

# Appendices

## Appendix 1: CLAWS 7 Tagset

APPGE	Possessive pronoun, pre-nominal (e.g., <i>my, your, our</i> )
AT	Article (e.g., <i>the, no</i> )
AT1	Singular article (e.g., <i>a, an, every</i> )
BCL	Before-clause marker (e.g., <i>in order (that), in order (to)</i> )
CC	Coordinating conjunction (e.g., <i>and, or</i> )
CCB	Adversative coordinating conjunction ( <i>but</i> )
CS	Subordinating conjunction (e.g., <i>if, because, unless, so, for</i> )
CSA	As (as conjunction)
CSN	Than (as conjunction)
CST	That (as conjunction)
CSW	Whether (as conjunction)
DA	After-determiner or post-determiner capable of pronominal function (e.g., <i>such, former, same</i> )
DA1	Singular after-determiner (e.g., <i>little, much</i> )
DA2	Plural after-determiner (e.g., <i>few, several, many</i> )
DAR	Comparative after-determiner (e.g., <i>more, less, fewer</i> )
DAT	Superlative after-determiner (e.g., <i>most, least, fewest</i> )
DB	Before determiner or pre-determiner capable of pronominal function ( <i>all, half</i> )
DB2	Plural before-determiner ( <i>both</i> )
DD	Determiner (capable of pronominal function) (e.g., <i>any, some</i> )
DD1	Singular determiner (e.g., <i>this, that, another</i> )
DD2	Plural determiner ( <i>these, those</i> )
DDQ	Wh-determiner ( <i>which, what</i> )
DDQGE	Wh-determiner, genitive ( <i>whose</i> )
DDQV	Wh-ever determiner ( <i>whichever, whatever</i> )

EX	Existential <i>there</i>
FO	Formula
FU	Unclassified word
FW	Foreign word
GE	Germanic genitive marker—(' or 's)
IF	<i>For</i> (as preposition)
II	General preposition
IO	<i>Of</i> (as preposition)
IW	<i>With, without</i> (as prepositions)
JJ	General adjective
JJR	General comparative adjective (e.g., <i>older, better, stronger</i> )
JJT	General superlative adjective (e.g., <i>oldest, best, strongest</i> )
JK	Catenative adjective ( <i>able</i> in <i>be able to</i> , <i>willing</i> in <i>be willing to</i> )
MC	Cardinal number, neutral for number ( <i>two, three...</i> )
MC1	Singular cardinal number ( <i>one</i> )
MC2	Plural cardinal number (e.g., <i>sixes, sevens</i> )
MCGE	Genitive cardinal number, neutral for number ( <i>two's, 100s</i> )
MCMC	Hyphenated number ( <i>40–50, 1770–1827</i> )
MD	Ordinal number (e.g., <i>first, second, next, last</i> )
MF	Fraction, neutral for number (e.g., <i>quarters, two-thirds</i> )
ND1	Singular noun of direction (e.g., <i>north, southeast</i> )
NN	Common noun, neutral for number (e.g., <i>sheep, cod, headquarters</i> )
NN1	Singular common noun (e.g., <i>book, girl</i> )
NN2	Plural common noun (e.g., <i>books, girls</i> )
NNA	Following noun of title (e.g., <i>M.A.</i> )
NNB	Preceding noun of title (e.g., <i>Mr., Prof.</i> )
NNL1	Singular locative noun (e.g., <i>Island, Street</i> )
NNL2	Plural locative noun (e.g., <i>Islands, Streets</i> )
NNO	Numeral noun, neutral for number (e.g., <i>dozen, hundred</i> )
NNO2	Numeral noun, plural (e.g., <i>hundreds, thousands</i> )
NNT1	Temporal noun, singular (e.g., <i>day, week, year</i> )
NNT2	Temporal noun, plural (e.g., <i>days, weeks, years</i> )
NUU	Unit of measurement, neutral for number (e.g., <i>in, cc</i> )
NUU1	Singular unit of measurement (e.g., <i>inch, centimeter</i> )
NUU2	Plural unit of measurement (e.g., <i>inches, centimeters</i> )
NP	Proper noun, neutral for number (e.g., <i>IBM, Andes</i> )
NP1	Singular proper noun (e.g., <i>London, Jane, Frederick</i> )
NP2	Plural proper noun (e.g., <i>Browns, Reagans, Koreas</i> )
NPD1	Singular weekday noun (e.g., <i>Sunday</i> )
NPD2	Plural weekday noun (e.g., <i>Sundays</i> )
NPM1	Singular month noun (e.g., <i>October</i> )
NPM2	Plural month noun (e.g., <i>Octobers</i> )
PN	Indefinite pronoun, neutral for number ( <i>none</i> )
PN1	Indefinite pronoun, singular (e.g., <i>anyone, everything, nobody, one</i> )
PNQO	Objective wh-pronoun ( <i>whom</i> )

