5)**

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
8	<i>alleged</i>	ADJ	531	1,065	25	4.55	110.96	Q2.2/A5.2	A
27	<i>desperate</i>	ADJ	1,403	2,815	23	3.03	56.51	E4.1- X5.2+	E
66	<i>needy</i>	ADJ	642	1,288	22	4.09	84.06	I1.1-	I
30	<i>displaced</i>	ADJ	3,945	7,915	116	3.95	442.75	A2.1+ M2	M
93	<i>stranded</i>	ADJ	712	1,429	29	4.34	120.57	S5-	S
36	<i>elderly</i>	ADJ	2,110	4,233	35	3.05	86.78	T3+++ T3+/S2c	T
13	<i>behalf</i>	PREP	660	1,324	24	4.18	94.48	S1.1.1	S
31	<i>donation</i>	SUBST	7,502	15,051	142	3.24	385.87	A9-	A
69	<i>plight</i>	SUBST	572	1,148	22	4.26	88.96	A12-	A
96	<i>sympathy</i>	SUBST	403	0.809	22	4.77	104.12	X2.5+/E1 E4.1- A6.1+	E
103	<i>violence</i>	SUBST	1,647	3,304	30	3.18	79.40	E3-	E
4	<i>abuse</i>	SUBST	1,311	2.63	22	3.06	54.98	E3- Q2.2 G2.2- A1.5.1/A5.1-	E
75	<i>rape</i>	SUBST	501	1,005	22	4.45	94.64	G2.1-/S3.2 A1.1.2	G
43	<i>fund-raiser</i>	SUBST	1,098	2,203	24	3.45	71.45	Z99	I
7	<i>aid</i>	SUBST	5,872	11,781	163	3.84	590.32	S8+	S
12	<i>assistance</i>	SUBST	7,529	15,105	146	3.33	430.81	S8+	S
24	<i>counseling</i>	SUBST	1,492	2,993	60	4.33	247.93	S8+ G2.1/S2mf S8+/Q2.2 M6 Q2.2 S7.1+ X4.2 K4 Q1.1 Q1.3/Y2 E4.2+ Q2.1 A1.8+ N5 E4.1+% S7.1- X4.1 C1/S2mf Q3 P1 S7.1-/S2mf A1.1.1 O1 A11.1+	S
15	<i>benefit</i>	SUBST	5,487	11,008	91	3.05	225.69	S8+ I1.1/G1.1	S
48	<i>helping</i>	SUBST	627	1,258	20	3.99	73.72	S8+	S
5	<i>advocate</i>	SUBST	1,174	2,355	21	3.16	54.86	Q2.2/S2mf G2.1/S2mf	S
55	<i>hurricane</i>	SUBST	84,157	168,842	4779	4.82	23285.87	W4	W

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	L.L. value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
29	<i>disaster</i>	SUBST	20,678	41,486	470	3.50	1436.14	A5.1---X9.2- A1.1.2/W4	W
91	<i>storm</i>	SUBST	38,469	77.18	862	3.48	736.58	W4 E3- A12-	W
101	<i>tsunami</i>	SUBST	1,356	2.721	115	5.40	646.30	W3/M4	W
40	<i>flood</i>	SUBST	10,325	20.715	176	3.09	445.21	W4 N5+	W
56	<i>hurricanes</i>	SUBST	2,689	5.395	82	3.93	295.32	W4	W
34	<i>earthquake</i>	SUBST	2,102	4.217	52	3.62	166.79	W3	W
63	<i>katrina</i>	SUBST	87,763	176.077	4427	4.65	20509.20	Z99	Z
99	<i>treat</i>	VERB	2,473	4.962	50	3.33	141.77	E4.2+ A9+	B
53	<i>house</i>	VERB	5,987	12.012	191	3.99	704.58	H1 P1/S5+c G1.1c	H
86	<i>shelter</i>	VERB	997	2	20	3.32	56.36	H1/A10- H4/H1c	H
47	<i>help</i>	VERB	34,763	69.744	1283	4.20	5109.03	S8+	S
6	<i>aid</i>	VERB	1,205	2.418	232	6.58	1706.14	S8+	S
14	<i>benefit</i>	VERB	3,340	6.701	229	5.10	1188.59	S8+ I1.1/G1.1	S
11	<i>assist</i>	VERB	3,172	6.364	176	4.79	838.94	S8+	S
78	<i>rescue</i>	VERB	3,414	6.849	56	3.03	137.72	S8+ S8+/S5+c	S
52	<i>honor</i>	VERB	1,011	2.028	23	3.50	70.20	A5.1+/A9- I1.3+/A9- A9- S7.1+/O2 G2.2+ S7.2+ L3 S7.4+ E2+ G1.1 Q2.2 M6 T2++ S7.1 M2 M8 N3 A3+ X2.1 A7 G1.2 S8+ I2.	S
90	<i>storm</i>	VERB	412	0.827	95	6.85	2615.09	W4 E3- A12-	W
57	<i>identify</i>	VERB	2,895	5.808	65	3.48	196.81	X2.2+ A10+ X2.5+	X

**Appendix E: Full collocation table for evacuee in the Katrina corpus (Chapter 5)**

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
6	150000	ADJ	278	0.502	30	5.90	189.68	N1[i1.2.1 T3[i1.2.1 T	N
69	300	ADJ	1,788	3.228	49	3.95	181.72	N1[i1.2.2 T3[i1.2.2 T	N
91	estimated	ADJ	1,928	3.48	43	3.66	143.11	X6/A7- N2/A7-	N
56	10000	ADJ	1,144	2.065	35	4.12	138.94	N1 T1.2 T3 T1.3 N3.2	N
17	25000	ADJ	461	0.832	27	5.02	137.05	N1[i4.4.1 I1[i4.4.1 N	N
61	2000	ADJ	1,192	2.152	36	4.07	136.13	N1[i4.4.2 I1[i4.4.2 N	N
98	1000	ADJ	2,036	3.675	43	3.55	133.63	N1[i4.4.3 I1[i4.4.3 N	N
44	5000	ADJ	782	1.412	28	4.31	115.04	I1[i8.4.4 N1[i8.4.4 N	N
31	200000	ADJ	569	1.027	25	4.61	112.64	I1[i12.4.4 N1[i12.4.4	N
46	20000	ADJ	696	1.256	24	4.26	96.85	I1[i16.4.4 N1[i16.4.4	N
60	1000000	ADJ	816	1.473	25	4.09	95.18	I1[i20.4.4 N1[i20.4.4	N
140	200	ADJ	2,617	4.724	38	3.01	92.31	I1[i24.4.4 N1[i24.4.4	N
41	4000	ADJ	602	1.087	22	4.34	91.24	I1[i28.4.4 N1[i28.4.4	N
51	1500	ADJ	679	1.226	22	4.17	86.12	T1.3[i35.3.1 T1.2[i35	N
42	30000	ADJ	549	0.991	20	4.34	82.82	T1.3[i35.3.2 T1.2[i35	N
102	400	ADJ	1,345	2.428	27	3.48	81.39	T1.2[i35.3.3 N1[i35.3	N
70	250	ADJ	790	1.426	22	3.95	79.79	N1 T1.2 T3 T1.3 N3.2	N
89	600	ADJ	988	1.783	23	3.69	75.61	N1[i40.4.2 I1[i40.4.2	N
82	3000	ADJ	913	1.648	22	3.74	73.78	N1[i40.4.3 I1[i40.4.3	N
119	fellow	ADJ	1,677	3.027	29	3.26	79.51	S5+ S2.2m P1/S2mf	S
93	elderly	ADJ	2,110	3.809	47	3.63	150.68	T3+++ T3+/S2c	T
133	refuge	SUBST	1,460	2.635	22	3.06	54.90	A15+/M7 A10-	A
59	shelter	SUBST	9,034	16.307	249	4.10	1067.87	H1/A10- H4/H1c	H
109	hotel	SUBST	9,730	17.563	135	3.38	530.11	H4/H1	H
10	influx	SUBST	627	1.132	54	5.58	316.29	M1	M

