A Stylo-Statistical Analysis of W. Somerset Maugham's Short Stories (4)

——*The Mixture as Before* (1940)——

WADA Koumei
YOSHIOKA Ken'ichi

1. Introduction

The Mixture as Before (hereafter referred to as *Mixture*) was published in 1940 as his next-to-last collection of short stories. As for the title of this volume, there is an ironic episode. Maugham tells about it in his preface:

> When my last volume of short stories was published *The Times* headed their review of it with the title *The Mixture as Before*. This of course was meant in a depreciatory sense, but I did not take it as such and so bold as to use it for the collection which I am now inviting the public to read.[1]

The phrase 'mixture as before' was, according to Stanley Archer, used at that time for reordering a supply of pipe tobacco, and was intended as an unfavorable tag.[2] The *Times* review of *Mixture* (8 June, 1940), on the other hand, was not bitter:

> As a description of the contents of *The Mixture as Before*, the title is not exact. The masterly method of concise and vivid statement is unchanged,...[3]

The purpose of our paper is to understand the features of the short stories by W. Somerset Maugham through a stylo-statistical analysis of *Mixture*. We will often refer to the results recorded in our previous papers on Maugham's collections such as *Orientations, Trembling, Casuarina, First Person,* and *Cosmopolitans*.[4]

The texts are extracted from W. Somerset Maugham, *The Complete Short Stories*, 3 vols. (London: Heinemann, 1967). Titles, the tokens, and the abbreviations of the ten stories are shown below in rank order with the story having the fewest words first:

The Three Fat Women of Antibes	(5,418 words, FAT)
The Treasure	(6,211 words, TRE)
The Lotus Eater	(6,498 words, LTU)
The Voice of Turtle	(6,602 words, VOI)
Gigoro and Gigolette	(6,942 words, GIG)
The Facts of Life	(7,728 words, FAC)
A Man with Conscience	(8,489 words, MWC)
An Official Position	(8,989 words, OFI)
The Lion's Skin	(9,332 words, LIO)
Lord Mountdrago	(9,691 words, LOR)

2. Quantitative Features of Vocabulary

(1) Word-length

Table 1 shows some characteristics of the tokens used in the ten stories. The mean of word-length is 4 letters±standard error (ranging 0.023 to 0.03) all through the ten stories, and the longest word in all the stories has 19 letters, with the median between 3.01 (OFI) and 3.14 (FAT).

Table 1 Word length

Text	Mean ±Standard Error	Max. Length	Median
FAT	4.37 ± 0.03	19	3.14
TRE	4.17 ± 0.029	18	3.04
LTU	4.01 ± 0.026	15	3.02
VOI	4.19 ± 0.026	19	3.03
GIG	4.09 ± 0.025	15	3.06
FAC	4.09 ± 0.024	16	3.03
MWC	4.21 ± 0.026	16	3.04
OFI	4.09 ± 0.023	18	3.01
LIO	4.13 ± 0.023	17	3.04
LOR	4.27 ± 0.024	16	3.05

As for the word-length, *Mixture* shows close approximation to *Orientations, Trembling, Casuarina,* and *Cosmopolitans,* which tells Maugham uses almost the same length of words in all his short stories.

The total number of the words ranging from one letter to four accounts for over 60% through all the texts. The longer words, on the other hand, consist mainly of hyphenated words or compound ones, such as 'conscience-stricken' (FAT), 'companion-secretary'

(VOI), and 'misunderstanding' (FAT).

The skewness and the coefficient of variation of the frequency distribution of the word-length are as follows:

Table 2 Skewness and the Coefficient of Variation

	FAT	TRE	LTU	VOI	GIG	FAC	MWC	OFI	LIO	LOR
skewness	.62	.52	.48	.51	.52	.52	.51	.50	.52	.53
coef. vari.(%)	50.3	52.3	52.6	55.4	50.8	51.1	55.9	52.9	52.8	56.3

As Table 2 shows, each story has a value of above fifty percent in skewness, which represents positively skew distributions. The properties of frequency distributions could be indicated by describing the shape of the frequency polygon which corresponds to them. In this case, 'tails' of all the polygons are to the right, towards the higher values on the x-axis.

As far as the word "I" is concerned, the ten stories have different dispersion: the frequency of "I" is low in OFI and FAT, and high in LTU, VOI, MWC, LIO, and LOR (Table 3). A chi-square test proves that the differences of the frequency of 'I' are statistically significant at the 0.1% level. The reason why the frequency of 'I' is high in those five stories is that three of them are written with the narrator 'I', and in the remaining two, LIO has a lot of dialogues and LOR has long dialogues in which Lord Mountdrago, the hero and a patient, tells his symptoms to the doctor, a psycho-analyst.

In *First person,* as the title shows, the percentage of the frequency of 'I' in relation to the token is considerably high (more than 2%) with the exception of one story (1.4).[5]

Table 3 The Frequency of "I" and its Percentage in Relation to the Token

	FAT	TRE	LTU	VOI	GIG	FAC	MWC	OFI	LIO	LOR	χ^2
"I"	48	89	208	177	124	70	221	10	182	257	377
%	0.9	1.4	3.2	2.7	1.8	0.9	2.6	0.1	2.0	2.7	

(2) **Type-Token Ratio (TTR)**

In our study of Maugham's vocabulary, we have been treating words in terms of the ratio of the type (the number of different words) to the token (the total number of words), that is, the relative frequency. Table 4 shows the type of the ten stories and their relative frequencies (TTR).