PNQS	Subjective wh-pronoun ( <i>who</i> )
PNQV	Wh-ever pronoun ( <i>whoever, whomever, whomsoever, whosoever</i> )
PNX1	Reflexive indefinite pronoun ( <i>oneself</i> )
PPGE	Nominal possessive personal pronoun (e.g., <i>mine, yours</i> )
PPH1	Third person singular neuter personal pronoun ( <i>it</i> )
PPHO1	Third person singular objective personal pronoun ( <i>him, her</i> )
PPHO2	Third person plural objective personal pronoun ( <i>them</i> )
PPHS1	Third person singular subjective personal pronoun ( <i>he, she</i> )
PPHS2	Third person plural subjective personal pronoun ( <i>they</i> )
PPIO1	First person singular objective personal pronoun ( <i>me</i> )
PPIO2	First person plural objective personal pronoun ( <i>us</i> )
PPIS1	First person singular subjective personal pronoun ( <i>I</i> )
PPIS2	First person plural subjective personal pronoun ( <i>we</i> )
PPX1	Singular reflexive personal pronoun (e.g., <i>yourself, itself</i> )
PPX2	Plural reflexive personal pronoun (e.g., <i>yourselves, themselves</i> )
PPY	Second person personal pronoun ( <i>you</i> )
RA	Adverb, after nominal head (e.g., <i>else, galore</i> )
REX	Adverb introducing appositional constructions (e.g., <i>namely, viz., e.g.</i> )
RG	Degree adverb (very, so, too)
RGQ	Wh-degree adverb ( <i>how</i> )
RGQV	Wh-ever degree adverb ( <i>however</i> )
RGR	Comparative degree adverb ( <i>more, less</i> )
RGT	Superlative degree adverb ( <i>most, least</i> )
RL	Locative adverb (e.g., <i>alongside, forward</i> )
RP	Prep. adverb, particle (e.g., <i>about, in</i> )
RPK	Prep. adv., catenative ( <i>about in be about to</i> )
RR	General adverb
RRQ	Wh-general adverb ( <i>where, when, why, how</i> )
RRQV	Wh-ever general adverb ( <i>wherever, whenever</i> )
RRR	Comparative general adverb (e.g., <i>better, longer</i> )
RRT	Superlative general adverb (e.g., <i>best, longest</i> )
RT	Quasi-nominal adverb of time (e.g., <i>now, tomorrow</i> )
TO	Infinitive marker ( <i>to</i> )
UH	Interjection (e.g., <i>oh, yes, um</i> )
VB0	<i>Be</i> , base form (finite, i.e., imperative, subjunctive)
VBDR	<i>Were</i>
VBDZ	<i>Was</i>
VBG	<i>Being</i>
VBI	<i>Be</i> , infinitive (To be or not... It will be ...)
VBM	<i>Am</i>
VBN	<i>Been</i>
VBR	<i>Are</i>
VBZ	<i>Is</i>
VD0	<i>Do</i> , base form (finite)
VDD	<i>Did</i>

