

The Quantitative Study of Style, Using Grammatical Tags : Short Stories of W. S. Maugham

Koumei Wada and Ken'ichi Yoshioka

1 Introduction

The objective of our research is to see if there is any difference among the short stories of W. Somerset Maugham (1874–1965) from the view points of grammatical tag, that is, 'part-of-speech'. Maugham is said to have written over one hundred and twenty short stories. Most of them were published in the collected form. We picked up two stories from each book of collected stories. The titles of the books or magazines, the titles of the stories, their notations, and their tokens are shown below :

Orientalions (1899)

<i>The Punctiliousness of Don Sebastian</i> (DON)	5,996 words
<i>Daisy</i> (DAI)	10,601 words

Unknown Magazines 1

<i>Lady Habbart</i> (1900) (HAB)	9,999 words
<i>The Fortunate Painter</i> (1906) (PAI)	3,095 words

The Trembling of a Leaf(1921)

<i>Rain</i> (RAI)	15,305 words
<i>Red</i> (RED)	8,931 words

The Casuarina Tree(1921)

<i>Before the Party</i> (BEF)	10,144 words
<i>The Letter</i> (LET)	13,146 words

Ashenden(1928)

<i>Giulia Lazzari</i> (GIU)	7,013 words
<i>His Excellency</i> (EXC)	13,298 words

The Gentleman in the Parlour (1930)

<i>Mabel</i> (MBL)	1,641 words
<i>Marriage of Convenience</i> (MRR)	5,352 words

First Person Singular (1931)

<i>Jane</i> (JAN)	9,574 words
-------------------	-------------

<i>The Creative Impulse</i> (IMP)	13,482 words
Ah King(1933)	
<i>The Book-Bag</i> (BAG)	14,850 words
<i>Neil MacAdams</i> (NEL)	17,925 words
Cosmopolitans(1936)	
<i>The Portrait of a Gentleman</i> (GTL)	2,156 words
<i>Home</i> (HOM)	1,837 words
The Mixture as Before (1940)	
<i>Lion's Skin</i> (LIO)	9,401 words
<i>Treasure</i> (TRE)	6,255 words
Creatures of Circumstance(1947)	
<i>The Colne's Lady</i> (COL)	7,188 words
<i>Sanatorium</i> (SAN)	9,852 words

2 Methodology

1) We assigned a grammatical (part-of-speech) tag to every word appearing in the 22 texts, using CLAWS 7, which is an automatic part-of-speech tagging system and has

Table 1 The revised tagset

AP	APPGE
AT	AT, ATI
B	BCL
C	CC, CCB, CS, CSA, CSN, CST, CSW
D	DA, DA 1, DA 2, DAR, DAT, DB, DB 2, DD, DD 1, DD 2, DDQ, DDQGE, DDQV
EX	EX
FO	FO
FU	FU
FW	FW
GE	GE
I	IF, II, IO, IW
J	JJ, JJR, JJT, JK
M	MC, MC 1, MC 2, MCGE, MCMC, MD, MF
NM	ND 1, NN, NN 1, NN 2, NNA, NNB, NNL 1, NNL 2, NNO, NNO 2, NNT 1, NNT 2, NNU, NNU 1, NNU 2
NP	NP, NP 1, NP 2, NPD 1, NPD 2, NPM 1, NPM 2
P	PN, PN 1, PNQO, PNQS, PNQV, PNX 1, PPGE, PPH 1, PPHO 1, PPHO 2, PPHS 1, PPHS 2, PPIO 1, PPIO 2, PPIS 1, PPIS 2, PPX 1, PPX 2, PPY
R	RA, REX, RG, RGQ, RGQV, RGR, RGT, RL, RP, RPK, RR, RRQ, RRQV, RRR, RRT, RT
TO	TO
UH	UH
VB	VB 0, VBDR, VBDZ, VBG, VBI, VBM, VBN, VBR, VBZ
VD	VD 0, VDD, VDG, VDI, VDN, VDZ
VH	VH 0, VHD, VHG, VHI, VHN, VHZ
VM	VM, VMK
VV	VV 0, VVD, VVG, VVGK, VVI, VVN, VVNK, VVZ
X	XX
Z	ZZ 1, ZZ 2

been developed by UCREL at Lancaster University. The tagset to be used with CLAWS 7 contains 137 tags. The CLAWS 7 tagset is shown in Appendix A. And a part of our tagged text is also shown in Appendix B. For our study, however, we summed up 137 tags to 26. We thought it would be better to classify the 137 tags into such groups as traditional parts of speech. Table 1 is the tagset we have revised for our research.