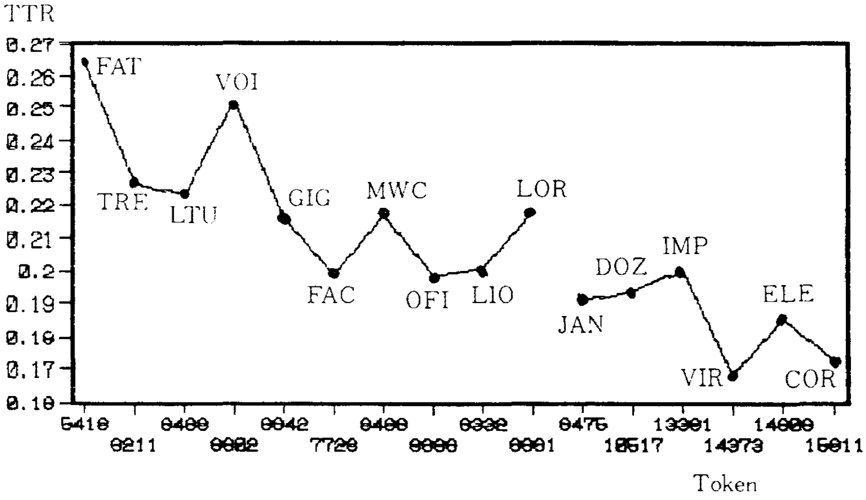
The type-token ratio (TTR) ranges from 0.198 (OFI) to 0.264 (FAT). For the comparison with *First Person,* TTRs are given in Figure 1.

Table 4 The Types and TTR

	FAT	TRE	LTU	VOI	GIG	FAC	MWC	OFI	LIO	LOR
Types	1,430	1,410	1,448	1,661	1,501	1,541	1,842	1,785	1,872	2,107
TTR	0.264	0.227	0.223	0.252	0.216	0.199	0.217	0.198	0.201	0.217

As a general tendency, the type increases with the increase of the token. But the rate of the type's increase is lower than that of the token, and a conspicuous fact is that TTR diminishes as the type increases.[6] In *Mixture*, however, TTRs are considerably high for their text sizes. For the understanding of this, we will compare the TTRs of *Mixture* with those of *First Person* (Figure 1).

Figure 1 The Relation of TTR to Token



Of all the stories, VOI and LOR have remarkably high TTR. Stories with high TTR in proportion to their tokens means that a fair variety of words are used in them. And this must be proved by another norm, 'the richness of vocabulary.'

(3) Richness of Vocabulary

In attempts to investigate a writer's style, several formulas for calculating a writer's richness of vocabulary have been proposed by M. E. Brunet (1980), M. D. Dugast (1980), P. Guiraud (1954), and C. Muller (1977).

Brunet's formula for calculating the richness index R(b) is:

$$R(b)=N^{(V-20)^{(-0.172)}}$$

(N is the tokens and V is the types.)

Guiraud's R(g) is:

$$R(g)=V/\sqrt{N}$$

Muller's R(m) is:

$$R(m)=(25-W)/1.5$$

(W is the richness index obtained from Brunet's formula.)

Dugast's R(d) is:

$$R(d)=\log^2 N/(\log N-\log V)$$

(He refers to his richness index as the Uber index.)

In accordance with these four formulas, we can get the index of the richness of vocabulary each story has. According to Dugast, the value of an index 18, 20, and 24 indicates respectively that a story is written with limited, average, and rich variety of words. Brunet, on the other hand, describes that the richness of vocabulary increases as the index number becomes lower. Guiraud and Muller do not give us precise standard by which we can evaluate the richness of vocabulary concretely.

The results obtained from the four formulas are shown in Table 5. We can see from R(d) that all the ten stories are written with a rich variety of words, while R(b) indicates that they fall into the range of rich to average variety of words.[7)]

There are of course definite differences of the index number coming from different ways of calculation among the four formulas, and also some disparity of the interpretation of the index number between R(b) and R(d). However, we can find some coincidence among the four when we compare the ranks each story occupies in Table 5. We can also find the stories taking the top highest ranks words are VOI, FAT, LOR, and MWC, and that these stories have a high TTR.

In conclusion, both TTR and the richness of vocabulary can be used as the scale of what degree of variety a writer uses words with. But the former tends to be under the

Table 5 Richness Index in Aascending Order of Tokens

Text	R(b)	Rank	R(g)	Rank	R(m)	Rank	R(d)	Rank
FAT	11.82	3	19.43	4	8.78	3	39.63	2
TRE	12.37	6	17.89	9	8.42	6	35.46	5
LTU	12.39	8	17.96	8	8.41	7	35.28	7
VOI	11.73	1	20.44	2	8.85	1	39.85	1
GIG	12.43	9	18.02	7	8.38	9	34.93	8
FAC	12.66	10	17.53	10	8.22	10	33.41	10
MWC	12.02	4	20.02	3	8.66	4	37.05	4
OFI	12.38	7	18.84	6	8.41	7	34.69	9
LIO	12.26	5	19.38	5	8.50	5	35.34	6
LOR	11.76	2	21.42	1	8.83	2	38.42	3

influence of the token, while the latter is not. As has been already noted, all the stories in *Mixture* have high TTR, of which the four stories tower above the others (Figure 1). The index of the richness of vocabulary proves that these stories are written with a fair variety of words.

(4) **Quantitative Characteristics of 'Type'**

We have calculated the mean frequency of the type (Table 6).