VDG	<i>Doing</i>
VDI	<i>Do</i> , infinitive (I may do... To do...)
VDN	<i>Done</i>
VDZ	<i>Does</i>
VH0	<i>Have</i> , base form (finite)
VHD	<i>Had</i> (past tense)
VHG	<i>Having</i>
VHI	<i>Have</i> , infinitive
VHN	<i>Had</i> (past participle)
VHZ	<i>Has</i>
VM	Modal auxiliary ( <i>can, will, would</i> , etc.)
VMK	Modal catenative ( <i>ought, used</i> )
VV0	Base form of lexical verb (e.g., <i>give, work</i> )
VVD	Past tense of lexical verb (e.g., <i>gave, worked</i> )
VVG	-ing participle of lexical verb (e.g., <i>giving, working</i> )
VVGK	-ing participle catenative ( <i>going</i> in <i>be going to</i> )
VVI	Infinitive (e.g., <i>to give... It will work...</i> )
VVN	Past participle of lexical verb (e.g., <i>given, worked</i> )
VVNK	Past participle catenative (e.g., <i>bound</i> in <i>be bound to</i> )
VVZ	-s form of lexical verb (e.g., <i>gives, works</i> )
XX	Not, n't
ZZ1	Singular letter of the alphabet (e.g., <i>a, b</i> )
ZZ2	Plural letter of the alphabet (e.g., <i>a's, b's</i> )

## Appendix 2: POS List

[nn*]	noun.all	[v?d*]	verb.ED
[v*]	verb.all	[v?n*]	verb.EN
[j*]	adj.all	[v?g*]	verb.ING
[r*]	adv.all	[vv*]	verb.[LEX]
[xx*]	neg.all	[vb*]	verb.[BE]
[at*]	art.all	[vd*]	verb.[DO]
[d*]	det.all	[vh*]	verb.[HAVE]
[p*]	pron.all	[j jr*]	adj.CMP
[app*]	poss.all	[j jt*]	adj.SPRL
[i*]	prep.all	[rp*]	adv.PRTCL
[c*]	conj.all	[rrq*]	adv.WH
[n*]	noun.all+	[pn1*]	pron.INDF
[*nn1*]	noun.SG	[pp*]	pron.PERS
[*nn2*]	noun.PL	[pnq*]	pron.WH
[*nn0*]	noun.CMN	[ppx*]	pron.REFL
[np*]	noun.+PROP	[mc*]	num.CARD
[nn*]	noun.-PROP	[md*]	num.ORD
[vv0*]	verb.BASE	[cc*]	conj.CRD
[v?i*]	verb.INF	[cs*]	conj.SUB
[vvi*]	verb.INF/LEX	[uh*]	interj
[vm*]	verb.MODAL	[y*]	PUNC
[v?z*]	verb.3SG		

## Appendix 3: Query Syntax

Simple slot			
Syntax	Meaning	Examples	Sample matches
Word	One exact word	Mysterious	Mysterious
[pos]	Part of speech (exact)	[v?g*]	Going, using
[pos*]	Part of speech (wildcard)	[v*]	Find, does, keeping, started
[lemma]	Lemmas (all forms of a word)	[sing] [tall]	Sing, singing, sang Tall, taller, tallest
[=word]	Synonyms [=strong]	Formidable, muscular, fervent	
[user:list]	Customized lists	[mark_davies@byu.edu:clothes]	Tie, shirt, blouse

Simple slot			
Syntax	Meaning	Examples	Sample matches
Wordword	Any of these words	stunninggorgeouscharming	Stunning, charming, gorgeous
*xx	Wildcard: * = any # letters	un*ly	Unlikely, unusually
X?xx	Wildcard: ? = one letter	s?ng	Sing, sang, song
X?xx*		s?ng*	Song, singer, songbirds
-word	NOT (followed by PoS, lemma, word, etc.)	-[nn*]	The, in, is
word.[pos]	Exact word and part of speech	strike.[v*]	Strike (only as a verb)
word*.[pos]	Substring and part of speech	dis*.[vvd]	Discovered, disappeared, discussed
[lemma].[pos]	Lemma and part of speech	[strike].[v*]	Strike, struck, striking
[=word].[pos]	Synonym and part of speech	[=beat].[v*]	Hit, strike, defeat (but not nouns, like <i>rhythm</i> or <i>drumming</i> )
[[=word]]	Synonym and lemma	[[=publish]]	Announced, circulating, publishes, issue (no part of speech specified, so some noun uses)
[[user:list]]	Customized list and lemma	[[mark_davies@byu.edu:clothes]]	Tie, tying, socks, socked, shirt, blouses (no part of speech specified, hence tying)
[[=word].[pos]	Synonym and lemma and part of speech	[[=clean].[v*]	Mop, scrubs, polishing
[[user:list]].[pos]	Customized list and lemma and part of speech	[[mark_davies@byu.edu:clothes]].[n*]	Tie, ties, sock, socks (i.e., just nouns)