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
97	arrival	SUBST 983	1.774	21	19	3.57	65.70	M1 T2-	M
36	thousand	SUBST 8,716	15.733	344	313	4.45	1479.97	N1 I1	N
134	hurricane	SUBST 84,157	151.909	1249	1123	3.04	3101.34	W4	W
124	katrina	SUBST 87,763	158.418	1430	1205	3.17	3792.58	Z99	Z
90	houston	SUBST 7,692	13.885	179	157	3.69	588.98	Z2	Z
26	astrodome	SUBST 985	1.778	47	42	4.72	219.50	H2 O4.4 Z2	Z
80	absorb	VERB 801	1.446	20	20	3.79	68.34	A1.1.1 T1.3+ X2.3+ X5	A
15	house	VERB 5,987	10.807	379	331	5.13	1985.29	H1 P1/S5+c G1.1c	H
12	shelter	VERB 997	1.8	80	80	5.47	457.08	H1/A10- H4/H1c	H
50	occupy	VERB 971	1.753	32	29	4.19	126.30	H4 A1.5.1 G3 X5.1+ A1	H
67	arrive	VERB 5,909	10.666	169	143	3.99	621.80	M1 T2-	M
131	stay	VERB 10,526	19	163	152	3.10	414.90	M8 T2++ H4	M
103	settle	VERB 2,555	4.612	51	49	3.47	153.22	M8 X9.2+ I1.2 H4 E4.2	M
35	transport	VERB 828	1.495	33	31	4.47	142.46	M3 M2 M4 M5	M
104	flee	VERB 2,205	3.98	44	42	3.47	132.14	M1	M
117	relocate	VERB 2,528	4.563	45	45	3.30	125.72	M6/A2.1+	M
77	scatter	VERB 1,328	2.397	34	34	3.83	117.85	M1 N3.6+/A2.1	M
38	enrol	VERB 1,455	2.626	57	43	4.44	244.13	S5+	P
81	interview	VERB 939	1.695	23	21	3.76	77.84	Q2.1	Q
95	assist	VERB 3,172	5.726	69	67	3.59	218.24	S8+	S
83	aid	VERB 1,205	2.175	29	28	3.74	97.16	S8+	S
128	welcome	VERB 1,998	3.607	32	31	3.15	83.32	Q2.2 S1.1.1	S
66	strand	VERB 692	1.249	20	20	4.00	73.92	Z2 S5-	S
107	land	VERB 1,515	2.735	29	29	3.41	84.87	W3 M7	W



**Appendix F: Full collocation table for survivor in the Katrina corpus (Chapter 5)**

	Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
42	<i>displaced</i>	ADJ	3,945	2.163	33	33	3.93	118.40	M1	M
25	<i>800</i>	ADJ	4,275	2.344	82	82	5.13	425.45	N1 T1.2 T3 T1.3 N3.2	N
58	<i>among</i>	PREP	11,812	6.476	65	38	3.33	183.19	Z5	Z
14	<i>cancer</i>	SUBST	1,976	1.083	90	65	6.38	622.11	B2-	B
10	<i>breast</i>	SUBST	420	0.23	23	17	6.64	167.51	B1 F1 O2	B
2	<i>holocaust</i>	SUBST	178	0.098	41	34	8.72	423.82	E3-	E
31	<i>\$50</i>	SUBST	1,404	0.77	20	20	4.70	92.06	I1	I
13	<i>acorn</i>	SUBST	534	0.293	25	24	6.42	174.07	L3	L
47	<i>line</i>	SUBST	11,959	6.557	94	94	3.84	326.71	O4.4 A4.1 Q1.3 X4.2 Q3 Q1.2 Q2.1/K4 S4 S7.1 A1.7+ A6.1- I3.1 M4@c G3@ K5.1@	Q
4	<i>helpline</i>	SUBST	158	0.087	21	21	7.92	191.67	Q1.3/S8+	Q
45	<i>wife</i>	SUBST	6,602	3.62	53	40	3.87	186.21	S4f	S
38	<i>daughter</i>	SUBST	5,228	2.866	46	39	4.00	169.47	S4f	S
50	<i>son</i>	SUBST	6,238	3.42	44	38	3.69	143.95	S4m	S
44	<i>hurricane</i>	SUBST	84,157	46.142	695	615	3.91	2498.40	W4	W
43	<i>katrina</i>	SUBST	87,763	48.119	734	591	3.93	2657.36	Z99	Z
37	<i>include</i>	VERB	18,694	10.25	181	147	4.14	700.85	A1.8+	A
57	<i>house</i>	VERB	5,987	3.283	38	38	3.53	116.90	H1 P1/S5+c G1.1c	H
32	<i>rescue</i>	VERB	3,414	1.872	45	41	4.59	200.51	S8+ S8+/S5+c	S
21	<i>rally</i>	VERB	901	0.494	21	21	5.41	116.93	S1.1.1 K5.1	S
40	<i>organize</i>	VERB	2,359	1.293	20	20	3.95	72.25	S7.1+ O4.1 S5+@	S
17	<i>search</i>	VERB	1,645	0.902	45	44	5.64	264.98	X2.4	X