Table 6 The Mean Frequency of Type

	FAT	TRE	LTU	VOI	GIG	FAC	MWC	OFI	LIO	LOR
Mean Freq.	4	4	5	4	5	5	5	5	5	5

The mean of frequencies of the types is almost the same through all the stories, though subject to the influence of the type (between 4 and 5).

We have also investigated the frequencies 'the' and 'a' (Table 7). The percentages of 'the' and 'a' show that they are used with almost the same frequency of use through the texts.

The mean of frequency of type and the frequencies of 'the' and 'a' could be constant features of Maugham's writing, though they seem to be very humble features.

Tabel 7 The Frequencies of 'a' and 'the' and their Percentages in Relation to the Tokens

	FAT	TRE	LTU	VOI	GIG	FAC	MWC	OFI	LIO	LOR	χ^2
'a'	172	216	184	192	217	224	235	305	285	244	21.1
(%)	(3.2)	(3.5)	(2.8)	(3.0)	(3.1)	(2.9)	(2.8)	(3.4)	(3.1)	(2.5)	
'the'	208	216	316	263	337	318	346	483	363	391	57.3
(%)	(3.8)	(3.4)	(4.9)	(4.0)	(4.9)	(4.1)	(4.1)	(5.4)	(3.9)	(4.0)	

Paying close attention to the words with the highest frequency of use, Guiraud proposed a formula named 'concentration of vocabulary.' Allen thought of this as a writer's degree of "repetitiveness" (Rep).[8)]

$$\text{Rep}=\sum_{i=1}^{50} f_i/N$$

(N is the token and f_i is a frequency of the i-th word counted down from the word with the highest frequency.)

The result of calculation with this formula is given in Table 8. The Rep of FAT, which has the lowest token, is 0.48 and that of LOR, which has the highest token, is 0.50. We can assume that if the token is not over 10,000 words, the Rep remain almost fifty.

Table 8 Repetitiveness

FAT	TRE	LTU	VOI	GIG	FAC	MWC	OFI	LIO	LOR
.48	.51	.50	.50	.50	.50	.50	.54	.49	.50

Yule (1944), on the other hand, paid a close attention to the other side of the word frequency distributions, and proposed the Characteristic K for 'concentration of vocabulary.' If we use the 1st and 2nd moment of a variable about zero, the Characteristic K can be written as follows:

$K = 10^4 \times (S_2 - S_1)/S_1^2$

(S_1 is $\Sigma F_x \cdot X$, S_2 is $\Sigma F_x \cdot X^2$.)

Yule's value is calculated on the assumption of The Poisson Law, while another index V_m is derived by Herdan (1956) with no assumption whatever about a stochastic process.

$V_m^2 = \{\Sigma F_x \cdot X^2 / (\Sigma F_x \cdot X)^2\} - 1/N$

(N is ΣF_x.)

The V_m can be obtained from dividing the coefficient of variation s/m by $\sqrt{N}$ (s is the standard deviation and m is $\Sigma F_x \cdot X/N$), and shows how the frequency that each type has deviates from the mean of frequency that all the types have. Table 9 shows Characteristic K, V_m, and entropy of the ten stories.[9]

Table 9 Characteristick K, Vm, and Entropy

	K	Vm	Entropy	Frequency of word used once (%)
FAT	89	0.092	8.63	869 (60.8)
TRE	96	0.095	8.45	826 (58.6)
LTU	100	0.097	8.47	871 (60.2)
VOI	94	0.095	8.63	1,028 (61.9)
GIG	89	0.092	8.65	838 (55.8)
FAC	102	0.098	8.49	852 (55.3)
MWC	92	0.094	8.66	1,110 (60.2)
OFI	132	0.113	8.36	1,005 (56.3)
LIO	88	0.092	8.65	1,151 (61.5)
LOR	96	0.096	8.72	1,287 (61.0)

We cannot see any characteristic tendency in Table 9. This may be due to the fact that all the tokens are less than 10,000. As for the stories with the token not going over 10.000, however, it may be a feature of Maugham's writing that around 60 % of all the words in each of the ten stories cansists of the words used once.

As far as the works whose token is over 10,000 are concerned, Characteristic K and Entropy are generally in reverse proportion. We can see that when the Chacteristic K goes over 100, the value of entropy tends to become lower even if tokens are within 10,000.

(5) Vocabulary Growth

An occurrence distribution of the tokens and types of the ten stories is illustrated in Figure 2. The first column of the y-axis is the number of tokens counted from the start of a text. The x-axis is the number of the types. As the last segment of each stories tends to be less than 500 words, the score is compensated as per 500-word segment. In the first 4 segments in all the stories, the score of "type" decreases sharply and then its curved line declines slowly. The marks of A, B and C in Figure 2 stands for the starting points of the beginning, the middle, and the end part of a story.

In Maugham's short stories such as *Cosmopolitans*, whose 29 stories are all written within 3,000 words, we can rather easily point out, in the graphical representation, those starting points. However, it is not an easy job to spot the turning points in *Mixture* in which all the stories go beyond 5,000 words. As a result, the shape of the curve needs to be collated with the development of a story.

Our collation reveals that there are three types in the development of a story, two of them having subclasses;

(1) The shape of the curve shows a continuous descent; the beginning part plays an introductory part, the middle is the main part of the story, and the end puts a conclusion to the story—FAT, MWC, and OFI.