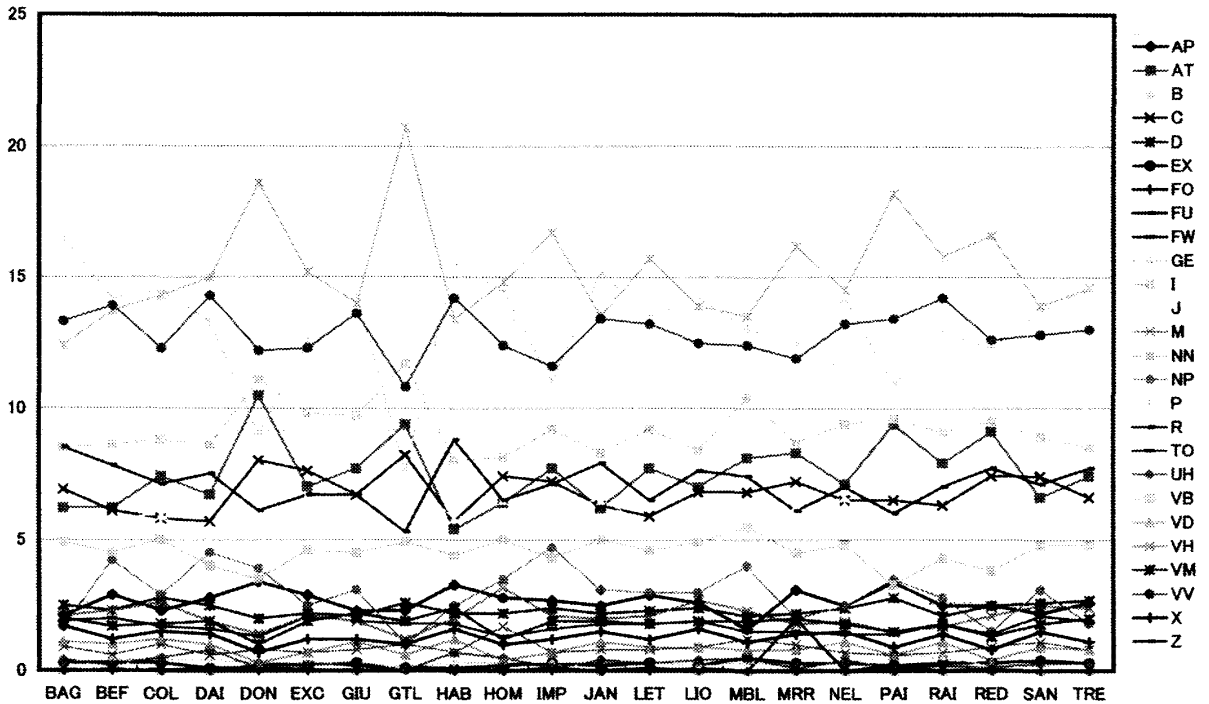


Fig. 1 (a) Frequency Polygon of Texts

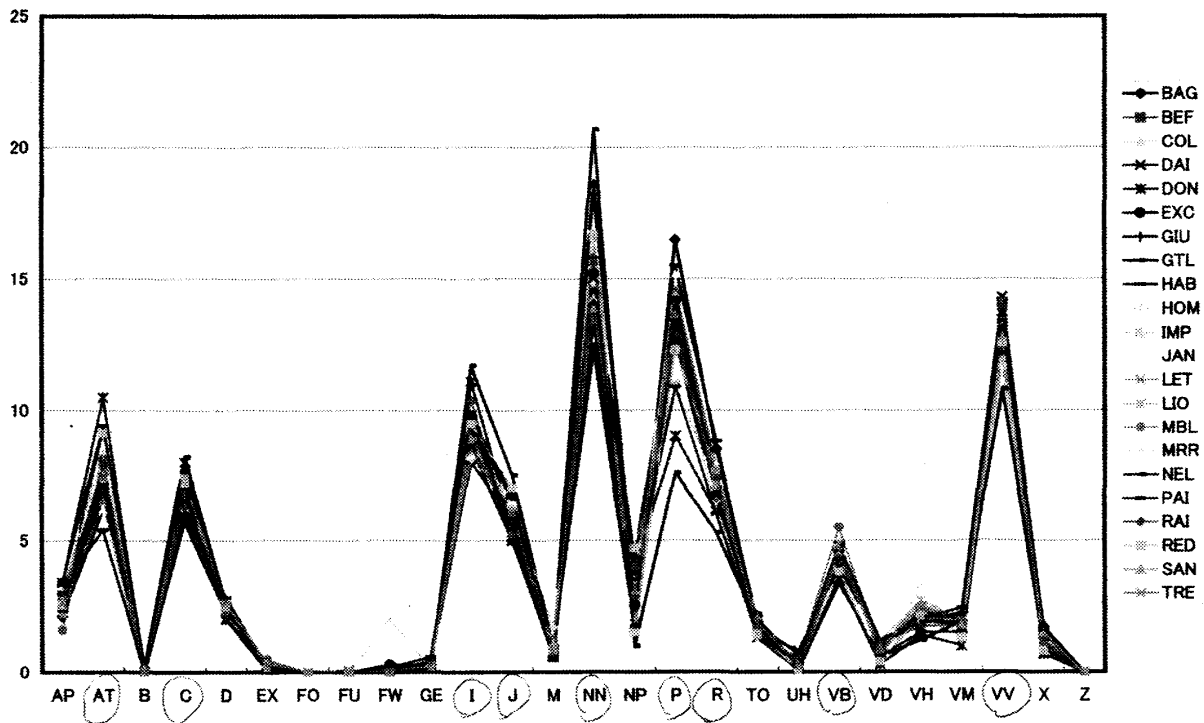


Fig. 1 (b) Frequency Polygon of Part-of-Speeches

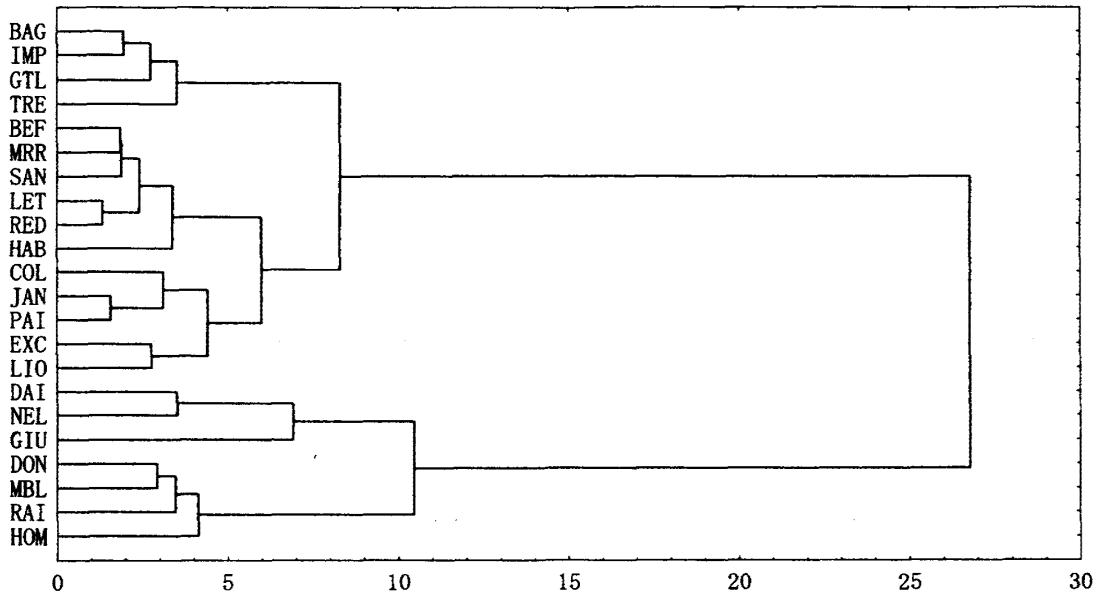


Fig. 2 Dendrogram Using Euclid Distance of Ward's Method

2) We counted the occurrences of grammatical tags in each text. Appendix C is the frequency distribution of grammatical tags (shown in percentage).

3) On the basis of the frequency distribution of grammatical tags in each text (Appendix C), we constructed frequency polygon of grammatical tags in each text (Figure 1).

4) Using grammatical tags as variables and the 22 texts as cases, we constructed a dendrogram to see if there appear some clusters of texts. In this case we made use of Euclid distance of Ward's method. The dendrogram is shown in Figure 2.

3 Results

1) It can be seen that there are three groups in the way the grammatical tags occur in the texts (see Figure 1): the first group's frequency is more than 10% , that of the second between 5 and 10%, and that of the third less than 5%. P (pronouns), VV (lexical verbs), and NN (nouns) belong to the first group, I (prepositions), R (adverbs), C (conjunctions), AT (articles), and J (adjectives) to the second group, the remaining 18 tags to the third group. The tag with the highest frequency is NN, followed by P, VV, I, respectively and AT. FO(formula) and Z (alphabet) do not occur in any text. As for FU(unclassified words), it occurs in two texts with the frequency of 0.1%.

2) Judging from the dendrogram (Figure 2), the 22 texts can be classified into three groups. In order to speculate the features each group has, we constructed three frequency polygons (Figure 3, a, b, c).

