

---

# A comparative study of the tagging of adverbs in modern English corpora

by Owen Nancarrow and Eric Atwell,  
Leeds University



## **A comparative study of the tagging of adverbs in modern English corpora**

Brown, LOB, BNC-Sampler and ICE-GB have related but different tagsets; adverbs are particularly different

Adverb is a “dustbin” category; subcategories are inconsistent between corpora, even within one corpus

We present a detailed analysis, grounded on “traditional” descriptions of adverbs in ELT textbooks

# Four corpora of modern English



UNIVERSITY OF LEEDS

A	B	C	D
1	The Brown Corpus	Brown	1964
2	The Lancaster-Oslo-Bergen Corpus	LOB	1978
3	The British National Corpus	BNC	1995
4	The International Corpus of English, British Component	ICE-GB	1998

Table 1  
Four corpora of modern English

# Four sets of related corpora



UNIVERSITY OF LEEDS

A	B	C	D	E	F	G
1	Brown	untagged	none	Brown1	1	1964
2	Tagged Brown	tagged	Brown	Brown2	1	1979
3	LOB	untagged	none	LOB1	1	1978
4	Tagged LOB	tagged	LOB	LOB2	1	1986
5	BNC	tagged	C5	BNC1	100	1995
6	BNC World	tagged	C5	BNC2	100	2000
7	BNC XML	tagged	C5	BNC3	100	2006
8	BNC Sampler	tagged	C7	BNCS	2	1999
9	BNC Baby	tagged	C5	BNCB	4	2004
10	ICE-GB	tagged and parsed	ICE	ICE-GB1	1	1998
11	ICE-GB 2 <sup>nd</sup> edition	tagged and parsed	ICE	ICE-GB2	1	2006

# Corpora in this paper



UNIVERSITY OF LEEDS

A	B	C	D	E
1	Brown2	Brown	B	1979
2	LOB2	LOB	L	1986
3	BNCS-W	the Sampler	S	1999
4	ICE-GB1	ICE	I	1998

## Section 63

There are seven kinds of adverbs:

1 of manner: e.g. *quickly, bravely, happily, hard, fast, well*

2 of place: e.g. *here, there, everywhere, up, down, near, by*

3 of time: e.g. *now, soon, yet, still, then, today*

4 of frequency: e.g. *twice, often, never, always, occasionally*

5 of degree: e.g. *very, fairly, rather, quite, too, hardly*

6 interrogative: e.g. *when? where? why?*

7 relative: e.g. *when, where, why*

(Thomson and Martinet, 1969: 38)



## Section 90

Some words can be used as either prepositions or adverbs ...

The most important words of this type are:

*in, on, up, down, off, near, through, along, across, under, round*

(Thomson and Martinet, 1969: 52)

# Manner adverbs



UNIVERSITY OF LEEDS

	Tag	Type total	Token total	Descriptive name
B	RB	1,764	36,602	adverb
L	RB	1,688	35,357	adverb
S	RR	1,473	28,207	general adverb
I	ADV(ge)	1,483	32,723	general adverb



# Type and token totals

	Corpus type total	Corpus token total	Adverb type total	Adverb token total
B	55,631	1,014,203	2,269	66,482
L	56,043	1,013,729	1,982	76,236
S	56,188	1,002,820	1,824	61,275
I	73,489	992,136	2,565	104,284
Ia			2,331	74,429

# Adverbs of time and place



UNIVERSITY OF LEEDS

<i>a Nominal adverbs</i>				
B	RN	3	9	nominal adverb
L	RN	15	4,333	nominal adverb
<i>b Adverbial nouns</i>				
B	NR	26	1,894	adverbial noun
L	NR	60	2,916	singular adverbial noun
<i>c Time adverbs</i>				
S	RT	33	4,635	quasi-nominal adverb of time
<i>d Place adverbs</i>				
S	RL	121	3,106	locative adverb

# 9 nominal adverbs in Brown



UNIVERSITY OF LEEDS

		token total
1	afar	2
2	here	4
3	then	3

# Inconsistent taggings in Brown



UNIVERSITY OF LEEDS

		token total
1a	outta here_RN	2
1b	out of here_RB	12
2a	around here_RN	1
2b	around here_RB	10
3a	from here_RN	1
3b	from here_RB	13

# Degree adverbs



UNIVERSITY OF LEEDS

*a Which may not follow the head*

B	QL	374	8,750	qualifier
L	QL	17	5,375	qualifier
S	RG	39	3,881	degree adverb
I	ADV(inten)	299	12,860	intensifier adverb

*b Which may follow the head*

B	QLP	4	263	post-qualifier
L	QLP	2	283	post-qualifier

# Inconsistent taggings in Brown



UNIVERSITY OF LEEDS

		tagged RB	tagged QL
1	exceedingly	1	5
2	sufficiently	2	23
3	terribly	3	5
4	unusually	5	5