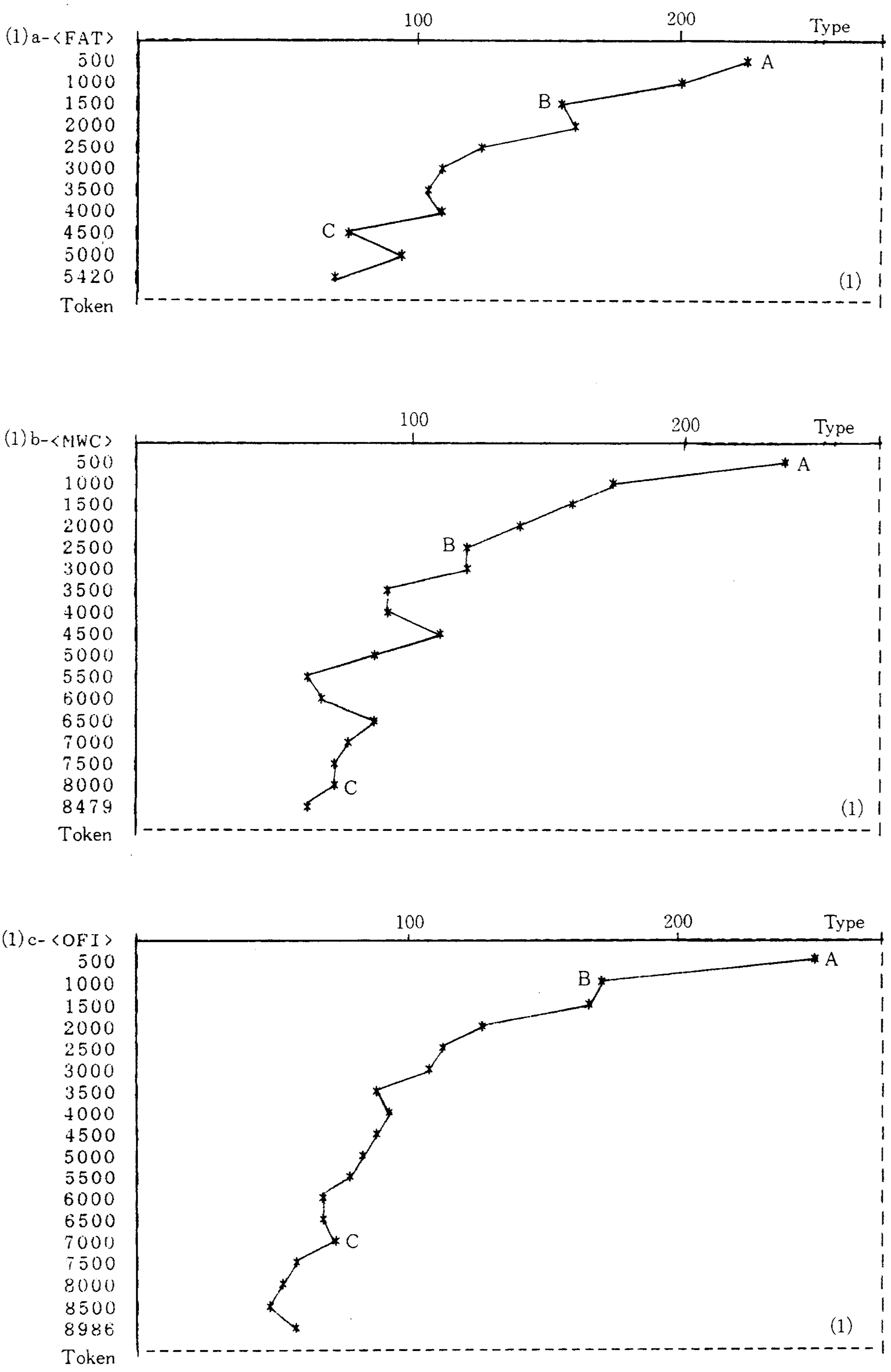
(2) The shape of the curve holds more or less evenness in the middle; (A) The beginning is rather short, while the middle is long and occupies the bulk of the story, and the middle can be further subdivided into a beginning(B—1), a middle(B—2), and an end(B—3)—TRE, LIO, and FAC. (B) The end is missing, with the middle subdivided into a beginning, a middle, and an end—GIG and VOI.

(3) The shape of the curve tends to rise in the end; (A) the story ends with an unexpected conclusion—LTU. (B) The beginning, via the middle, reverts to the end —LOR.

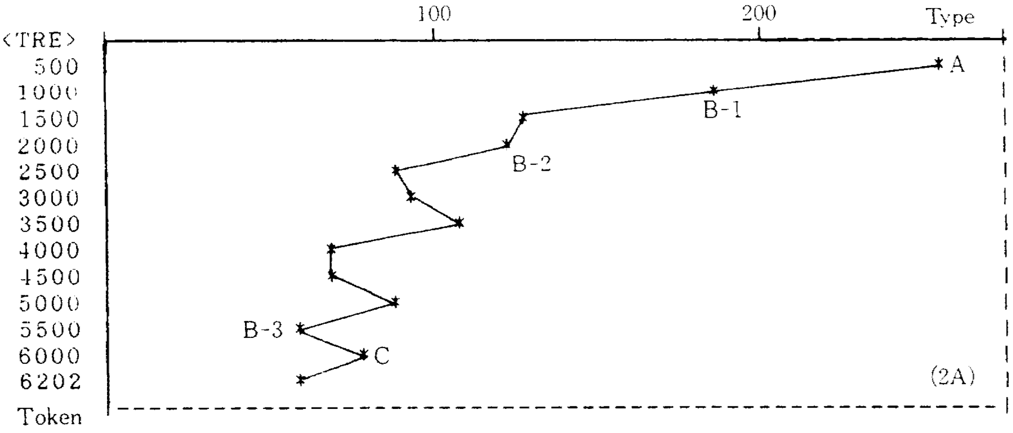
(6) Dendrogram

The number of words that occur in any of the ten stories is 224. Almost all of these words consist of function words. Based upon the number of these words, the Euclidian distances can be obtained. The distance from one story to another in the dendrogram is generally considered to denote the degree of resemblance between any two stories taken out of the ten stories. From the distance we can conclude that there is a close re-