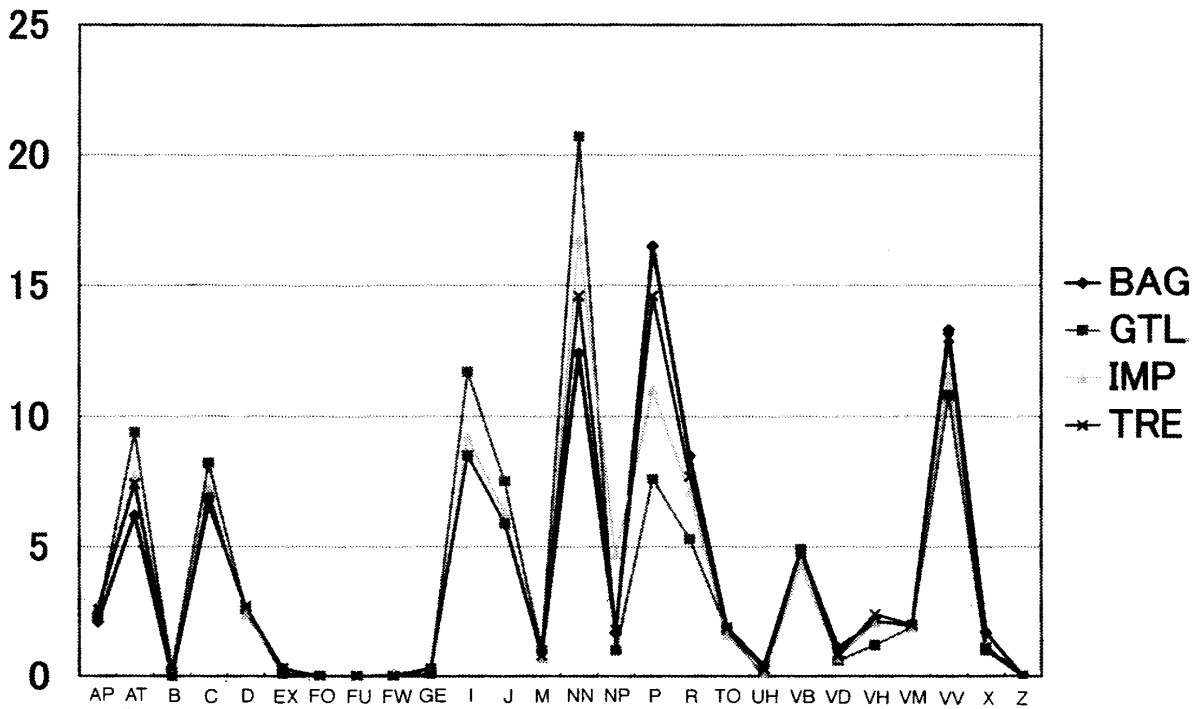


Fig. 3 (a) Frequency Polygon of Part-of-Speeches

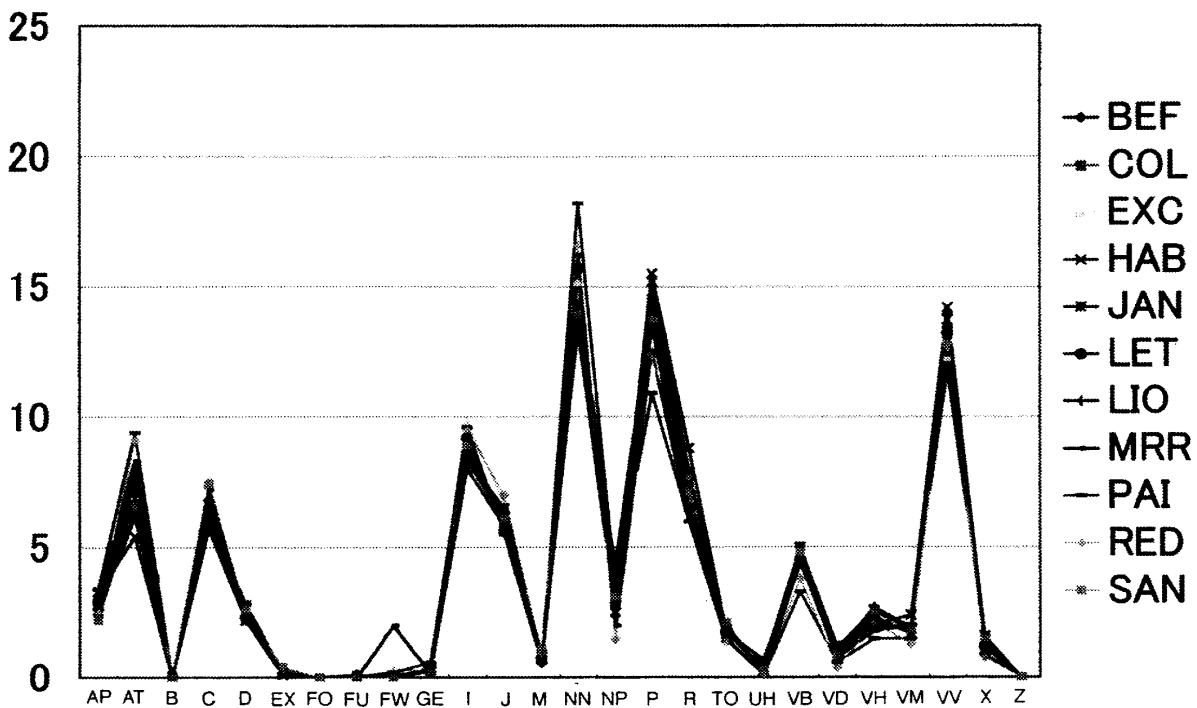


Fig. 3 (b)

There seems a general tendency that the works in group A (Figure 3, a) have higher frequency in almost all the tags than those of the other two groups. The difference between group B (Figure 3, b) and group C (Figure 3, c) can be seen in that the disparity between the highest and the lowest frequency of each tag in the texts of group B is

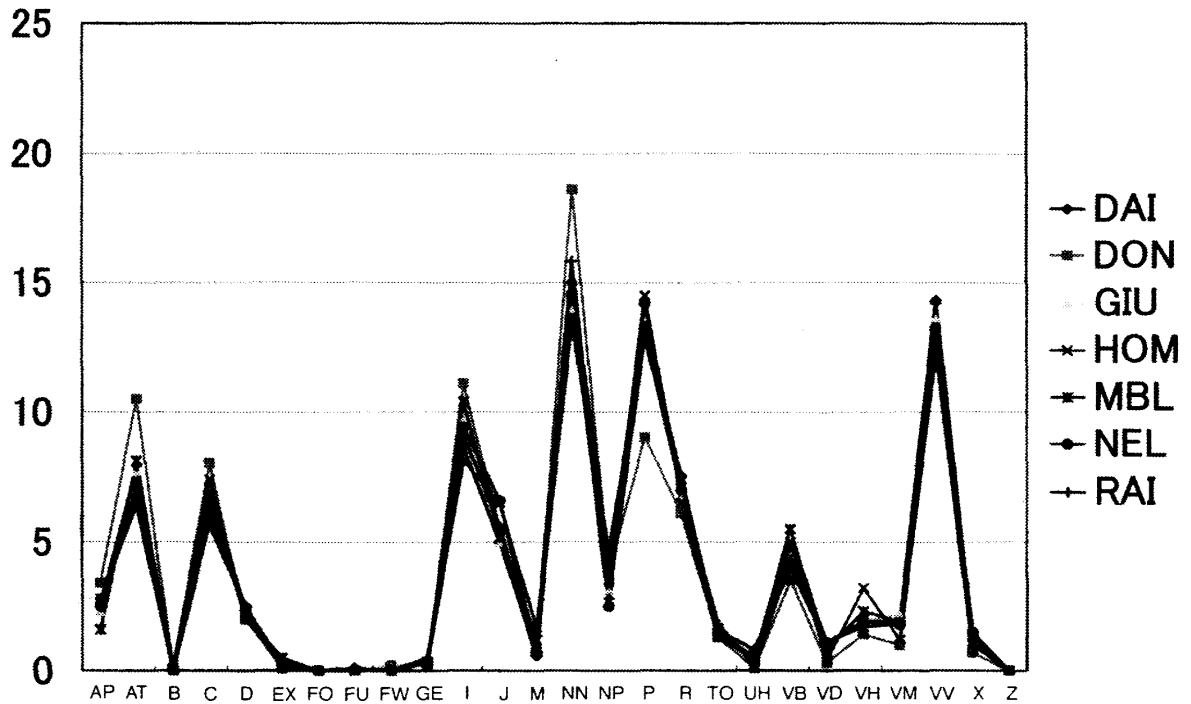


Fig. 3 (c)