Multiple slots	
Syntax	Examples
fastquickrapid [nn*]	Fast food, rapid transit
Pretty -[nn*]	Pretty smart' pretty as (but not pretty <i>girl</i> , pretty <i>picture</i> , etc.)
[get] her to [v*]	Get her to stay, got her to sleep
.!,: nevertheless [p*] [v*]	. Nevertheless it is ; nevertheless he said
[break] the [nn*]	Break the law, broke the story
[[beat]].[v*] * [nn*]	Beat the Yankees, beaten to death
[=gorgeous] [nn*]	Beautiful woman, attractive wife
[put] on [ap*] [mark_davies@byu.edu:clothes].[n*]	Put on her hat, putting on my pants

### Appendix 4: Systems for Statistics (Figs. A1 and A2)

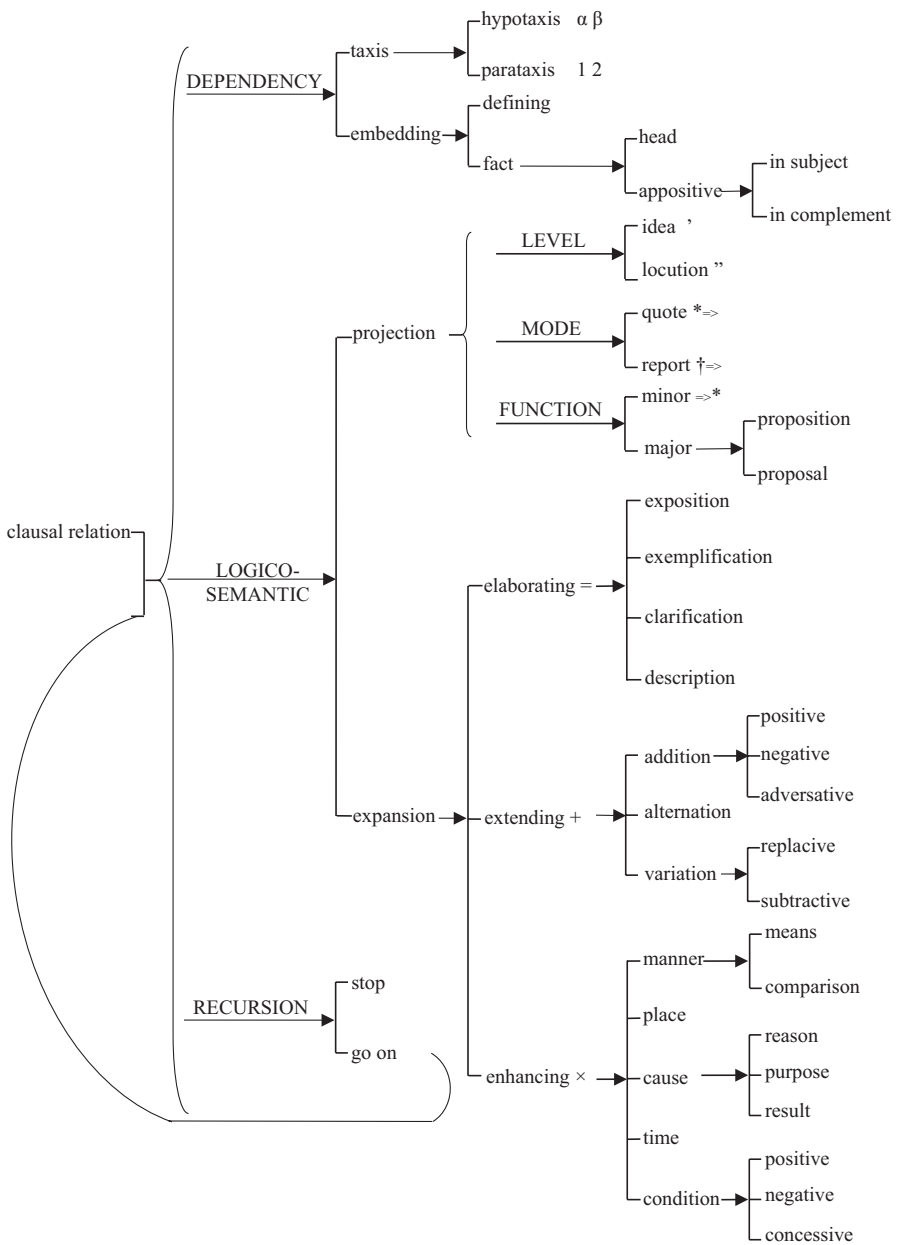


Fig. A1 System network of clausal relations in English

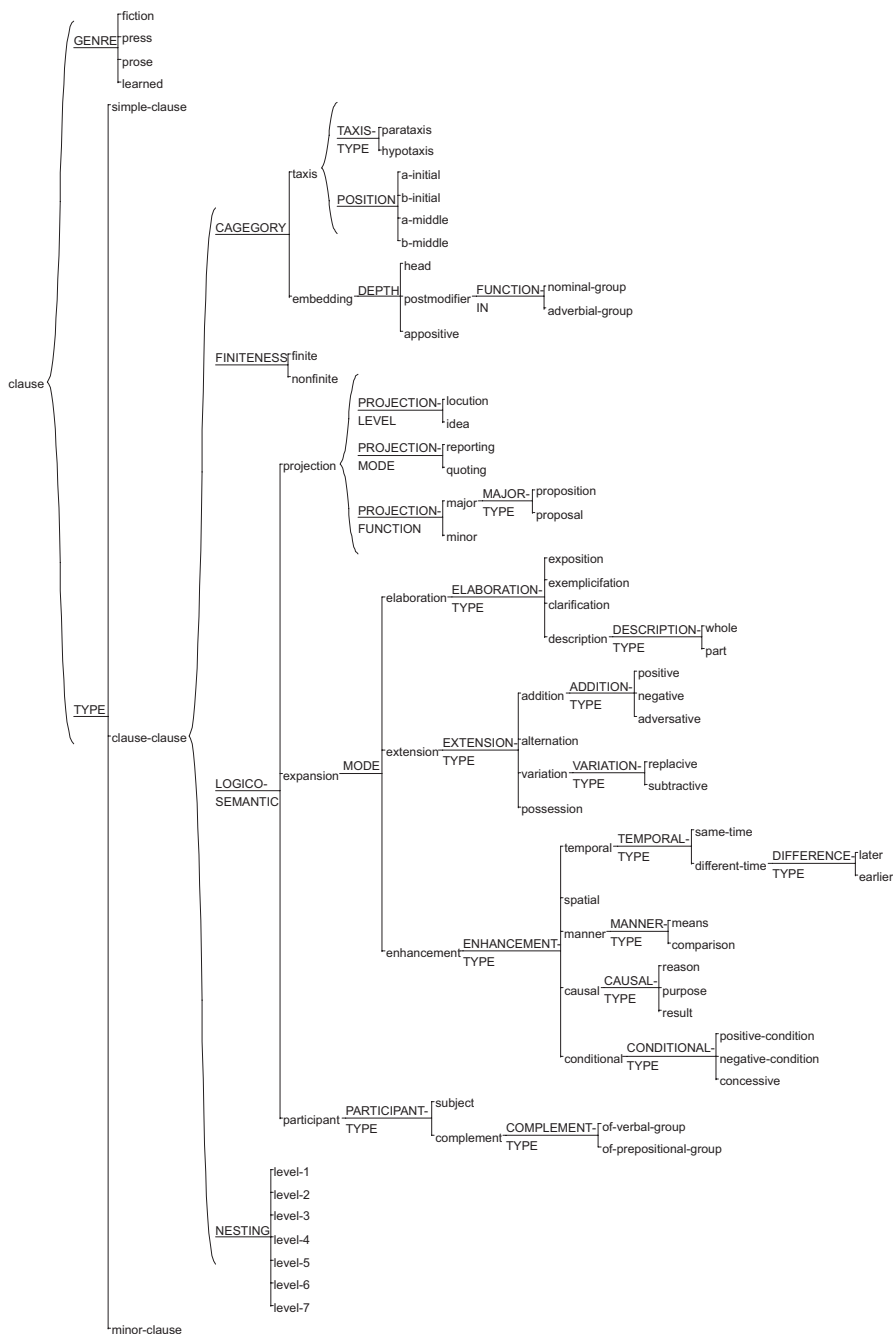


Fig. A2 System network of clause combining working in the UAM Corpus Tool