**Appendix G: Full collocation table for UNCONTESTED uses of refugee in the Katrina corpus (Chapter 6)**

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
19	<u>ADJ</u>	1,403	0.288	8	7	4.797	37.753	E4.1- X5.2+	E
3	<u>ADJ</u>	461	0.095	19	18	7.65	164.466	T1.3[i2.3.1 T1.2]i2.3.1 N1	N
12	<u>ADJ</u>	1,144	0.235	11	11	5.551	63.22	N1 T1.2 T3 T1.3 N3.2	N
4	<u>ADJ</u>	210	0.043	7	7	7.344	57.557	T1.3[i2.3.1 T1.2]i2.3.1 N1	N
2	<u>ADJ</u>	81	0.017	5	4	8.235	47.378	T1.3[i2.3.1 T1.2]i2.3.1 N1	N
10	<u>ADJ</u>	696	0.143	7	7	5.615	40.799	N1 T1.2 T3 T1.3 N3.2	N
22	<u>ADJ</u>	2,036	0.418	9	9	4.429	38.096	N1 T1.2 T3 T1.3 N3.2	N
20	<u>ADJ</u>	913	0.187	5	5	4.739	23.245	N1 T1.2 T3 T1.3 N3.2	N
26	<u>ADJ</u>	1,788	0.367	6	6	4.032	22.264	N1 T1.2 T3 T1.3 N3.2	N
35	<u>ADJ</u>	2,617	0.537	6	6	3.483	18	N1 T1.2 T3 T1.3 N3.2	N
8	<u>ADJ</u>	611	0.125	9	9	6.166	59.281	F1-/B1 S1.2.2+ A1.5.1/N5+	S
11	<u>ADJ</u>	712	0.146	7	7	5.582	40.491	S5-	S
31	<u>SUBST</u>	2,488	0.51	7	7	3.778	23.651	F1	F
18	<u>SUBST</u>	870	0.179	6	6	5.071	30.543	H2	H
16	<u>SUBST</u>	3,751	0.77	27	24	5.133	139.921	Z99	M
9	<u>SUBST</u>	985	0.202	14	14	6.114	91.239	H2 O4.4 Z2	M
7	<u>SUBST</u>	627	0.129	10	10	6.281	67.443	M1	M
38	<u>SUBST</u>	6,148	1.261	11	11	3.125	28.184	M3fn	M
14	<u>SUBST</u>	8,716	1.788	75	71	5.39	415.309	N1 I1	N
28	<u>SUBST</u>	5,569	1.142	18	18	3.978	65.631	N1	N
1	<u>SUBST</u>	89	0.018	6	6	8.357	57.901	M3/N5+	N
37	<u>SUBST</u>	4,124	0.846	8	8	3.241	21.599	N5	N
29	<u>SUBST</u>	2,968	0.609	9	9	3.886	31.698	S5+c N5++	S
36	<u>SUBST</u>	4,000	0.821	9	9	3.455	26.725	A12-	S
30	<u>SUBST</u>	7,692	1.578	22	22	3.801	75.193	Z2	W

Tagged lemma	POS	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
39 <i>katrina</i>	<u>SUBST</u>	87,763	18.003	155	145	3.106	396.06	Z99	Z
17 <i>house</i>	<u>VERB</u>	5,987	1.228	43	41	5.13	222.738	H1 P1/S5+c G1.1c	H
13 <i>shelter</i>	<u>VERB</u>	997	0.205	9	9	5.46	50.582	H1/A10- H4/H1c	H
24 <i>arrive</i>	<u>VERB</u>	5,909	1.212	21	18	4.115	80.272	M1 T2-	M
6 <i>camp</i>	<u>VERB</u>	228	0.047	5	5	6.739	36.855	M7 S5+c	M
21 <i>scatter</i>	<u>VERB</u>	1,328	0.272	6	6	4.461	25.611	M1 N3.6+/A2.1	M
34 <i>fill</i>	<u>VERB</u>	6,016	1.234	14	12	3.504	42.475	N5.1+ I3.1 Q1.2@	N
23 <i>pack</i>	<u>VERB</u>	1,524	0.313	6	6	4.263	24.043	O2 N5 S5+c	N
25 <i>enrol</i>	<u>VERB</u>	1,455	0.299	5	5	4.066	18.799	S5+	P
5 <i>huddle</i>	<u>VERB</u>	280	0.057	7	7	6.93	53.544	S5+c	S
27 <i>gather</i>	<u>VERB</u>	2,385	0.489	8	8	4.032	29.683	S5+ A9+ N5+/A2.1 X2.5+ A1.1.1%	S
33 <i>rescue</i>	<u>VERB</u>	3,414	0.7	9	6	3.684	29.32	S8+ S8+/S5+c	S
15 <i>strand</i>	<u>VERB</u>	692	0.142	5	5	5.138	25.915	Z2 S5-	S
40 <i>invite</i>	<u>VERB</u>	2,964	0.608	5	5	3.04	12.252	Q2.2	S

**Appendix H: Full collocation table for “people with (AIDS|A.I.D.S.|HIV|H.I.V.|GRID|G.R.I.D.)” in the AIDS/HIV corpus  
(Chapter 7)**

Collocate lemma	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
3 <i>quarantining</i>	266	0.044	12	11	8.101	111.339	Z99	A
13 <i>exclude</i>	3,304	0.543	15	15	4.788	70.493	A1.8-	A
11 <i>quarantine</i>	2,022	0.332	13	13	5.29	69.89	A1.8-/B2-	A
9 <i>stigmatize</i>	630	0.104	6	6	5.857	36.79	A5.1-G2.2-S9% A4.1 Q1.1 Q2.2/E2-04.2-Q1.2 A5/P1 N3.1 I1 Z1m S9/O1.2 Q2.2	A
34 <i>isolate</i>	3,323	0.546	5	5	3.195	13.038	A6.1-A1.8- S5- A10+	A
18 <i>infection</i>	48,601	7.984	151	134	4.241	602.705	B2- <del>i</del> 1.2.2 B2-	B
14 <i>afflict</i>	2,849	0.468	11	11	4.555	48.318	E6-	E
6 <i>homeless</i>	20,874	3.429	300	235	6.451	2097.017	H4-	H
8 <i>low-income</i>	4,916	0.808	56	54	6.116	364.942	I1.1-	I
37 <i>poor</i>	34,008	5.587	48	45	3.103	121.677	I1.1-A5.1-N5- E4.1- X9.1-	I
7 <i>indigent</i>	1,182	0.194	14	14	6.172	92.118	I1.1-A5.1-I1.1--- N5- E4.1- X9.1-	I
22 <i>needy</i>	3,158	0.519	8	8	3.947	28.688	I1.1-	I
2 <i>homebound</i>	725	0.119	123	123	10.012	1483.944	M1 M6	M
4 <i>citywide</i>	2,353	0.387	55	55	7.153	437.226	A1.8+ P1	M
30 <i>many</i>	155,846	25.602	275	257	3.425	809.593	N5+	N
1 <i>235470</i>	7	0.001	5	5	12.15	79.201	N1 T1.2 T3 T1.3 N3.2	N
16 <i>5000</i>	3,094	0.508	10	10	4.298	40.34	N1	N
21 <i>600</i>	3,896	0.64	10	10	3.966	36.253	N1 T1.2 T3 T1.3 N3.2	N
25 <i>10000</i>	4,030	0.662	9	9	3.765	30.076	N1 T1.2 T3 T1.3 N3.2	N
38 <i>1000</i>	7,980	1.311	11	11	3.069	27.39	N1 T1.2 T3 T1.3 N3.2	N

Collocate lemma		No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
20	<i>200000</i>	1,921	0.316	5	5	3.986	18.247	N1 T1.2 T3 T1.3 N3.2	N
23	<i>40000</i>	2,059	0.338	5	5	3.886	17.322	N1 T1.2 T3 T1.3 N3.2	N
35	<i>2000</i>	4,069	0.668	6	6	3.166	15.651	N1 T1.2 T3 T1.3 N3.2	N
31	<i>250</i>	2,851	0.468	5	5	3.416	14.584	N1 T1.2 T3 T1.3 N3.2	N
33	<i>3000</i>	3,048	0.501	5	5	3.32	13.722	N1 T1.2 T3 T1.3 N3.2	N
5	<i>coalition</i>	10,980	1.804	251	230	7.121	1987.09	S5+c	S
17	<i>help</i>	126,312	20.75	395	377	4.251	1585.338	B2-[j]1.2.1 S8+	S
12	<i>treat</i>	37,046	6.086	169	158	4.795	799.269	E4.2+ A9+	S
10	<i>assist</i>	7,524	1.236	63	63	5.672	372.346	S8+	S
27	<i>serve</i>	44,483	7.307	88	85	3.59	276.663	K5.1	S
29	<i>protect</i>	23,278	3.824	45	44	3.557	139.474	S8+/A15+ A10-	S
28	<i>deserve</i>	6,182	1.016	12	12	3.563	37.266	S1.1.4+	S
19	<i>enable</i>	3,821	0.628	10	10	3.994	36.592	S8+	S
32	<i>counsel</i>	5,203	0.855	9	9	3.396	25.922	G2.1/S2mf S8+/Q2.2	S
26	<i>for</i>	1,501,448	246.65	3128	2686	3.665	10462.9	Z5	Z
36	<i>or</i>	421,297	69.208	605	557	3.128	1563.403	Z5	Z
15	<i>against</i>	88,120	14.476	304	279	4.392	1275.861	Z5	Z
24	<i>toward</i>	21,509	3.533	50	50	3.823	172.084	Z5	Z