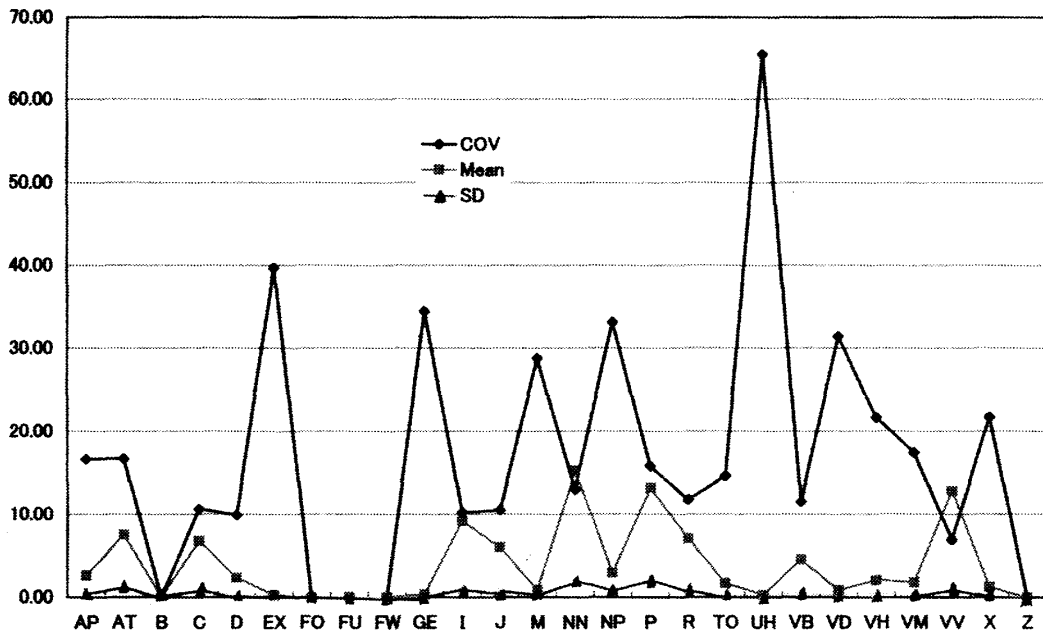


Fig. 4 Coefficient of Variation

smaller than that in the texts of group C. Though the frequency of each tag is generally low in group C, the disparity between the highest and the lowest frequency of each tag is wider than the other two groups. We worked out the coefficient of variation to know the

disparity between the highest and the lowest frequency of each grammatical tag (Figure 4). LET and RED are much alike in their frequency spectra.

4 Discussion

It had been our original expectation that we would be able to confirm the first four stories (DON, DAI, HAB, and PAI) would group together when we carried out cluster analysis on the basis of the grammatical tags. These four stories were written in the first part of Maugham's apprenticeship, from 1897 to 1907²⁾. R. A. Cordel wrote they are interesting only as the youthful work of one who later tried seriously to improve his style. The prose has little of the simplicity, suppleness, and rhythm of his later work³⁾. Maugham also stated "they are so immature. . . and are best forgotten."⁴⁾ But, in fact, there seems almost no difference between the stories written in his apprenticeship and all the rest of the stories, which were written after 1921, from the view-point of grammatical tags. As Figure 2 shows, almost all the grammatical tags are used at a fixed ratio in every text: the first group more than 10%, the second group between 5 and 10%, and the third group less than 5%.

We had expected the chronological development of Maugham's style. However, in terms of grammatical tags, it can be seen that he did not vary his style from his early age to old age. For our further study, we think we have to use more detailed grammatical tags and his word habits as raw data so that we can see the broader picture of Maugham's style.

Notes

- 1) CLAWS (the Constituent Likelihood Automatic Word-tagging System.)

The CLAWS tagging system consists of five separate stages applied successively to a text to be tagged. The first step is pre-editing where the text is cleaned and verticalised (one word is printed above another). This is followed by candidate tag assignment, where each possible tag that might apply to a word is assigned in descending order of likelihood. Thirdly, multi-word units such as idioms are tagged as single items. The fourth step is tag disambiguation: this stage inspects all cases where a word has been assigned more than one tag, and attempts to choose a preferred tag by considering the context in which the word appears, and assessing the probability of any particular sequence of tags. The final phase is manual post-editing, in which erroneous tagging decisions made by the computer are corrected by human editors. (Michael P. Oakes, *Statistics for Corpus Linguistics* (Edinburgh: Edinburgh UP., 1998), p. 81)

UCREL (Unit for Computer Research on the English Language) claims: CLAWS has consistently achieved 96–97% accuracy. The precise degree of accuracy varies according to the type of text. Judged in terms of major categories, the system has an approximate error-rate

of only 1.5%, with c. 3.3% ambiguities unresolved, within the BNC (British National Corpus).

As to the works of Maugham, accuracy does not come up to 96–97%. And it took more than six months to post-edit all his short stories

- 2) Forrest D. Burt, *W. Somerset Maugham* (Boston : Twayne Publishers, 1985), p. 18.
- 3) Richard A. Cordel, *Somerset Maugham* (Bloomington : Indiana U. P.), p. 166.
- 4) Maugham says in the preface of the first volume of *The Complete Short Stories of W. Somerset Maugham Vols. 3* (1951): In my early youth I wrote a number, but they are so immature that I have preferred not to print them. A few are in a book that has long remained out of print, a few others are scattered in various magazines.

In 1969, four years after Maugham's death, *Seventeen Lost Stories* by W. Somerset Maugham was published in New York. This contains the 6 stories which made up his earliest book, *Orientalism* (1899). This book, just as Maugham said above, refers to 'a book that has long remained out of print', and the remaining 11 stories refer to 'a few others. . . .'

References

- Dixon, P. and Mannion, D. (1993), 'Goldsmith's Periodical Essays : A Statistical Analysis of Eleven Doubtful Cases', *Literary and Linguistic Computing*, Vol. 8, No. 1.
- Garside, R., Leech, G. and Sampson, G. (ed.) (1987), *The Computational Analysis of English : A corpus-based approach*, London : Longman.
- Kenny, A. J. P. (1982), *The Computation of Style*, Oxford : Pergamon Press.
- McEnery, A. M. and Wilson, A. (1996), *Corpus Linguistics*, Edinburgh : Edinburgh UP.
- Merriam, T. (1997), 'Invalidation Reappraised', *Computers and Humanities* 30 : 417–431.
- Oakes, P. O. (1998), *Statistics for Corpus Linguistics*, Edinburgh : Edinburgh UP.