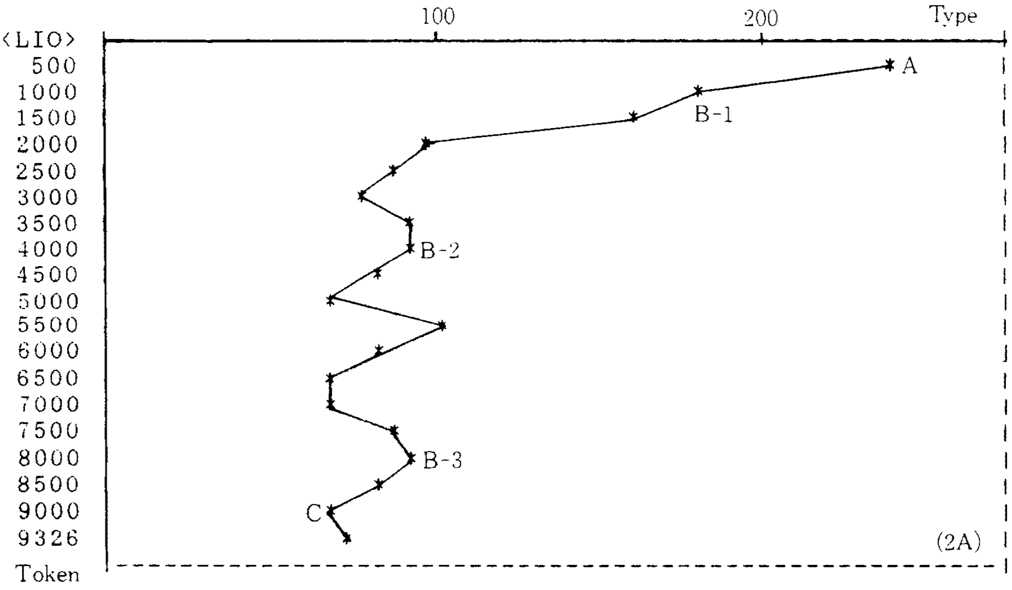
Figure 2 The Vocabulary Growth



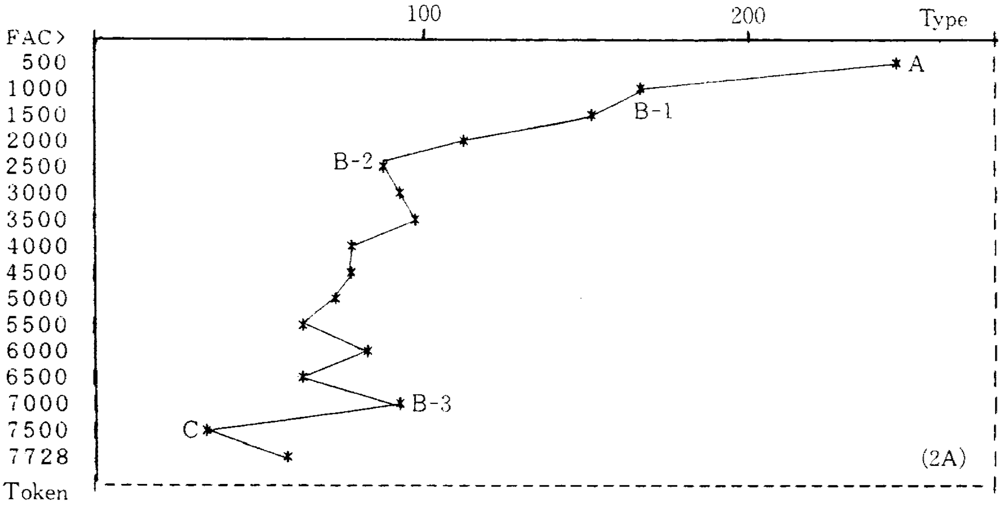
(2A)a- <TRE>

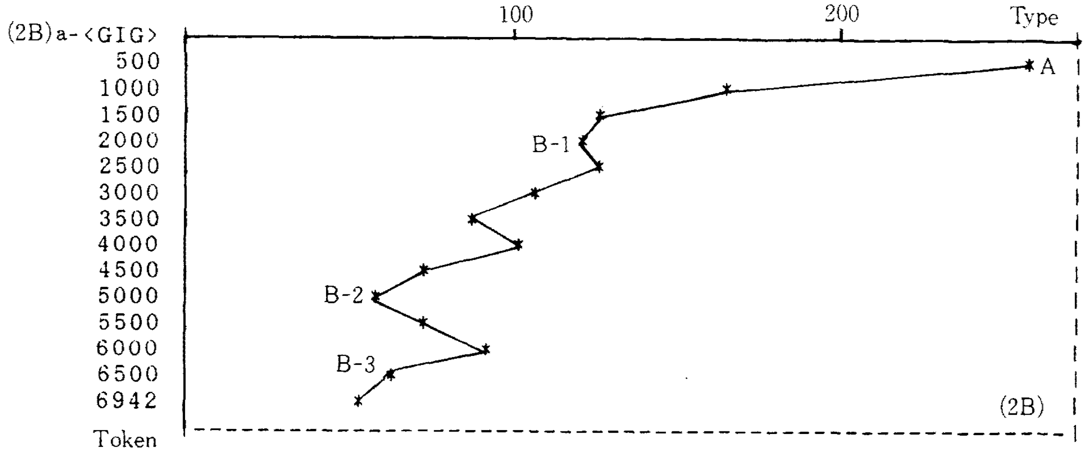

100
200
Type
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
6202
Token
A
B-1
B-2
B-3
C
(2A)