**Appendix I: Full collocation table for patient in the AIDS/HIV corpus (Chapter 8)**

	Collocate lemma	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
5	<i>quarantining</i>	266	0.224	37	37	7.369	309.62	Z99	A
30	<i>quarantine</i>	2,022	1.702	58	54	5.09	298.101	A1.8-/B2-	A
62	<i>exclusively</i>	1,698	1.43	29	26	4.342	119.817	A14	A
115	<i>isolate</i>	3,323	2.798	34	33	3.603	107.477	A6.1-A1.8-S5-A10+	A
70	<i>shun</i>	1,493	1.257	23	23	4.194	90.537	S1.2.1-A1.9	A
135	<i>confidentiality</i>	3,183	2.68	28	25	3.385	80.885	A10-	A
12	<i>isolating</i>	189	0.159	11	11	6.111	72.042	A6.1-A1.8-S5-A10+	A
140	<i>severely</i>	2,459	2.07	21	21	3.343	59.414	A5.1-S1.2.5+S1.2.1-	A
141	<i>commonly</i>	2,359	1.986	20	20	3.332	56.212	A6.2+	A
163	<i>solely</i>	1,689	1.422	12	10	3.077	29.876	A14	A
131	<i>classify</i>	1,111	0.935	10	8	3.418	29.238	A4.1	A
24	<i>treat</i>	37,046	31.189	1356	1172	5.442	7641.663	E4.2+ A9+	B
86	<i>azt</i>	12,963	10.913	175	153	4.003	645.228	B3	B
58	<i>hospice</i>	7,784	6.553	141	138	4.427	598.956	B3/H1c	B
142	<i>ill</i>	15,769	13.276	132	128	3.314	369.781	B2- A5.1-	B
43	<i>afflict</i>	2,849	2.399	66	63	4.782	311.806	E6-	B
123	<i>nursing</i>	9,438	7.946	91	76	3.518	278.155	B3 A9+ S8+	B
109	<i>pneumonia</i>	6,022	5.07	63	60	3.635	201.961	B2-	B
67	<i>hospitalize</i>	2,542	2.14	40	37	4.224	159.053	B3 A1.1.1 S6+ A1.7+/G	B
169	<i>transplant</i>	8,949	7.534	61	50	3.017	148.537	B3	B
15	<i>hospitalized</i>	428	0.36	22	22	5.932	138.631	Z99	B
32	<i>pneumocystis</i>	967	0.814	27	27	5.052	137.226	Z99	B
38	<i>blindness</i>	1,131	0.952	28	27	4.878	135.59	X3.4-	B
117	<i>pediatric</i>	4,183	3.522	42	42	3.576	131.306	B3	B
17	<i>cytomegalovirus</i>	422	0.355	19	18	5.741	114.503	B2- L2 Y2	B

Collocate lemma		No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
74	<i>kaposi</i>	1,975	1.663	29	26	4.124	111.271	Z99	B
87	<i>antiretroviral</i>	2,166	1.824	29	29	3.991	106.2	Z99	B
68	<i>lymphoma</i>	1,654	1.393	26	21	4.223	103.093	B2-04.2-G2.2-	B
122	<i>marrow</i>	3,285	2.766	32	31	3.532	98.316	B1 L3	B
20	<i>cmv</i>	403	0.339	16	14	5.559	92.523	N1 T1.2 T3 T1.3 N3.2	B
22	<i>retinitis</i>	387	0.326	15	13	5.525	86.058	B2-04.3	B
69	<i>opportunistic</i>	1,279	1.077	20	20	4.215	79.137	G2.2-	B
37	<i>leper</i>	513	0.432	13	13	4.912	63.674	B2-/S2mf	B
118	<i>anemia</i>	1,803	1.518	18	17	3.568	56.1	B2- X5.2-	B
170	<i>chemotherapy</i>	3,236	2.724	22	22	3.013	53.242	B3	B
46	<i>medi-cal</i>	501	0.422	11	10	4.705	50.589	B3	B
126	<i>sarcoma</i>	1,836	1.546	17	15	3.459	50.429	B2-04.2-G2.2-	B
92	<i>ulcer</i>	1,108	0.933	14	14	3.908	49.812	B2-	B
60	<i>unapproved</i>	557	0.469	10	10	4.415	42.238	Z99	B
107	<i>neurological</i>	1,106	0.931	12	11	3.688	39.085	B3 B1	B
150	<i>ddi</i>	1,765	1.486	14	14	3.236	37.634	Z99	B
90	<i>ribavirin</i>	772	0.65	10	9	3.944	36.061	I2.1/S2mf G1.1/X2.2+/ T1 T1.2 T1.3 T1.1.2 N	B
148	<i>hospitalization</i>	1,594	1.342	13	13	3.276	35.68	T1 T1.2 T1.3 T1.1.2 N	B
146	<i>dementia</i>	1,220	1.027	10	9	3.283	27.403	B2-/X1	B
161	<i>desperate</i>	3,716	3.129	27	27	3.109	68.715	E4.1- X5.2+	E
53	<i>meal</i>	8,283	6.973	160	150	4.52	699.398	F1	F
34	<i>appetite</i>	1,869	1.574	51	51	5.018	256.978	F1/B1 X7+	F
9	<i>brownny</i>	202	0.17	14	13	6.363	96.515	O4.3	F
96	<i>discrimination</i>	11,010	9.269	129	122	3.799	441.077	A6.1-G2.2-	G
116	<i>discriminate</i>	2,548	2.145	26	26	3.599	82.202	A6.1-/X2.2 S1.1.1	G
111	<i>homeless</i>	20,874	17.574	217	193	3.626	693.967	H4-	H