# Interrogative and relative adverbs



UNIVERSITY OF LEEDS

<i>a Wh- general adverbs</i>				
B	WRB	20	4,569	wh- adverb
L	WRB	15	5,076	WH-adverb
S	RRQ	14	2,058	wh- general adverb
I	ADV(wh)	16	2,440	wh- adverb
<i>b Wh- degree adverbs</i>				
B	WQL	2	181	wh- qualifier
S	RGQ	1	229	wh- degree adverb
<i>c Wh-ever general adverbs</i>				
S	RRQV	5	70	wh-ever general adverb
<i>d Wh-ever degree adverbs</i>				
S	RGQV	2	28	wh-ever degree adverb
<i>e Relative adverbs</i>				
I	ADV(rel)	8	954	relative adverb

# Tagging schemes for B, L, S



UNIVERSITY OF LEEDS

Brown			
WRB	when	whenever	
	where	wherever	
	why		
	how	however	
WQL	how	however	
LOB			
WRB	when	whenever	
	where	wherever	
	why		
	how	however	
Sampler			
RRQ	when	whenever	RRQV
	where	wherever	
	why		
	how	however	
RGQ	how	however	RGQV



## ICE

ICE				
ADV(rel)	when	when	whenever	ADV(wh)
	where	where	wherever	
		why		
	how	how	however	

# Adverb or preposition



UNIVERSITY OF LEEDS

<i>a Phrasal adverbs</i>				
B	RP	12	6,039	adverb/particle
L	RP	28	8,700	adverbial particle
S	RP	21	6,873	prepositional adverb, particle
I	ADV(phras)	61	7,211	phrasal adverb
<i>b Adverbs with prepositional homographs</i>				
L	RI	21	571	adverb (homograph of preposition)
<i>c Catenative prepositional adverbs</i>				
S	RPK	1	40	prepositional adverb, catenative

# New adverb subclasses in S,I



UNIVERSITY OF LEEDS

<i>a Post-nominal adverbs</i>				
S	RA	50	502	adverb after nominal head
<i>b Nominal introducer adverbs</i>				
S	REX	11	454	adverb introducing nominal construction
<i>c Focusing adverbs</i>				
I	ADV(excl)	16	3,970	exclusive adverb
I	ADV(partic)	23	964	particularizer adverb
I	ADV(add)	22	3,349	additive adverb

# Negative *not*



UNIVERSITY OF LEEDS

B	*	3	4,615	not, n't
L	XNOT	8	7,454	not, n't
S	XX	5	6,140	not, n't

# Existential *there*



UNIVERSITY OF LEEDS

B	EX	3	2,169	existential there
L	EX	3	2,793	existential there
S	EX	3	2,215	existential there
I	EX THERE	1	3,444	existential there

I	CONNEC(ge)	198	18,944	general connective
I	CONNEC(appos)	42	1,666	appositive connective
I	REACT	126	10,911	reaction signal



# Adverbial inflections: manner adverbs

UNIVERSITY OF LEEDS

<i>a Comparative inflections</i>				
B	RBR	26	1,187	comparative adverb
L	RBR	25	1,357	comparative adverb
S	RRR	25	1,107	comparative general adverb
I	ADV(ge,comp)	24	826	comparative general adverb
<i>b Superlative inflections</i>				
B	RBT	12	101	superlative adverb
L	RBT	12	103	superlative adverb
S	RRT	16	109	superlative general adverb
I	ADV(ge,sup)	10	49	superlative general adverb

# Adverbial inflections: degree adverbs



UNIVERSITY OF LEEDS

<i>a Comparative inflections</i>				
S	RGR	2	1,032	comparative degree adverb
I	ADV(inten,comp)	14	1,291	comparative intensifier adverb
<i>b Superlative inflections</i>				
S	RGT	2	589	superlative degree adverb
I	ADV(inten,sup)	8	545	superlative intensifier adverb

# Comparative inflection of another adverbial subclass



UNIVERSITY OF LEEDS

I	ADV(add,comp)	1	3	comparative additive adverb
---	---------------	---	---	-----------------------------

# Plural inflections of adverbs



UNIVERSITY OF LEEDS

B	NRS	6	17	plural adverbial noun
L	NRS	8	84	plural adverbial noun

# Genitive inflections of adverbs



UNIVERSITY OF LEEDS

## *a Of manner adverbs*

B	RB\$	1	9	-
L	RB\$	1	6	adverb + genitive

## *b Of time and place adverbs*

B	NR\$	13	77	possessive adverbial noun
L	NR\$	13	48	singular adverbial noun + genitive

# Adverb ditto tag in LOB



UNIVERSITY OF LEEDS

L	RB <sup>11</sup>	61	1,785	adverb ditto tag
---	------------------	----	-------	------------------

# Contracted negative tags in Brown



UNIVERSITY OF LEEDS

	Tag	Type total	Token total	Example
1	BED*	1	22	were'nt
2	BEDZ*	1	155	wasn't
3	BEM*	1	9	aint't
4	BER*	2	48	aren't
5	BEZ*	3	117	isn't
6	DO*	2	488	don't
7	DOD*	2	405	didn't
8	DOZ*	2	90	doesn't
9	HV*	2	42	haven't
10	HVD*	1	100	hadn't
11	HVZ*	2	22	hasn't
12	MD*	11	867	couldn't