Appendix A : The C 7 tagset

APPGE	possessive pronoun, pre-nominal (e.g. my, your, our)
AT	article (e.g. the, no)
ATI	singular article (e.g. a, an, every)
BCL	before-clause marker (e.g. in order (that), in order (to))
CC	co-ordinating conjunction (e.g. and, or)
CCB	adversative co-ordinating conjunction (but)
CS	subordinating conjunction (e.g. if, because, unless, so, for)
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. such, former, same)
DA 1	singular after-determiner (e.g. little, much)
DA 2	plural after-determiner (e.g. few, several, many)
DAR	comparative after-determiner (e.g. more, less, fewer)
DAT	superlative after-determiner (e.g. most, least, fewest)
DB	before determiner or pre-determiner capable of pronominal function (all, half)
DB 2	plural before-determiner (both)
DD	determiner (capable of pronominal function) (e.g. any, some)
DD 1	singular determiner (e.g. this, that, another)
DD 2	plural determiner (these, those)

DDQ	wh-determiner (which, what)
DDQGE	wh-determiner, genitive (whose)
DDQV	wh-ever determiner (whichever, whatever)
EX	existential there
FO	formula
FU	unclassified word
FW	foreign word
GE	germanic genitive marker-(‘or’ s)
IF	for (as preposition)
II	general preposition
IO	of (as preposition)
IW	with, without (as prepositions)
JJ	general adjective
JJR	general comparative adjective (e.g. older, better, stronger)
JJT	general superlative adjective (e.g. oldest, best, strongest)
JK	catenative adjective (able in be able to, willing in be willing to)
MC	cardinal number, neutral for number (two, three. .)
MC 1	singular cardinal number (one)
MC 2	plural cardinal number (e.g. sixes, sevens)
MCGE	genitive cardinal number, neutral for number (two’s, 100’s)
MCMC	hyphenated number (40–50, 1770–1827)
MD	ordinal number (e.g. first, second, next, last)
MF	fraction, neutral for number (e.g. quarters, two-thirds)
NDI	singular noun of direction (e.g. north, south-east)
NN	common noun, neutral for number (e.g. sheep, cod, headquarters)
NN 1	singular common noun (e.g. book, girl)
NN 2	plural common noun (e.g. books, girls)
NNA	following noun of title (e.g. M.A.)
NNB	preceding noun of title (e.g. Mr., Prof.)
NNL 1	singular locative noun (e.g. Island, Street)
NNL 2	plural locative noun (e.g. Islands, Streets)
NNO	numeral noun, neutral for number (e.g. dozen, hundred)
NN 02	numeral noun, plural (e.g. hundreds, thousands)
NNT 1	temporal noun, singular (e.g. day, week, year)
NNT 2	temporal noun, plural (e.g. days, weeks, years)
NNU	unit of measurement, neutral for number (e.g. in, cc)
NNU 1	singular unit of measurement (e.g. inch, centimetre)
NNU 2	plural unit of measurement (e.g. ins, feet)
NP	proper noun, neutral for number (e.g. IBM, Andes)
NP 1	singular proper noun (e.g. London, Jane, Frederick)
NP 2	plural proper noun (e.g. Browns, Reagans, Koreas)
NPD 1	singular weekday noun (e.g. Sunday)
NPD 2	plural weekday noun (e.g. Sundays)
NPM 1	singular month noun (e.g. October)
NPM 2	plural month noun (e.g. Octobers)
PN	indefinite pronoun, neutral for number (none)
PN 1	indefinite pronoun, singular (e.g. anyone, everything, nobody, one)
PNQO	objective wh-pronoun (whom)
PNQS	subjective wh-pronoun (who)
PNQV	wh-ever pronoun (whoever)
PNX 1	reflexive indefinite pronoun (oneself)
PPGE	nominal possessive personal pronoun (e.g. mine, yours)
PPH 1	3 rd person sing. neuter personal pronoun (it)
PPHO 1	3 rd person sing. objective personal pronoun (him, her)
PPHO 2	3 rd person plural objective personal pronoun (them)
PPHS 1	3 rd person sing. subjective personal pronoun (he, she)
PPHS 2	3 rd person plural subjective personal pronoun (they)
PPIO 1	1 st person sing. objective personal pronoun (me)
PPI 02	1 st person plural objective personal pronoun (us)
PPIS 1	1 st person sing. subjective personal pronoun (I)
PPIS 2	1 st person plural subjective personal pronoun (we)
PPX 1	singular reflexive personal pronoun (e.g. yourself, itself)
PPX 2	plural reflexive personal pronoun (e.g. yourselves, themselves)
PPY	2 nd person personal pronoun (you)
RA	adverb, after nominal head (e.g. else, galore)

REX	adverb introducing appositional constructions (namely, e.g.)
RG	degree adverb (very, so, too)
RGQ	wh-degree adverb (how)
RGQV	wh-ever degree adverb (however)
RGR	comparative degree adverb (more, less)
RGT	superlative degree adverb (most, least)
RL	locative adverb (e.g. alongside, forward)
RP	prep. adverb, particle (e.g. about, in)
RPIC	prep. adv., catenative (about in be about to)
RR	general adverb
RRQ	wh-general adverb (where, when, why, how)
RRQV	wh-ever general adverb (wherever, whenever)
RRR	comparative general adverb (e.g. better, longer)
RRT	superlative general adverb (e.g. best, longest)
RT	quasi-nominal adverb of time (e.g. now, tomorrow)
TO	infinitive marker (to)
UH	interjection (e.g. oh, yes, um)
VB 0	be, base form (finite i.e. imperative, subjunctive)
VBDR	were
VBDZ	was
VBG	being,
VBI	be, infinitive (To be or not. . . It will be)
VBM	am
VBN	been
VBR	are
VBZ	is
VDO	do, base form (finite)
VDD	did
VDG	doing
VDI	do, infinitive (I may do. . . To do. . .)
VDN	done
VDZ	does
VH 0	have, base form (finite)
VHD	had (past tense)
VHG	having
VHI	have, infinitive
VHN	had (past particle)
VHZ	has
VM	modal auxiliary (can, will, would, etc)
VMK	modal catenative (ought, used)
VV 0	base form of lexical verb (e.g. give, work)
VVD	past tense of lexical verb (e.g. gave, worked)
VVG	-ing particle of lexical verb (e.g. giving, working)
VVGK	-ing particle catenative (going in be going to)
VVI	infinitive (e.g. to give. . . It will work. . .)
VVN	past participle of lexical verb (e.g. given, worked)
VVNK	past participle catenative (e.g. bound in be bound to)
VVZ	-s form of lexical verb (e.g. gives, works)
XX	not, n't
ZZ 1	singular letter of the alphabet (e.g. A, b)
ZZ 2	plural letter of the alphabet (e.g. A's, b's)

Appendix B: A part of a tagged text

^ <BAG>__NULL
 ^ Some_DD people_NN read_VV 0 for_IF instruction_NN 1 ,_,
 which_DDQ
 is_VBZ praiseworthy_JJ ,_, and_CC some_DD for_IF pleasure_NN 1
 ,_, which_DDQ
 is_VBZ innocent_JJ ,_, but_CCB not_XX a_AT 1 few_DA 2 read_VV 0
 from_II habit_NN 1
 ,_, and_CC I_PPIS 1 suppose_VV 0 that_CST this_DD 1 is_VBZ
 neither_RR innocent_JJ
 nor_CC praiseworthy_JJ ._.
 ^ Of_IO that_DD 1 lamentable_JJ company_NN 1 am_VBM I_PPIS 1 ._.