Collocate lemma		No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
147	<i>bed</i>	12,607	10.614	103	86	3.279	283.851	H5 H3/L3	H
52	<i>occupy</i>	2,488	2.095	49	48	4.548	215.932	H4 A1.5.1 G3 X5.1+ A1	H
113	<i>residence</i>	4,926	4.147	51	49	3.62	162.717	H1 H4	H
166	<i>residential</i>	3,411	2.872	24	20	3.063	59.716	H4	H
16	<i>indigent</i>	1,182	0.995	55	51	5.788	335.732	I1.1- A5.1- I1.1--- N	I
112	<i>low-income</i>	4,916	4.139	51	50	3.623	162.891	I1.1-	I
103	<i>impoverished</i>	1,867	1.572	21	20	3.74	70.011	I1.1- A5.1-	I
42	<i>destitute</i>	473	0.398	11	11	4.788	51.776	I1.1-	I
164	<i>reimbursement</i>	1,831	1.542	13	13	3.076	32.539	S8+	I
36	<i>dying</i>	5,578	4.696	143	140	4.928	703.999	L1- T2-	L
44	<i>terminally</i>	2,521	2.122	58	56	4.772	273.216	L1- T2-	L
120	<i>survival</i>	4,859	4.091	48	45	3.553	148.76	A3+/T2++ L1+	L
83	<i>baboon</i>	1,253	1.055	17	17	4.01	62.817	L2mfn	L
151	<i>terminal</i>	2,905	2.446	23	22	3.233	61.921	M5/H1 M4/H1 O3 Y2	L
33	<i>deathbed</i>	366	0.308	10	10	5.02	50.412	L1-	L
152	<i>surviving</i>	1,399	1.178	11	11	3.223	29.577	A3+/T2++ L1+	L
8	<i>homebound</i>	725	0.61	66	62	6.757	493.496	M1 M6	M
155	<i>thousand</i>	22,142	18.641	165	162	3.146	427.958	N1 I1	N
88	<i>4000</i>	2,245	1.89	30	28	3.989	109.827	I1 I11.4.4 N1 I11.4.4	N
171	<i>increasing</i>	5,932	4.994	40	38	3.002	96.518	N5+/A2.1	N
167	<i>10000</i>	4,030	3.393	28	28	3.045	68.862	I1 I19.4.4 N1 I19.4.4	N
85	<i>caseload</i>	1,184	0.997	16	16	4.005	58.953	N5	N
66	<i>disproportionate</i>	885	0.745	14	13	4.232	55.719	N5.2+	N
134	<i>rising</i>	1,921	1.617	17	15	3.394	49.169	M1 N5+/A2.1 N3.2+/A2.	N
106	<i>40%</i>	1,011	0.851	11	11	3.692	36.095	N1 I13.4.1 I1 I13.4.1 N	N
153	<i>6000</i>	1,724	1.451	13	13	3.163	33.75	I1 I11.4.4 N1 I11.4.4	N



	Collocate lemma	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
105	<i>care</i>	106,939	90.031	1169	969	3.699	3855.428	S8+ G1.1/S8+ B3 A1.3	S
56	<i>advocate</i>	17,701	14.902	332	264	4.478	1432.859	Q2.2/S2mf G2.1/S2mf	S
127	<i>advocacy</i>	6,407	5.394	59	58	3.451	175.239	Q2.2/G2.1 Q2.2	S
125	<i>counsel</i>	5,203	4.38	50	46	3.513	152.557	G2.1/S2mf S8+/Q2.2	S
121	<i>behalf</i>	4,290	3.612	42	42	3.54	129.534	S1.1.1	S
168	<i>comfort</i>	4,106	3.457	28	27	3.018	68.068	S8+	S
144	<i>visiting</i>	2,558	2.154	21	21	3.286	58.083	S1.1.1 M1	S
64	<i>counselse</i>	749	0.631	12	12	4.25	47.942	Z99	S
154	<i>susceptible</i>	1,475	1.242	11	11	3.147	28.343	S1.2.5-	S
110	<i>advanced</i>	4,123	3.471	43	42	3.631	137.608	T3- A5.1+ Y1 A12- T4-	T
61	<i>prolong</i>	1,800	1.515	32	32	4.4	134.653	T1.3++	T
11	<i>late-stage</i>	275	0.232	17	17	6.198	113.54	N4	T
41	<i>38-year-old</i>	504	0.424	12	12	4.822	57.043	Z99	T
143	<i>known</i>	3,164	2.664	26	26	3.287	71.677	X2.2+ S3.2/B1%	X
156	<i>experimental</i>	7,519	6.33	56	54	3.145	145.007	X2.4 Y1	Y
49	<i>getty</i>	1,538	1.295	32	32	4.627	144.259	A2.1+	Z
51	<i>clare</i>	849	0.715	17	15	4.572	75.444	Z1fZ2	Z
104	<i>bergalis</i>	990	0.834	11	10	3.722	36.487	Z99	Z
145	<i>kimberly</i>	1,464	1.233	12	11	3.283	32.884	Z99	Z

**Appendix J: Full collocation table for victim in the AIDS/HIV corpus (Chapter 8)**

Collocate lemma		No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
4	<i>quarantine</i>	2,022	0.511	62	60	6.923	473.827	A1.8-/B2-	A
1	<i>quarantining</i>	266	0.067	35	32	9.025	372.58	Z99	A
51	<i>acquire</i>	9,352	2.363	36	36	3.929	128.717	A9+	A
20	<i>plight</i>	2,141	0.541	15	15	4.793	70.615	A12-	A
54	<i>treat</i>	37,046	9.36	130	125	3.796	443.546	E4.2+ A9+	B
22	<i>carrier</i>	5,529	1.397	38	38	4.766	178.05	M3 M5/G3fn B2-/S2mf	B
23	<i>dying</i>	5,578	1.409	38	38	4.753	177.426	L1- T2-	B
65	<i>hospice</i>	7,784	1.967	22	22	3.484	66.124	B3/H1c	B
29	<i>handicapped</i>	1,730	0.437	11	11	4.653	49.72	B2- S8- B2-/S2c	B
49	<i>hemophilic</i>	3,301	0.834	13	13	3.962	46.955	B2-/S2mf B2- O4.2- X5.2- E3- S2mfc	B
74	<i>suffering</i>	4,447	1.124	12	12	3.417	34.952	E4.1- B2- X3 A5.1-/A2.1 S7.4+	B
10	<i>compassion</i>	5,033	1.272	60	58	5.56	345.619	E1 X2.5+/S1.2	E
24	<i>sympathy</i>	2,282	0.577	15	14	4.701	68.839	X2.5+/E1 E4.1- A6.1+	E
6	<i>discrimination</i>	11,010	2.782	193	171	6.116	1259.973	A6.1- G2.2-	G
8	<i>discriminate</i>	2,548	0.644	38	31	5.883	235.513	A6.1-/X2.2 S1.1.1	G
62	<i>protect</i>	23,278	5.882	69	66	3.552	213.555	S8+/A15+ A10-	G
46	<i>prohibit</i>	5,185	1.31	21	20	4.003	77.159	S7.4-	G
42	<i>homeless</i>	20,874	5.274	93	88	4.14	358.692	H4-	H
35	<i>quilt</i>	6,254	1.58	35	32	4.469	150.067	H5	H
41	<i>residence</i>	4,926	1.245	23	20	4.208	90.75	H1 H4	H
15	<i>indigent</i>	1,182	0.299	12	12	5.328	65.151	I1.1- A5.1- I1.1--- N5- E4.1- X9.1-	I
82	<i>employer</i>	9,079	2.294	23	23	3.326	64.513	I3.1/S2mf	I
95	<i>employment</i>	5,815	1.469	12	11	3.03	29.245	I3.1	I
72	<i>funeral</i>	12,551	3.171	34	33	3.422	99.634	L1-/S1.1.1	L