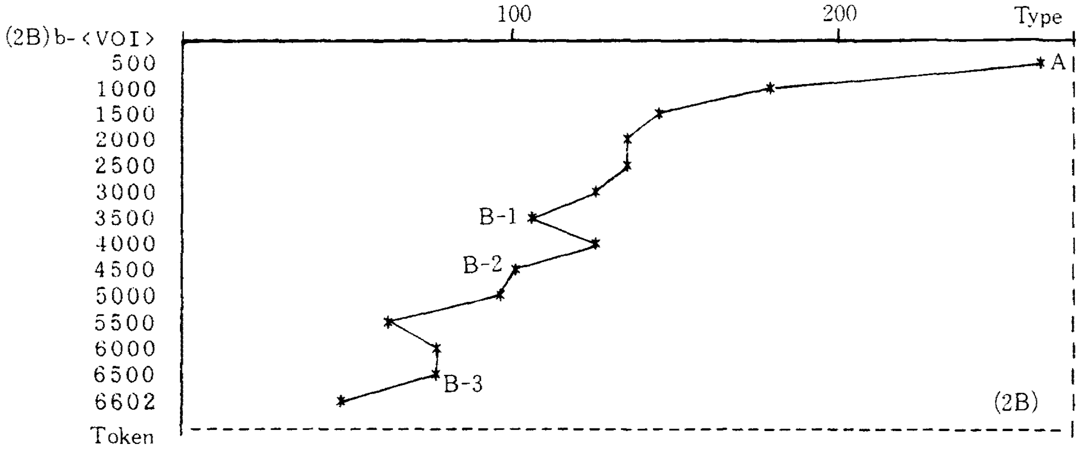
(2A)b- <LIO>

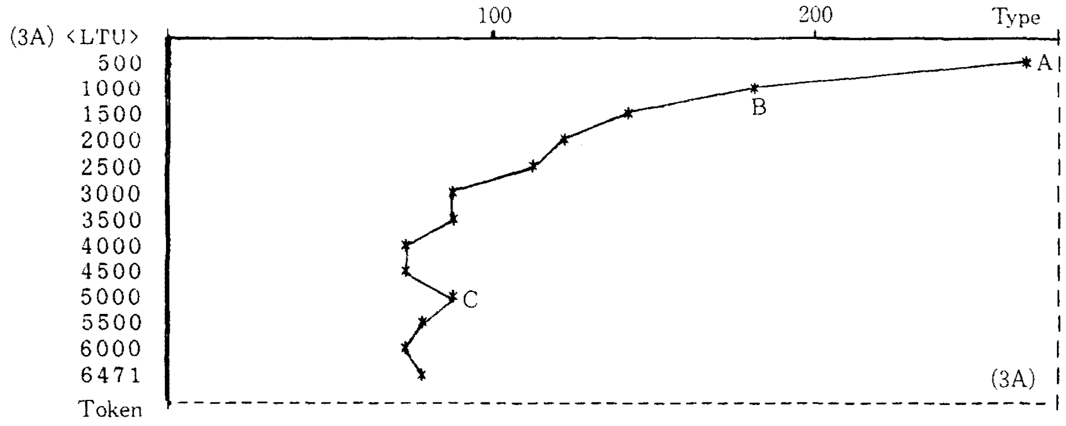

100
200
Type
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
6500
7000
7500
8000
8500
9000
9326
Token
A
B-1
B-2
B-3
C
(2A)