# Combined adverb tags in Brown



UNIVERSITY OF LEEDS

	Tag	Type total	Token total	Example	Relevant table
a adverb					
1	RB+BEZ	2	13	here's	4
2	RB+CS	2	3	soon's	
b adverbial noun					
1	NR+MD	1	1	today'll	6
c wh-adverb					
1	WRB+BER	1	1	where're	11
2	WRB+BEZ	2	14	where's	
3	WRB+DO	1	1	howda	
4	WRB+DOD	2	6	how'd	
5	WRB+DOD*	1	1	whyn't	
6	WRB+DOZ	1	1	how's	
8	WRB+IN	1	1	why'n	
7	WRB+MD	1	1	where'd	
d adverb/particle					
1	RP+IN	2	4	outta	14
e existential <i>there</i>					
1	EX+BEZ	1	105	there's	17
2	EX+HVD	1	3	there'd	
3	EX+HVZ	1	2	there's	
4	EX+MD	2	4	there'd	
f comparative adverb					
1	RBR+CS	1	1	more'n	19



# Tags with negative feature in ICE

	Class of tag	Type total	Token total
1	PRON	33	1,754
2	V	23	841
3	AUX	83	5,349

# Discontinuous adverb tags in ICE



UNIVERSITY OF LEEDS

1	ADV(ge,disc1)	2	2	discontinuous general adverb
2	ADV(inten,disc1)	4	4	discontinuous intensifier adverb
3	CONNEC(ge,disc1)	3	3	discontinuous general connective
4	ADV(inten,comp,disc1)	1	2	discontinuous comparative intensifier adverb

# Ignored adverb tags in ICE



UNIVERSITY OF LEEDS

1	ADV(ge,ignore)	100	415
2	ADV(inten,ignore)	47	221
3	ADV(wh,ignore)	4	134
4	ADV(rel,ignore)	2	15
5	ADV(phras,ignore)	14	49
6	ADV(excl,ignore)	4	123
7	ADV(partic,ignore)	4	9
8	ADV(add,ignore)	5	19
9	EXTHERE(ignore)	1	168
10	CONNEC(ge,ignore)	33	407
11	CONNEC(appos,ignore)	9	18
12	REACT(ignore)	7	58
13	ADV(ge,comp,ignore)	2	3
14	ADV(inten,comp,ignore)	2	23
15	ADV(inten,sup,ignore)	1	4

# Anomalous adverb tags in ICE



UNIVERSITY OF LEEDS

1	ADV	23	50	(anomalous) adverb
2	CONNEC	24	389	(anomalous) connective
3	CONNEC(ignore)	3	7	ignored (anomalous) connective
4	ADV(comp)	7	10	anomalous comparative
5	ADV(sup)	1	1	anomalous superlative

# Synoptic table

1	2	3	4	Manner (3)
5	6			Time & place (4)
7	8			
		9		
		10		
11	12	13	14	Degree (5)
15	16			
17	18	19	20	Interrogative and relative (6)
21		22		
		23		
		24		
			25	
26	27	28	29	Adverb or preposition (7)
	30			
		31		
		32		New subclasses (8)
		33	43	
			34	
			35	
			36	
37	38	39		New classes (9)
40	41	42	43	
			44	
		45	45	
			46	
47	48	49	50	Inflections (10)
51	52	53	54	Comp. and imp. manner
		55	56	Comp. and imp. degree
		57	58	
			59	Comp. additive
60	61			Plural adbl noun
62	63			Genitive manner
64	65			Genitive adbl noun
	66			Ditto manner (11)
	67			Dtic. manner (14)
	68			Ignored manner (13)
	69			Anomalous (16)

RB	RB	RR	ADV(ge)	4
RN	RN			6
NR	NL			
		RT		
		RL		
QL	QL	RG	ADV(inten)	9
QLP	QLP			
WRB	WRB	RRQ	ADV(wh)	11
WQL		RRQ		
		RRQV		
		RRQV		
			ADV(rel)	
RP	RP	RP	ADV(phrase)	14
	RI			
		RPK		
		RA		15
		REX	CONNEC(assort)	
			ADV(excl)	
			ADV(partic)	
			ADV(add)	
*	XNOT	XX		16
EX	EX	EX	EXTHERE	17
			CONNEC(ge)	18
		REX	CONNEC(assort)	
			REACT	
RBR	RBR	RRR	ADV(ge,comp)	19
RBT	RBT	RRT	ADV(ge,imp)	
		ROR	ADV(inten,comp)	20
		RGT	ADV(inten,imp)	
			ADV(add,comp)	21
NRS	NRS			22
RES	RES			23
NRS	NRS			
	RF'			24
			ADV(ge,excl)	28
			ADV(ge,ignore)	29
			ADV	30



LEEDS

Other studies have included comparisons between English corpus tagsets (eg van Halteren 1999, Atwell et al 2000, Jurafsky and Martin 2000), but none to our knowledge has focused on adverbs, or examined differences of sub-categorizations in such detail.

Tagset “standards” should include this level of detail.

The approach in this paper provides a methodology to follow in examining sub-categorizations in other corpus tagsets, and/or other grammatical categories.