## Appendix 5: Table of Corpus Data Statistics (Table A1)

**Table A1** Genre distributions of clause combining in the Brown Corpus

Feature	Fiction		Press		Prose		Learned	
	Percent	N	Percent	N	Percent	N	Percent	N
Total units	N = 548		N = 644		N = 637		N = 595	
Simple-clause	23.0	126	17.1	110	20.3	129	17.1	102
Clause-clause	76.5	419	82.9	534	79.7	508	82.9	493
Minor-clause	0.5	3	0.0	0	0.0	0	0.0	0
Category	N = 419		N = 534		N = 508		N = 493	
Taxis	80.4	337	67.6	361	70.1	356	67.7	334
Embedding	19.6	82	32.4	173	29.9	152	32.3	159
Taxis	N = 337		N = 361		N = 356		N = 334	
Parataxis	51.0	172	27.4	99	36.8	131	31.4	105
Hypotaxis	49.0	165	72.6	262	63.2	225	68.6	229
Position	N = 337		N = 361		N = 356		N = 334	
a-initial	77.4	261	81.4	294	92.4	329	81.4	272
b-initial	19.0	64	16.1	58	7.3	26	18.0	60
a-middle	1.8	6	0.6	2	0.0	0	0.6	2
b-middle	1.8	6	1.9	7	0.3	1	0.0	0
Depth	N = 82		N = 173		N = 152		N = 159	
Head	37.8	31	37.0	64	17.8	27	18.2	29
Post-modifier	62.2	51	57.8	100	82.2	125	74.2	118