Collocate lemma		No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
38	<i>concentration</i>	2,204	0.557	11	11	4.304	44.533	X5.1+ N5+ I2.2% O1.2	N
86	<i>classroom</i>	6,219	1.571	15	15	3.255	40.745	P1/H2	P
60	<i>plea</i>	3,347	0.846	10	9	3.564	30.857	Q2.2	Q
50	<i>homosexual</i>	23,117	5.841	89	88	3.93	318.897	S3.2	S
79	<i>contact</i>	23,544	5.949	61	56	3.358	173.833	S1.1.1 S2mf Q2.2 O3	S
2	<i>memorialize</i>	423	0.107	19	18	7.474	159.843	H4 Q2.2 Q1.2 A1.1.1 Q2.1 S1.1.1 L1+/T2+++ N5.2+ K3 T1.1.1 A5.1+++ X4.1	S
21	<i>behalf</i>	4,290	1.084	30	30	4.791	141.463	S1.1.1	S
93	<i>advocate</i>	17,701	4.473	37	37	3.048	91.092	Q2.2/S2mf G2.1/S2mf	S
17	<i>commemorate</i>	1,418	0.358	13	13	5.181	68.057	S1.1.1	S
63	<i>deserve</i>	6,182	1.562	18	15	3.527	55.161	S1.1.4+	S
53	<i>dedicate</i>	4,554	1.151	16	14	3.798	54.532	Q2.2 A9-	S
88	<i>assist</i>	7,524	1.901	18	18	3.243	48.468	S8+	S
77	<i>tribute</i>	5,347	1.351	14	14	3.373	40.039	S7.2+	S
80	<i>embrace</i>	5,473	1.383	14	14	3.34	39.514	S3.2 S8+	S
57	<i>casual</i>	3,808	0.962	12	12	3.641	38.356	S1.2.1+ N6- O4.1	S
14	<i>14-year-old</i>	1,278	0.323	13	13	5.331	70.633	Z99	T
16	<i>13-year-old</i>	1,296	0.328	13	13	5.311	70.282	Z99	T
66	<i>teenage</i>	4,674	1.181	13	13	3.46	38.641	T3-	T
96	<i>identify</i>	17,002	4.296	35	33	3.026	85.315	X2.2+ A10+ X2.5+	X
32	<i>known</i>	3,164	0.799	19	18	4.571	83.794	X2.2+ S3.2/B1%	X
44	<i>innocent</i>	3,901	0.986	17	17	4.108	64.732	G2.1+ X9.1- G2.2+	X
31	<i>suspected</i>	1,785	0.451	11	11	4.608	49.084	X2.1/A5.1- A7	X
75	<i>against</i>	88,120	22.265	236	209	3.406	688.346	Z5	Z
11	<i>ryan</i>	7,339	1.854	83	82	5.484	469.565	Z1mf	Z

**Appendix K: Full collocation table for sufferer in the AIDS/HIV corpus (Chapter 8)**

Collocate lemma	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
2 <i>plight</i>	2,141	0.126	12	12	6.576	85.501	A12-	A
25 <i>treat</i>	37,046	2.177	29	25	3.736	96.479	E4.2+ A9+	B
4 <i>compassion</i>	5,033	0.296	18	18	5.928	112.321	E1 X2.5+ /S1.2	E
5 <i>discrimination</i>	11,010	0.647	39	38	5.914	243.304	A6.1- G2.2-	G
27 <i>protect</i>	23,278	1.368	15	15	3.455	44.413	S8+ /A15+ A10-	G
24 <i>ban</i>	16,267	0.956	13	13	3.766	43.557	S7.4-	G
14 <i>homeless</i>	20,874	1.226	28	27	4.513	121.681	H4-	H
36 <i>housing</i>	20,138	1.183	10	10	3.079	24.833	H4	H
33 <i>poor</i>	34,008	1.998	19	19	3.249	51.448	I1.1- A5.1- N5- E4.1- X9.1-	I
26 <i>majority</i>	13,728	0.807	10	10	3.632	31.96	N5+++c	N
20 <i>advocate</i>	17,701	1.04	18	17	4.113	68.683	Q2.2/S2mf G2.1/S2mf	S
23 <i>homosexual</i>	23,117	1.358	19	18	3.806	64.856	S3.2	S
15 <i>fellow</i>	8,981	0.528	12	12	4.507	51.999	S2.2m P1/S2mf	S
30 <i>against</i>	88,120	5.178	53	50	3.356	150.871	Z5	Z
34 <i>toward</i>	21,509	1.264	12	12	3.247	32.359	Z5	Z

**Appendix L: Full collocation table for carrier in the AIDS/HIV corpus (Chapter 8)**

Collocate lemma	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
3 quarantine	2,022	0.071	19	19	8.07	174.793	A1.8-/B2-	A
13 potential	16,686	0.583	19	19	5.025	95.516	A7+	A
30 infect	41,551	1.453	16	16	3.461	47.624	B2-	B
36 victim	32,982	1.153	11	11	3.254	29.907	A1.1.2/G2.1/S2mf B2-/-	B
11 discrimination	11,010	0.385	17	15	5.465	95.598	A6.1-G2.2-	G
19 protect	23,278	0.814	16	16	4.297	64.908	S8+/A15+ A10-	G
38 spread	34,391	1.202	11	10	3.194	28.861	F1 N3.6 N3.3 N5+/A2.1	N
24 partner	29,227	1.022	14	13	3.776	47.19	S3.1/S2mf S3.2/S2mf S	S
26 contact	23,544	0.823	11	11	3.74	36.65	S1.1.1 S2mf Q2.2 O3	S
2 suspected	1,785	0.062	19	15	8.25	179.529	X2.1/A5.1- A7	X
9 identify	17,002	0.594	28	28	5.558	161.072	X2.2+ A10+ X2.5+	X
4 knowingly	1,210	0.042	11	11	8.023	100.544	X2.2+	X
5 known	3,164	0.111	11	10	6.636	79.287	X2.2+ S3.2/B1%	X
28 against	88,120	3.081	34	32	3.464	101.303	Z5	Z

**Appendix M: Full collocation table for “(gay|gays|homosexual|homosexuals)” in the AIDS/HIV 1980s subcorpus (Chapter 9)**

	Collocate lemma	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
3	<i>openly</i>	663	3.415	219	176	6.003	1471.337	A10+	A
22	<i>activity</i>	4,383	22.576	309	258	3.775	1063.925	A1.1.1	A
17	<i>primarily</i>	1,386	7.139	154	152	4.431	668.638	A13.2	A
21	<i>movement</i>	2,377	12.244	168	110	3.778	579.119	M1 S5+c K2 A2.1+	A
29	<i>active</i>	2,904	14.958	162	147	3.437	485.56	X5.2+ A1.1.1	A
26	<i>mostly</i>	1,671	8.607	101	99	3.553	317.938	A13.2	A
8	<i>predominantly</i>	327	1.684	59	57	5.13	315.771	A13.2	A
13	<i>liberation</i>	454	2.339	62	49	4.729	295.38	A1.7-	A
14	<i>exclusively</i>	381	1.963	51	45	4.7	240.855	A14	A
23	<i>mainly</i>	924	4.759	63	60	3.727	212.773	A13.2	A
39	<i>high-risk</i>	1,639	8.442	68	66	3.01	166.842	A15- A1.4	A
36	<i>violence</i>	1,979	10.194	90	72	3.142	235.76	E3-	E
10	<i>intravenous</i>	4,379	22.556	620	588	4.781	3002.589	799	F
38	<i>user</i>	5,449	28.067	234	202	3.06	588.852	A1.5.1/S2mf	F
31	<i>abuser</i>	1,626	8.375	79	75	3.238	216.492	G2.2-/S2mf A5.1-/S2mf	F
19	<i>right</i>	17,925	92.329	1423	911	3.946	5237.17	S7.4+/H2[i]1.2.1 S7.4+ G2.2+ M6 G1.2	G
7	<i>activist</i>	2,238	11.528	410	318	5.152	2208.711	G1.2/S2mf	G
27	<i>discrimination</i>	3,723	19.177	224	182	3.546	703.349	A6.1- G2.2-	G
18	<i>discriminate</i>	739	3.807	62	59	4.026	234.382	A6.1-/X2.2 S1.1.1	G
4	<i>bathhouse</i>	773	3.982	197	142	5.629	1204.522	B4/M4 N5.1+ O2 Y2 A1.1.1 H1 A5.1+/A2.1 N3.2+/A2.1 P1/S5+c G1.1cH4	H