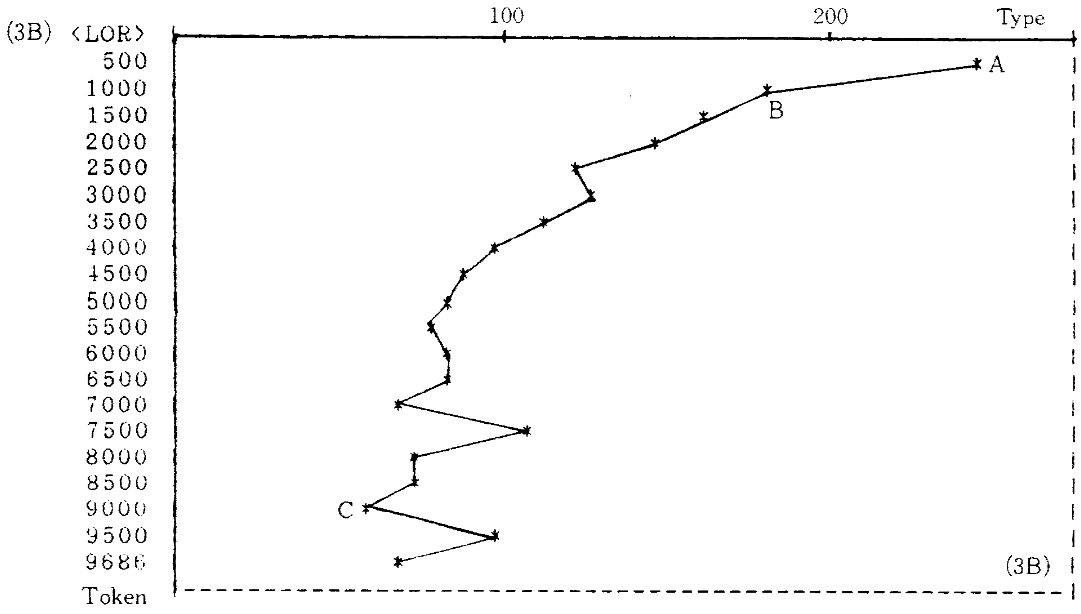
(2A)c- <FAC>


100
200
Type
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
6500
7000
7500
7728
Token
A
B-1
B-2
B-3
C
(2A)


(2B)a-<GIG>
100
200
Type
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
6500
6942
Token
A
B-1
B-2
B-3
(2B)

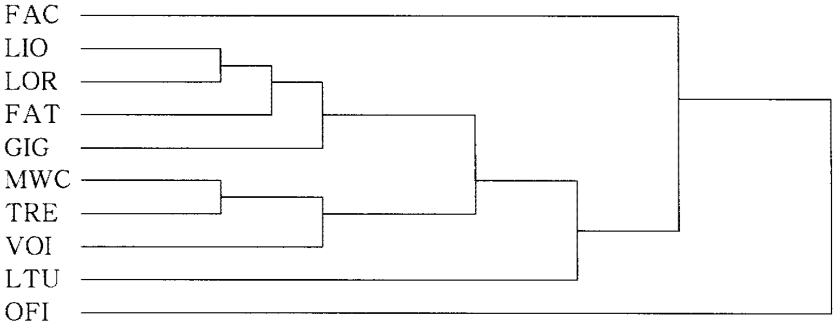

(2B)b-<VOI>
100
200
Type
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
6500
6602
Token
A
B-1
B-2
B-3
(2B)


(3A) <LTU>
100
200
Type
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
5500
6000
6471
Token
A
B
C
(3A)



semblance between LIO and LOR. Moreover, TRE and MWC are in pretty close relation. These distances are illustrated in the form of the dendrogram in Figure 3.

Figure 3 Dendrogram Using Euclidian Distance Data



The dendrogram (Figure 4) is based on the correlation coefficient (Table 11). Judging from Figure 4, the pair of FAC and OFI, and that of MWC and LOR have the highest degree of resemblance in terms of the occurrence of the same words between any two stories taken out of the ten.

Figure 4 Dendrogram Using Correlation Coefficient

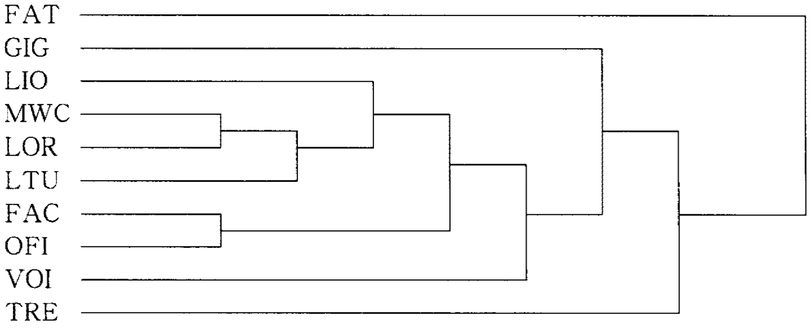


Table 11 Correlation Coefficient

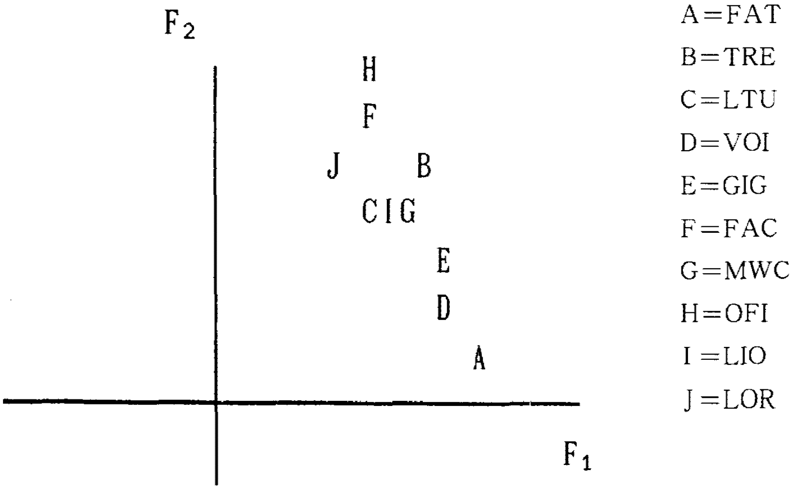
	FAT	TRE	LTU	VOI	GIG	FAC	MWC	OFI	LIO
TRE	0.884								
LTU	0.806	0.895							
VOI	0.914	0.936	0.928						
GIG	0.919	0.923	0.938	0.942					
FAC	0.782	0.949	0.918	0.884	0.903				
MWC	0.846	0.930	0.969	0.954	0.945	0.933			
OFI	0.771	0.925	0.898	0.850	0.884	0.975	0.915		
LIO	0.853	0.949	0.951	0.943	0.954	0.956	0.967	0.923	
LOR	0.765	0.909	0.959	0.909	0.908	0.951	0.974	0.926	0.965

To seek the resemblance among the ten stories from another norm, we have analysed the values given in Table 11 via a factor analysis. Communalities and factor loadings are given in Table 12, and the resemblance are illustrated with communalities based F_1 and F_2 on the two-dimensional plane (Figure 5).