Appositive	0.0	0	5.2	9	0.0	0	7.5	12
Function-in	N = 51		N = 100		N = 125		N = 118	
Nominal-group	84.3	43	94.0	94	99.2	124	99.2	117
Adverbial-group	15.7	8	6.0	6	0.8	1	0.8	1
Finiteness	N = 419		N = 534		N = 508		N = 493	
Finite	70.6	296	67.4	360	44.3	225	60.9	300
Nonfinite	29.4	123	32.6	174	55.7	283	39.1	193
Logico-semantic	N = 419		N = 534		N = 508		N = 493	
Projection	20.5	86	25.8	138	5.9	30	14.0	69
Expansion	72.1	302	62.2	332	88.8	451	80.1	395
Participant	7.4	31	12.0	64	5.3	27	5.9	29
Projection-level	N = 86		N = 138		N = 30		N = 69	
Locution	53.5	46	81.2	112	93.3	28	52.2	36
Idea	46.5	40	18.8	26	6.7	2	47.8	33
Projection-mode	N = 86		N = 138		N = 30		N = 69	
Reporting	55.8	48	81.2	112	100.0	30	79.7	55
Quoting	44.2	38	18.8	26	0.0	0	20.3	14
Projection-function	N = 86		N = 138		N = 30		N = 69	
Major	96.5	83	99.3	137	100.0	30	100.0	69
Minor	3.5	3	0.7	1	0.0	0	0.0	0

(continued)