^ Conversation_NN 1 after_II a_AT 1 time_NNT 1 bores_VVZ me_PPIO 1
 ,,, games_NN 2
 tire_VV 0 me_PPIO 1 ,,, and_CC my_APPGE own_DA thoughts_NN 2 ,,,
 which_DDQ
 we_PPIS 2 are_VBR told_VVN are_VBR the_AT unfailing_JJ
 resource_NN 1 of_IO a_AT 1
 sensible_JJ man_NN 1 ,,, have_VH 0 a_AT 1 tendency_NN 1 to_TO
 run_VVI dry_JJ ._.
 ^ Then_RT I_PPIS 1 fly_VV 0 to_II my_APPGE book_NN 1 as_II the_AT
 opium-smoker_NN 1 to_II his_APPGE pipe_NN 1 ._.
 ^ I_PPIS 1 would_VM sooner_RRR read_VVI the_AT catalogue_NN 1
 of_IO the_AT
 army_NN 1 and_CC navy_NN 1 stores_NN 2 or_CC Bradshaw_NP 1 's_GE
 guide_NN 1
 than_CSN nothing_PN 1 at_RR 21 all_RR 22 ,,, and_CC indeed_RR

Appendix C : A frequency distribution of grammatical tags

BAG			(17) R	806	(7.8%)
(1) AP	322	(2.1%)	(18) TO	210	(2.0%)
(2) AT	943	(6.2%)	(19) UH	19	(0.2%)
(3) B	2	(0.0%)	(20) VB	462	(4.5%)
(4) C	1051	(6.9%)	(21) VD	105	(1.0%)
(5) D	375	(2.5%)	(22) VH	232	(2.3%)
(6) EX	39	(0.3%)	(23) VM	171	(1.7%)
(7) FO	0	(0.0%)	(24) VV	1429	(13.9%)
(8) FU	0	(0.0%)	(25) X	127	(1.2%)
(9) FW	7	(0.0%)	(26) Z	0	(0.0%)
(10) GE	43	(0.3%)			
(11) I	1288	(8.5%)	COL		
(12) J	897	(5.9%)	(1) AP	168	(2.3%)
(13) M	133	(0.9%)	(2) AT	547	(7.4%)
(14) N	1883	(12.4%)	(3) B	0	(0.0%)
(15) NP	251	(1.7%)	(4) C	432	(5.8%)
(16) P	2504	(16.5%)	(5) D	205	(2.8%)
(17) R	1297	(8.5%)	(6) EX	26	(0.3%)
(18) TO	284	(1.9%)	(7) FO	0	(0.0%)
(19) UH	60	(0.4%)	(8) FU	0	(0.0%)
(20) VB	752	(4.9%)	(9) FW	0	(0.0%)
(21) VD	170	(1.1%)	(10) GE	30	(0.4%)
(22) VH	312	(2.1%)	(11) I	658	(8.8%)
(23) VM	301	(2.0%)	(12) J	442	(5.9%)
(24) VV	2030	(13.3%)	(13) M	71	(1.0%)
(25) X	266	(1.7%)	(14) N	1066	(14.3%)
(26) Z	1	(0.0%)	(15) NP	216	(2.9%)
			(16) P	1077	(14.5%)
BEF			(17) R	530	(7.1%)
(1) AP	296	(2.9%)	(18) TO	125	(1.7%)
(2) AT	642	(6.2%)	(19) UH	34	(0.5%)
(3) B	6	(0.1%)	(20) VB	369	(5.0%)
(4) C	628	(6.1%)	(21) VD	89	(1.2%)
(5) D	237	(2.3%)	(22) VH	197	(2.6%)
(6) EX	27	(0.3%)	(23) VM	131	(1.8%)
(7) FO	0	(0.0%)	(24) VV	919	(12.3%)
(8) FU	0	(0.0%)	(25) X	110	(1.5%)
(9) FW	5	(0.0%)	(26) Z	0	(0.0%)
(10) GE	38	(0.4%)			
(11) I	883	(8.6%)	DAI		
(12) J	612	(5.9%)	(1) AP	311	(2.8%)
(13) M	64	(0.6%)	(2) AT	734	(6.7%)
(14) N	1417	(13.7%)	(3) B	4	(0.0%)
(15) NP	431	(4.2%)	(4) C	629	(5.7%)
(16) P	1459	(14.2%)	(5) D	275	(2.5%)

(6) EX	13	(0.1%)	(14) N	2033	(15.2%)
(7) FO	0	(0.0%)	(15) NP	336	(2.5%)
(8) FU	16	(0.1%)	(16) P	1671	(12.5%)
(9) FW	0	(0.0%)	(17) R	899	(6.7%)
(10) GE	47	(0.4%)	(18) TO	276	(2.1%)
(11) I	944	(8.6%)	(19) UH	32	(0.2%)
(12) J	548	(5.0%)	(20) VB	617	(4.6%)
(13) M	62	(0.6%)	(21) VD	97	(0.7%)
(14) N	1641	(15.0%)	(22) VH	286	(2.1%)
(15) NP	497	(4.5%)	(23) VM	249	(1.9%)
(16) P	1454	(13.3%)	(24) VV	1639	(12.3%)
(17) R	822	(7.5%)	(25) X	158	(1.2%)
(18) TO	179	(1.6%)	(26) Z	6	(0.0%)
(19) UH	93	(0.8%)			
(20) VB	440	(4.0%)	GIU		
(21) VD	106	(1.0%)	(1) AP	337	(2.3%)
(22) VH	212	(1.9%)	(2) AT	1106	(7.7%)
(23) VM	206	(1.9%)	(3) B	4	(0.0%)
(24) VV	1569	(14.3%)	(4) C	963	(6.7%)
(25) X	156	(1.4%)	(5) D	306	(2.1%)
(26) Z	2	(0.0%)	(6) EX	41	(0.3%)
			(7) FO	0	(0.0%)
DON			(8) FU	3	(0.0%)
(1) AP	203	(3.4%)	(9) FW	31	(0.2%)
(2) AT	630	(10.5%)	(10) GE	39	(0.3%)
(3) B	4	(0.1%)	(11) I	1403	(9.7%)
(4) C	479	(8.0%)	(12) J	727	(5.0%)
(5) D	119	(2.0%)	(13) M	120	(0.8%)
(6) EX	6	(0.1%)	(14) N	2016	(14.0%)
(7) FO	0	(0.0%)	(15) NP	447	(3.1%)
(8) FU	1	(0.0%)	(16) P	2059	(14.3%)
(9) FW	11	(0.2%)	(17) R	969	(6.7%)
(10) GE	20	(0.3%)	(18) TO	280	(1.9%)
(11) I	665	(11.1%)	(19) UH	47	(0.3%)
(12) J	331	(5.5%)	(20) VB	648	(4.5%)
(13) M	45	(0.8%)	(21) VD	157	(1.1%)
(14) N	1115	(18.6%)	(22) VH	269	(1.9%)
(15) NP	236	(3.9%)	(23) VM	318	(2.2%)
(16) P	536	(9.0%)	(24) VV	1970	(13.6%)
(17) R	365	(6.1%)	(25) X	174	(1.2%)
(18) TO	75	(1.3%)	(26) Z	1	(0.0%)
(19) UH	5	(0.1%)			
(20) VB	211	(3.5%)	GTL		
(21) VD	19	(0.3%)	(1) AP	50	(2.3%)
(22) VH	82	(1.4%)	(2) AT	202	(9.4%)
(23) VM	58	(1.0%)	(3) B	4	(0.2%)
(24) VV	733	(12.2%)	(4) C	176	(8.2%)
(25) X	39	(0.7%)	(5) D	55	(2.6%)
(26) Z	0	(0.0%)	(6) EX	3	(0.1%)
			(7) FO	0	(0.0%)
EXC			(8) FU	0	(0.0%)
(1) AP	387	(2.9%)	(9) FW	0	(0.0%)
(2) AT	938	(7.0%)	(10) GE	2	(0.1%)
(3) B	8	(0.1%)	(11) I	250	(11.7%)
(4) C	1009	(7.6%)	(12) J	160	(7.5%)
(5) D	287	(2.2%)	(13) M	22	(1.0%)
(6) EX	33	(0.2%)	(14) N	443	(20.7%)
(7) FO	0	(0.0%)	(15) NP	22	(1.0%)
(8) FU	0	(0.0%)	(16) P	163	(7.6%)
(9) FW	41	(0.3%)	(17) R	113	(5.3%)
(10) GE	38	(0.3%)	(18) TO	38	(1.8%)
(11) I	1314	(9.8%)	(19) UH	0	(0.0%)
(12) J	897	(6.7%)	(20) VB	105	(4.9%)
(13) M	97	(0.7%)	(21) VD	13	(0.6%)