Collocate lemma		No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
20	<i>bar</i>	4,016	20.686	307	233	3.892	1105.171	F2/H1c F2/H2 F2/H5 O2 O4.4 S8-G2.1c K2	H
33	<i>population</i>	5,755	29.643	273	236	3.203	736.442	S2/N5c	N
12	<i>man</i>	26,981	138.976	3711	2160	4.739	17823.12	S2.2m	S
11	<i>community</i>	13,840	71.288	1948	1213	4.772	9426.827	S5+c	S
2	<i>lesbian</i>	2,433	12.532	1010	554	6.333	7361.786	S3.2/S2.1f	S
1	<i>bisexual</i>	1,396	7.191	797	598	6.792	6501.588	S3.2/B1	S
6	<i>male</i>	3,566	18.368	692	580	5.236	3815.282	S2.2	S
37	<i>leader</i>	9,069	46.713	390	272	3.062	982.731	S7.1+/S2mf A5.1+++ /S2mf Q4.2 G2.1/S2mf	S
24	<i>heterosexual</i>	4,622	23.807	282	253	3.566	893.016	S3.2	S
16	<i>straight</i>	1,520	7.829	181	153	4.531	811.408	O4.4 M6 N3.8+ A5.2+ N4	S
32	<i>couple</i>	4,624	23.818	221	149	3.214	599.142	N5+ S4c	S
25	<i>advocate</i>	2,804	14.443	170	140	3.557	536.195	Q2.2/S2mf G2.1/S2mf	S
15	<i>lifestyle</i>	843	4.342	112	94	4.689	527.244	S1.1.1	S
35	<i>relationship</i>	3,761	19.372	174	143	3.167	461.323	S3.1 S3.2 A2.2 S4	S
9	<i>promiscuous</i>	465	2.395	69	65	4.848	340.676	S3.2	S
5	<i>cater</i>	287	1.478	57	50	5.269	316.898	S8+ F1	S
28	<i>engage</i>	1,714	8.829	101	95	3.516	313.09	A1.1.1 S1.1.3+ X5.2+ I3.1 G3@	S
34	<i>encounter</i>	1,227	6.32	57	52	3.173	151.514	S3.1 A1.1.1	S
30	<i>among</i>	17,647	90.897	933	713	3.36	2707.109	Z5	Z

**Appendix N: Full collocation table for “(gay|gays|homosexual|homosexuals)” in the AIDS/HIV 1990s subcorpus (Chapter 9)**

	Collocate lemma	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
4	<i>openly</i>	2,305	8.791	1140	831	7.019	9515.581	A10+	A
25	<i>movement</i>	7,182	27.392	470	342	4.101	1815.345	M1 S5+c K2 A2.1+	A
19	<i>liberation</i>	986	3.761	134	117	5.155	715.299	A1.7-	A
6	<i>closeted</i>	195	0.744	87	75	6.87	701.815	A10- H5	A
15	<i>exclusively</i>	808	3.082	78	64	4.662	361.454	A14	A
14	<i>predominantly</i>	691	2.635	70	66	4.731	331.229	A13.2	A
40	<i>closet</i>	1,760	6.713	73	70	3.443	218.403	H5	A
53	<i>primarily</i>	2,429	9.264	76	75	3.036	188.298	A13.2	A
47	<i>exclude</i>	1,712	6.53	62	56	3.247	169.993	A1.8-	A
1	<i>bashing</i>	177	0.675	102	89	7.239	895.453	E3- Q2.2	E
22	<i>pride</i>	2,534	9.665	190	162	4.297	784.49	E4.2+/S1.2 S7.2+ S1.2.3+ O4.2	E
52	<i>condemn</i>	1,571	5.992	52	50	3.117	134.087	Q2.2/E2- Q2.2/A1.1.2 G2.1	E
29	<i>intravenous</i>	3,320	12.662	198	192	3.967	728.882	Z99	F
30	<i>right</i>	45,904	175.076	2736	1675	3.966	10089.14	S7.4+ G2.2+ M6 G1.2	G
27	<i>military</i>	11,328	43.204	730	490	4.079	2798.236	G3	G
26	<i>ban</i>	9,146	34.882	593	442	4.087	2279.91	S7.4-	G
23	<i>discrimination</i>	4,997	19.058	363	302	4.251	1476.294	A6.1- G2.2-	G
51	<i>agenda</i>	3,709	14.146	126	111	3.155	330.843	Q1.2/X7+	G
46	<i>ordinance</i>	1,534	5.851	56	48	3.259	322.788	G2.1	G
49	<i>lobby</i>	3,369	12.849	116	100	3.174	307.409	H2 G1.2/S5+c	G
34	<i>discriminate</i>	1,358	5.179	76	69	3.875	270.431	A6.1-/X2.2 S1.1.1	G
32	<i>activism</i>	965	3.681	56	53	3.927	203.169	G1.2/X7/A2.1	G
50	<i>armed</i>	1,901	7.25	65	53	3.164	171.426	G3 B1	G