**Table A1** (continued)

Feature	Fiction		Press		Prose		Learned	
	Percent	N	Percent	N	Percent	N	Percent	N
Major	N = 83		N = 137		N = 30		N = 69	
Proposition	88.0	73	99.3	136	100.0	30	98.6	68
Proposal	12.0	10	0.7	1	0.0	0	1.4	1
Mode	N = 302		N = 332		N = 451		N = 395	
Elaboration	20.9	63	37.7	125	37.5	169	37.5	148
Extension	37.7	114	27.4	91	29.9	135	28.1	111
Enhancement	41.4	125	34.9	116	32.6	147	34.4	136
Elaboration	N = 63		N = 125		N = 169		N = 148	
Exposition	11.1	7	1.6	2	1.8	3	2.0	3
Exemplification	3.2	2	0.0	0	0.6	1	0.0	0
Clarification	3.2	2	0.8	1	0.0	0	0.0	0
Description	82.5	52	97.6	122	97.6	165	98.0	145
Description	N = 52		N = 122		N = 165		N = 145	
Whole	5.8	3	2.5	3	4.8	8	4.8	7
Part	94.2	49	97.5	119	95.2	157	95.2	138
Extension	N = 114		N = 91		N = 135		N = 111	
Addition	96.5	110	91.2	83	90.4	122	92.8	103
Alternation	0.9	1	4.4	4	6.7	9	2.7	3
Variation	1.8	2	2.2	2	1.5	2	2.7	3
Possession	0.9	1	2.2	2	1.5	2	1.8	2
Addition	N = 110		N = 83		N = 122		N = 103	
Positive	86.4	95	81.9	68	98.4	120	92.2	95
Negative	0.0	0	1.2	1	0.0	0	1.0	1
Adversative	13.6	15	16.9	14	1.6	2	6.8	7
Variation	N = 2		N = 2		N = 2		N = 3	
Replacive	100.0	2	100.0	2	100.0	2	100.0	3
Subtractive	0.0	0	0.0	0	0.0	0	0.0	0
Enhancement	N = 125		N = 116		N = 147		N = 136	
Temporal	48.8	61	24.1	28	18.4	27	23.5	32
Spatial	4.0	5	9.5	11	6.1	9	5.1	7
Manner	11.2	14	7.8	9	4.8	7	18.4	25
Causal	28.8	36	45.7	53	57.1	84	36.0	49
Conditional	7.2	9	12.9	15	13.6	20	16.9	23
Temporal	N = 61		N = 28		N = 27		N = 32	
Same-time	45.9	28	71.4	20	40.7	11	78.1	25
Different-time	54.1	33	28.6	8	59.3	16	21.9	7
Difference	N = 33		N = 8		N = 16		N = 7	
Later	90.9	30	25.0	2	6.2	1	42.9	3
Earlier	9.1	3	75.0	6	93.8	15	57.1	4

(continued)

**Table A1** (continued)

Feature	Fiction		Press		Prose		Learned	
	Percent	N	Percent	N	Percent	N	Percent	N
Manner	N = 14		N = 9		N = 7		N = 25	
Means	0.0	0	44.4	4	71.4	5	84.0	21
Comparison	100.0	14	55.6	5	28.6	2	16.0	4
Causal	N = 36		N = 53		N = 84		N = 49	
Reason	11.1	4	24.5	13	9.5	8	36.7	18
Purpose	41.7	15	66.0	35	89.3	75	42.9	21
Result	47.2	17	9.4	5	1.2	1	20.4	10
Conditional	N = 9		N = 15		N = 20		N = 23	
Positive-condition	33.3	3	60.0	9	45.0	9	47.8	11
Negative-condition	0.0	0	6.7	1	5.0	1	0.0	0
Concessive	66.7	6	33.3	5	50.0	10	52.2	12
Participant	N = 31		N = 64		N = 27		N = 29	
Subject	25.8	8	28.1	18	3.7	1	27.6	8
Complement	74.2	23	71.9	46	96.3	26	72.4	21
Complement	N = 23		N = 46		N = 26		N = 21	
Of-verbal-group	87.0	20	26.1	12	11.5	3	14.3	3
Of-prepositional-group	13.0	3	73.9	34	88.5	23	85.7	18
Nesting	N = 419		N = 534		N = 508		N = 493	
Level-1	41.8	175	36.1	193	34.1	173	41.0	202
Level-2	37.9	159	40.4	216	42.3	215	42.2	208
Level-3	13.4	56	15.2	81	16.9	86	12.4	61
Level-4	4.5	19	6.0	32	5.3	27	3.9	19
Level-5	1.7	7	1.9	10	1.0	5	0.4	2
Level-6	0.5	2	0.4	2	0.2	1	0.2	1
Level-7	0.2	1	0.0	0	0.2	1	0.0	0