(22) VH	26	(1.2%)
(23) VM	41	(1.9%)
(24) VV	230	(10.8%)
(25) X	21	(1.0%)
(26) Z	0	(0.0%)

HAB

(1) AP	334	(3.3%)
(2) AT	549	(5.4%)
(3) B	0	(0.0%)
(4) C	580	(5.7%)
(5) D	229	(2.2%)
(6) EX	10	(0.1%)
(7) FO	0	(0.0%)
(8) FU	2	(0.0%)
(9) FW	1	(0.0%)
(10) GE	29	(0.3%)
(11) I	810	(8.0%)
(12) J	586	(5.8%)
(13) M	69	(0.7%)
(14) N	1360	(13.4%)
(15) NP	257	(2.5%)
(16) P	1578	(15.5%)
(17) R	899	(8.8%)
(18) TO	187	(1.8%)
(19) UH	71	(0.7%)
(20) VB	446	(4.4%)
(21) VD	122	(1.2%)
(22) VH	208	(2.0%)
(23) VM	244	(2.4%)
(24) VV	1451	(14.2%)
(25) X	162	(1.6%)
(26) Z	0	(0.0%)

HOM

(1) AP	53	(2.8%)
(2) AT	119	(6.4%)
(3) B	0	(0.0%)
(4) C	138	(7.4%)
(5) D	41	(2.2%)
(6) EX	4	(0.2%)
(7) FO	0	(0.0%)
(8) FU	0	(0.0%)
(9) FW	0	(0.0%)
(10) GE	5	(0.3%)
(11) I	152	(8.1%)
(12) J	121	(6.5%)
(13) M	31	(1.7%)
(14) N	277	(14.8%)
(15) NP	65	(3.5%)
(16) P	271	(14.5%)
(17) R	122	(6.5%)
(18) TO	25	(1.3%)
(19) UH	10	(0.5%)
(20) VB	93	(5.0%)
(21) VD	8	(0.4%)
(22) VH	59	(3.2%)
(23) VM	22	(1.2%)
(24) VV	231	(12.4%)
(25) X	19	(1.0%)
(26) Z	0	(0.0%)

IMP

(1) AP	365	(2.7%)
(2) AT	1043	(7.7%)
(3) B	12	(0.1%)
(4) C	975	(7.2%)
(5) D	329	(2.4%)
(6) EX	46	(0.3%)
(7) FO	0	(0.0%)
(8) FU	0	(0.0%)
(9) FW	10	(0.1%)
(10) GE	46	(0.3%)
(11) I	1247	(9.2%)
(12) J	862	(6.3%)
(13) M	90	(0.7%)
(14) N	2273	(16.7%)
(15) NP	634	(4.7%)
(16) P	1502	(11.0%)
(17) R	972	(7.1%)
(18) TO	216	(1.6%)
(19) UH	17	(0.1%)
(20) VB	580	(4.3%)
(21) VD	97	(0.7%)
(22) VH	286	(2.1%)
(23) VM	255	(1.9%)
(24) VV	1577	(11.6%)
(25) X	164	(1.2%)
(26) Z	0	(0.0%)

JAN

(1) AP	242	(2.5%)
(2) AT	606	(6.2%)
(3) B	0	(0.0%)
(4) C	620	(6.3%)
(5) D	216	(2.2%)
(6) EX	20	(0.2%)
(7) FO	0	(0.0%)
(8) FU	0	(0.0%)
(9) FW	1	(0.0%)
(10) GE	29	(0.3%)
(11) I	814	(8.3%)
(12) J	631	(6.5%)
(13) M	75	(0.8%)
(14) N	1314	(13.5%)
(15) NP	306	(3.1%)
(16) P	1482	(15.2%)
(17) R	775	(7.9%)
(18) TO	178	(1.8%)
(19) UH	35	(0.4%)
(20) VB	485	(5.0%)
(21) VD	106	(1.1%)
(22) VH	192	(2.0%)
(23) VM	188	(1.9%)
(24) VV	1304	(13.4%)
(25) X	147	(1.5%)
(26) Z	1	(0.0%)

LET

(1) AP	389	(2.9%)
(2) AT	1028	(7.7%)
(3) B	10	(0.1%)
(4) C	791	(5.9%)
(5) D	311	(2.3%)
(6) EX	40	(0.3%)