	Collocata lemma	No. in whole corpus	Expected collocata frequency	Observed collocata frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
42	<i>elected</i>	1,379	5.259	52	44	3.306	146.422	G1.2/X7+	G
33	<i>bar</i>	9,876	37.667	559	441	3.891	2002.039	F2/H1c F2/H2 F2/H5 O2 O4.4 S8	H
9	<i>bathroom</i>	499	1.903	85	75	5.481	494.416	B4/M4 N5.1+ O2 Y2 A1.1.1 H1 A	H
48	<i>character</i>	13,325	50.821	481	277	3.243	1316.51	S2mf K4/S2mf S1.2 S2mf/A6.2-	K
36	<i>publication</i>	3,138	11.968	155	108	3.695	514.615	Q4	Q
37	<i>literature</i>	2,477	9.447	109	76	3.528	338.146	Q4.1 Q1.2	Q
18	<i>portrayal</i>	880	3.356	70	55	4.382	297.199	C1 Q2.2	Q
13	<i>depiction</i>	537	2.048	56	53	4.773	268.285	C1 Q2.2	Q
3	<i>lesbian</i>	10,965	41.82	5832	2535	7.124	49954.26	S3.2/B1	S
16	<i>man</i>	74,991	286.012	6933	3901	4.599	31670.41	S2.2m	S
31	<i>community</i>	42,615	162.532	2512	1701	3.95	9207.674	S5+c	S
5	<i>bisexual</i>	1,887	7.197	889	653	6.949	7302.45	S3.2/B1	S
17	<i>activist</i>	9,819	37.449	893	714	4.576	4032.952	G1.2/S2mf	S
20	<i>male</i>	8,626	32.899	679	557	4.367	2869.95	S2.2	S
12	<i>straight</i>	4,884	18.627	561	502	4.913	2799.599	O4.4 A5.2+ N4 A7+ A6.2+ S3.2	S
38	<i>couple</i>	12,533	47.8	541	412	3.501	1659.513	N5+ S4c	S
21	<i>heterosexual</i>	5,095	19.432	393	349	4.338	1644.939	S3.2	S
11	<i>lifestyle</i>	2,484	9.474	291	243	4.941	1463.755	S1.1.1	S
43	<i>advocate</i>	9,046	34.501	340	300	3.301	955.602	Q2.2	S
41	<i>marriage</i>	8,205	31.294	320	206	3.354	921.129	S4	S
7	<i>ordination</i>	255	0.973	52	35	5.741	909.58	S9	S
2	<i>african-</i>	154	0.587	87	87	7.211	758.688	Z2 Z2/S2mf	S
24	<i>tolerance</i>	1,150	4.386	79	74	4.171	312.542	S7.4+ O4.1	S
35	<i>acceptance</i>	1,773	6.762	88	80	3.702	292.958	A9+ X2.5+ S1.2.1+ S7.4+	S
19	<i>cater</i>	746	2.845	59	54	4.374	249.82	S8+ F1	S
39	<i>latino</i>	1,374	5.24	57	38	3.443	170.551	Z2/Q3 Z2 Z2/S2mf A5.4- Q3 Y2	S

	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
45	1,544	5.889	57	54	3.275	158.286	A2.2/X2.1 S8-	S
8	407	1.552	75	42	5.594	154.36	S9 X6+ G2.1	S
28	1,318	5.027	83	81	4.045	314.273	A6.2+/X2.1	X
44	2,817	10.744	105	93	3.289	293.44	S2 A6.1+ Q1.2	X

**Appendix O: Full collocation table for “(gay | gays | homosexual | homosexuals)” in the AIDS/HIV 2000s subcorpus (Chapter 9)**

	Collocate lemma	No. in whole corpus	Expected collocate frequency	Observed collocate frequency	No. of texts	MI score	LL value	USAS tag candidates (automatic)	USAS broad tag (adjusted)
2	<i>openly</i>	1,659	4,946	869	575	7.457	7820.232	A10+	A
21	<i>movement</i>	5,208	15,528	262	199	4.077	999.989	M1 S5+c K2 A2.1+	A
4	<i>closeted</i>	190	0.567	69	63	6.928	554.282	A10- H5	A
26	<i>oppose</i>	3,164	9,434	123	97	3.705	408.79	S8-/E2- A6.1-	A
11	<i>liberation</i>	825	2.46	76	71	4.949	381.192	A1.7-	A
29	<i>mainstream</i>	1,606	4.788	51	46	3.413	150.231	A6.2+	A
40	<i>equal</i>	1,990	5.933	51	47	3.104	130.338	A6.1+++	A
31	<i>abortion</i>	4,628	13.799	144	123	3.383	418.841	B3 T2- X9.2-	B
9	<i>pride</i>	2,258	6.732	222	168	5.043	1143.141	E4.2+/S1.2 S7.2+ S1.2.3+ O4.2	E
22	<i>right</i>	29,004	86.477	1457	854	4.075	5566.453	Z4 A5.3+ A4.2+ M6 A13.3 A13.2	G
18	<i>activist</i>	7,011	20.904	389	294	4.218	1558.981	G1.2/S2mf	G
14	<i>ban</i>	4,324	12.892	295	229	4.516	1302.06	S7.4-	G
20	<i>discrimination</i>	2,288	6.822	123	110	4.172	485.194	A6.1- G2.2-	G
35	<i>advocate</i>	5,839	17.409	167	141	3.262	460.006	Q2.2	G
13	<i>legalize</i>	717	2.138	62	54	4.858	303.009	G2.1+ G2.1+/A2.2 G2.2+/A2.2 G	G
33	<i>republicans</i>	3,508	10.459	103	62	3.3	288.613	G1.2/S2mf	G
39	<i>agenda</i>	3,256	9.708	85	69	3.13	220.062	Q1.2/X7+	G
42	<i>advocacy</i>	3,182	9.487	79	77	3.058	197.419	Q2.2/G2.1 Q2.2	G
23	<i>constitutional</i>	1,002	2.988	50	46	4.065	189.995	G1.1	G
36	<i>amendment</i>	2,171	6.473	61	53	3.236	166.026	A2.1+	G
24	<i>bar</i>	7,869	23.462	352	254	3.907	1264.156	F2/H1c F2/H2 F2/H5 O2 O4.4 S8	H
5	<i>bathroom</i>	204	0.608	53	46	6.445	383.624	B4/M4 N5.1+ O2 Y2 A1.1.1 H1 A	H

38	<i>character</i>	8,527	25,424	228	146	3,165	600,278	S2mf K4/S2mf S1.2 S2mf/A6.2-	K
34	<i>publication</i>	1,748	5,212	51	39	3,291	142,304	Q4	Q
1	<i>lesbian</i>	5,233	15,602	2909	1396	7,543	26710.28	S3.2/B1	S
10	<i>man</i>	48,909	145,824	4708	2306	5,013	24162.87	S2.2m	S
7	<i>marriage</i>	7,782	23,202	1249	617	5,75	7719.065	S4	S
3	<i>bisexual</i>	1,628	4,854	842	569	7,439	7547.028	S3.2/B1	S
27	<i>community</i>	35,989	107,303	1395	865	3,701	4637.957	S5+c	S
8	<i>straight</i>	3,096	9,231	407	336	5,462	2341.063	O4.4 M6 N3.8+ A5.2+ N4	S
17	<i>couple</i>	9,463	28,214	545	372	4,272	2224.352	N5+ S4c	S
41	<i>black</i>	25,065	74,732	636	368	3,089	1615.823	O4.3 G2.2- A5.1- E4.1- S5+	S
19	<i>male</i>	4,850	14,461	261	218	4,174	1030.282	S2.2	S
12	<i>acceptance</i>	1,031	3,074	94	74	4,934	469.509	A9+ X2.5+ S1.2.1+ S7.4+	S
15	<i>lifestyle</i>	1,609	4,797	99	85	4,367	416.652	S1.1.1	S
16	<i>heterosexual</i>	1,671	4,982	99	95	4,313	409.298	S3.2	S
25	<i>adoption</i>	2,338	6,971	99	58	3,828	345.044	A9+ S4	S
6	<i>african-</i>	333	0,993	55	55	5,792	342.899	Z2 Z2/S2mf	S
32	<i>bishop</i>	2,824	8,42	85	58	3,336	242.027	S9/S7.1+/S2mf	S
30	<i>parade</i>	2,667	7,952	83	73	3,384	241.42	K1/M1/S5+ G3/S5+ S5+ N5+ I2.2	S
43	<i>priest</i>	3,059	9,121	74	44	3,02	181.501	S9/S2mf	S
37	<i>prostitute</i>	1,997	5,954	56	54	3,233	152.215	S3.2+/I3.1/S2mf G2.2-/S2mf	S
28	<i>among</i>	24,310	72,481	872	580	3,589	2769.703	Z5	Z