(7) FO	0	(0.0%)	(15) NP	67	(4.0%)
(8) FU	10	(0.1%)	(16) P	215	(13.0%)
(9) FW	2	(0.0%)	(17) R	123	(7.4%)
(10) GE	40	(0.3%)	(18) TO	25	(1.5%)
(11) I	1235	(9.2%)	(19) UH	8	(0.5%)
(12) J	751	(5.6%)	(20) VB	92	(5.5%)
(13) M	111	(0.8%)	(21) VD	13	(0.8%)
(14) N	2090	(15.7%)	(22) VH	38	(2.3%)
(15) NP	406	(3.0%)	(23) VM	31	(1.9%)
(16) P	1809	(13.5%)	(24) VV	205	(12.4%)
(17) R	872	(6.5%)	(25) X	19	(1.1%)
(18) TO	243	(1.8%)	(26) Z	0	(0.0%)
(19) UH	41	(0.3%)			
(20) VB	620	(4.6%)	MRR		
(21) VD	118	(0.9%)	(1) AP	167	(3.1%)
(22) VH	281	(2.1%)	(2) AT	445	(8.3%)
(23) VM	234	(1.8%)	(3) B	0	(0.0%)
(24) VV	1756	(13.2%)	(4) C	385	(7.2%)
(25) X	163	(1.2%)	(5) D	119	(2.2%)
(26) Z	2	(0.0%)	(6) EX	17	(0.3%)
			(7) FO	0	(0.0%)
			(8) FU	0	(0.0%)
LIO			(9) FW	108	(2.0%)
(1) AP	248	(2.6%)	(10) GE	12	(0.2%)
(2) AT	673	(7.0%)	(11) I	465	(8.7%)
(3) B	0	(0.0%)	(12) J	336	(6.3%)
(4) C	652	(6.8%)	(13) M	55	(1.0%)
(5) D	231	(2.4%)	(14) N	869	(16.2%)
(6) EX	15	(0.2%)	(15) NP	107	(2.0%)
(7) FO	0	(0.0%)	(16) P	673	(12.5%)
(8) FU	0	(0.0%)	(17) R	325	(6.1%)
(9) FW	6	(0.1%)	(18) TO	80	(1.5%)
(10) GE	24	(0.2%)	(19) UH	8	(0.1%)
(11) I	807	(8.4%)	(20) VB	239	(4.5%)
(12) J	579	(6.0%)	(21) VD	42	(0.8%)
(13) M	83	(0.9%)	(22) VH	97	(1.8%)
(14) N	1335	(13.9%)	(23) VM	108	(2.0%)
(15) NP	288	(3.0%)	(24) VV	636	(11.9%)
(16) P	1366	(14.2%)	(25) X	74	(1.4%)
(17) R	726	(7.6%)	(26) Z	0	(0.0%)
(18) TO	181	(1.9%)			
(19) UH	36	(0.4%)	NEL		
(20) VB	475	(4.9%)	(1) AP	467	(2.5%)
(21) VD	82	(0.9%)	(2) AT	1294	(7.1%)
(22) VH	264	(2.7%)	(3) B	2	(0.0%)
(23) VM	180	(1.9%)	(4) C	1198	(6.5%)
(24) VV	1199	(12.5%)	(5) D	431	(2.4%)
(25) X	157	(1.6%)	(6) EX	48	(0.3%)
(26) Z	0	(0.0%)	(7) FO	0	(0.0%)
			(8) FU	0	(0.0%)
MBL			(9) FW	2	(0.0%)
(1) AP	26	(1.6%)	(10) GE	64	(0.3%)
(2) AT	134	(8.1%)	(11) I	1724	(9.4%)
(3) B	0	(0.0%)	(12) J	1201	(6.6%)
(4) C	112	(6.8%)	(13) M	124	(0.7%)
(5) D	35	(2.1%)	(14) N	2663	(14.5%)
(6) EX	9	(0.5%)	(15) NP	457	(2.5%)
(7) FO	0	(0.0%)	(16) P	2607	(14.2%)
(8) FU	0	(0.0%)	(17) R	1276	(7.0%)
(9) FW	0	(0.0%)	(18) TO	253	(1.4%)
(10) GE	5	(0.3%)	(19) UH	70	(0.4%)
(11) I	172	(10.4%)	(20) VB	873	(4.8%)
(12) J	86	(5.2%)	(21) VD	208	(1.1%)
(13) M	20	(1.2%)	(22) VH	350	(1.9%)
(14) N	224	(13.5%)			

(23) VM	322	(1.8%)	(2) AT	822	(9.1%)
(24) VV	2414	(13.2%)	(3) B	0	(0.0%)
(25) X	274	(1.5%)	(4) C	663	(7.4%)
(26) Z	0	(0.0%)	(5) D	225	(2.5%)
			(6) EX	29	(0.3%)
PAI			(7) FO	0	(0.0%)
(1) AP	107	(3.4%)	(8) FU	0	(0.0%)
(2) AT	294	(9.4%)	(9) FW	0	(0.0%)
(3) B	0	(0.0%)	(10) GE	15	(0.2%)
(4) C	204	(6.5%)	(11) I	855	(9.5%)
(5) D	87	(2.8%)	(12) J	629	(7.0%)
(6) EX	6	(0.2%)	(13) M	83	(0.9%)
(7) FO	0	(0.0%)	(14) N	1490	(16.6%)
(8) FU	0	(0.0%)	(15) NP	135	(1.5%)
(9) FW	6	(0.2%)	(16) P	1104	(12.3%)
(10) GE	20	(0.6%)	(17) R	689	(7.7%)
(11) I	302	(9.6%)	(18) TO	123	(1.4%)
(12) J	172	(5.5%)	(19) UH	13	(0.1%)
(13) M	20	(0.6%)	(20) VB	339	(3.8%)
(14) N	573	(18.2%)	(21) VD	37	(0.4%)
(15) NP	109	(3.5%)	(22) VH	185	(2.1%)
(16) P	344	(10.9%)	(23) VM	120	(1.3%)
(17) R	190	(6.0%)	(24) VV	1137	(12.6%)
(18) TO	43	(1.4%)	(25) X	75	(0.8%)
(19) UH	3	(0.1%)	(26) Z	0	(0.0%)
(20) VB	103	(3.3%)			
(21) VD	18	(0.6%)	SAN		
(22) VH	46	(1.5%)	(1) AP	218	(2.2%)
(23) VM	48	(1.5%)	(2) AT	670	(6.6%)
(24) VV	421	(13.4%)	(3) B	2	(0.0%)
(25) X	27	(0.9%)	(4) C	747	(7.4%)
(26) Z	0	(0.0%)	(5) D	264	(2.6%)
			(6) EX	37	(0.4%)
RAI			(7) FO	0	(0.0%)
(1) AP	392	(2.5%)	(8) FU	0	(0.0%)
(2) AT	1232	(7.9%)	(9) FW	0	(0.0%)
(3) B	0	(0.0%)	(10) GE	22	(0.2%)
(4) C	989	(6.3%)	(11) I	896	(8.9%)
(5) D	337	(2.1%)	(12) J	619	(6.1%)
(6) EX	44	(0.3%)	(13) M	111	(1.1%)
(7) FO	0	(0.0%)	(14) N	1404	(13.9%)
(8) FU	0	(0.0%)	(15) NP	309	(3.1%)
(9) FW	0	(0.0%)	(16) P	1381	(13.7%)
(10) GE	71	(0.5%)	(17) R	720	(7.1%)
(11) I	1426	(9.1%)	(18) TO	215	(2.1%)
(12) J	882	(5.6%)	(19) UH	31	(0.3%)
(13) M	114	(0.7%)	(20) VB	480	(4.8%)
(14) N	2483	(15.8%)	(21) VD	92	(0.9%)
(15) NP	435	(2.8%)	(22) VH	253	(2.5%)
(16) P	2017	(12.9%)	(23) VM	182	(1.8%)
(17) R	1103	(7.0%)	(24) VV	1294	(12.8%)
(18) TO	283	(1.8%)	(25) X	156	(1.5%)
(19) UH	36	(0.2%)	(26) Z	0	(0.0%)
(20) VB	682	(4.3%)			
(21) VD	169	(1.1%)	TRE		
(22) VH	264	(1.7%)	(1) AP	162	(2.6%)
(23) VM	278	(1.8%)	(2) AT	466	(7.4%)
(24) VV	2221	(14.2%)	(3) B	0	(0.0%)
(25) X	226	(1.4%)	(4) C	419	(6.6%)
(26) Z	1	(0.0%)	(5) D	170	(2.7%)
			(6) EX	17	(0.3%)
RED			(7) FO	0	(0.0%)
(1) AP	223	(2.5%)	(8) FU	0	(0.0%)
			(9) FW	0	(0.0%)

(11) I	542	(8.5%)
(12) J	372	(5.9%)
(13) M	51	(0.8%)
(14) N	925	(14.6%)
(15) NP	117	(1.8%)
(16) P	928	(14.6%)
(17) R	486	(7.7%)
(18) TO	122	(1.9%)
(19) UH	19	(0.3%)
(20) VB	304	(4.8%)
(21) VD	48	(0.8%)
(22) VH	150	(2.4%)
(23) VM	129	(2.0%)
(24) VV	825	(13.0%)
(25) X	71	(1.1%)
(26) Z	0	(0.0%)