Table 12 Communalities and Factor Loadings

	COM	F_1	F_2	F_3	F_4	F_5
FAT	0.984	0.920	0.195	-0.315	0.013	-0.002
TRE	0.975	0.712	0.528	-0.406	-0.072	0.136
LTU	0.967	0.550	0.431	-0.690	0.044	-0.018
VOI	0.978	0.735	0.306	-0.572	-0.102	0.076
GIG	0.986	0.736	0.361	-0.532	0.164	0.007
FAC	0.987	0.545	0.676	-0.473	0.021	0.092
MWC	0.985	0.601	0.440	-0.655	-0.037	0.004
OFI	0.994	0.542	0.723	-0.415	0.018	-0.068
LIO	0.983	0.623	0.487	-0.581	0.056	0.133
LOR	0.981	0.485	0.536	-0.674	-0.006	0.059

Figure 5 Resemblance based communalities



From the factor analysis we cannot attach the significance to the x-axis nor to the y-axis. However, we can see that A (FAT), H (OFI), and J (LOR) are rather far from each other in the Figure. They are indeed quite different from each other in the light of their contents of stories. We can also see that C (LTU), I (LIO), and G (MWC) may be classified into the same group.

3. Central Tendency and Variability of Sentence-length

We have tried to scrutinize the features of Maugham's style from the standpoint of sentence length.

The central tendency viz. the mode, the median and the mean are shown in Table 13.

Table 13 Central Tendency and Variability

	Total of sentences	Mean length in words(SD)	Coeff. varia.	Mode	Median	Frequency. of one-word sentence
FAT	426	12.7 (9.9)	0.78	6	9.1	2
TRE	470	13.2 (10.5)	0.49	6	9.0	2
LTU	430	15.1 (11.6)	0.77	6	11.5	10
VOI	444	14.9 (12.0)	0.81	6	10.6	3
GIG	608	11.4 (9.2)	0.80	6	8.1	3
FAC	536	14.4 (11.8)	0.82	6	10.2	5
MWC	503	16.9 (12.2)	0.73	5	13.6	4
OFI	525	17.1 (13.8)	0.80	5	12.9	5
LIO	600	15.6 (12.9)	0.83	6	10.7	6
LOR	624	15.5 (12.8)	0.82	4	11.5	7

The mean sentence length in each story of *Mixture* ranges from 11 to 17 words and the mean of the total stories is 14.6 words, while that of *First Person* in our previous paper is 13.1 words.

Compared with Maugham's previous collections of short stories, the mean of sentence length in *Mixture* will be not shorter nor longer.

Though OFI and MWC fall into the longest group, 'and' and commas are used with high frequency in them. (Table 14).

All the stories contain some one-word sentences, but they are almost dialogues.

From the probability obtained by a χ^2, the differences of frequencies are statistically significant at the 0.01 level.

Table 14 The Frequency of "and", "but", comma and semicolon

	FAT	TRE	LTU	VOI	GIG	FAC	MWC	OFI	LIO	LOR	χ^2
and	190	175	213	192	229	214	247	280	255	273	15.03
but	32	38	66	65	40	62	56	56	78	65	22.42
comma	260	271	348	355	434	366	377	385	521	412	66.82
semicolon	13	5	19	20	13	60	58	46	55	60	79.55

The longest sentence in the whole stories is found in OFI (95 words):

'And yet never was there a more good-natured man than Louis Remire; she blamed him for the money he spent at the cafe, she accused him of wasting it on other women; well, in his position he had opportunities now and then, and as any man would he took them, and he was easy with his money, he never minded paying a round of drinks for his friends, and when a girl who had been nice to him wanted a new hat or a pair of silk stockings he wasn't the man to say no.'

4. Conclusion

Mixture is said to be one of Maugham's best and happiest collection. His craftsmanship and his style were now perfected.[10] We have investigated some characteristic features of his style from the standpoint of stylometrics:

1. Word-length: Maugham uses almost the same length of words (4 letter word) in *Mixture* as he did in his previous collections. He often uses long words, but they consist mainly of hyphenated words or compound ones.
2. TTR and the Richness of Vocabulary: All the stories are written with a fair variety of words. And some (VOI, FAT, LOR, and MWC) stand out conspicuously high from the viewpoint of the richness of vocabulary as well as from that of TTR.
3. The mean of frequencies of type: between 4 and 5.
4. The frequencies of 'the' and 'a': The articles 'the' and 'a' are used with almost the same percentage of frequency through the ten stories.
5. Repetitiveness: no less than 0.48 nor more than 0.50.
6. Vocabulary growth: Some stories do not observe strictly the form of short stories Maugham says he favours; that is, a plot should have a beginning, a middle, and an end. In those stories, the end part of a story is missing. Instead the middle part is long and can be subdivided into a beginning, a middle, and an end.
7. Resemblance: Since all the stories are written by the same writer and their contents are not substantially different from each other, we cannot illustrate remarkably the

resembrance and the difference among the stories on the two-dimensional plane. Some of them, however, are obviously similar and some are placed apart, which could be proved by their contents.

8 . Sentence length: Maugham writes sentences with the mean length between 11 and 17 words as he did in his previous collections of short stories. We can see that he makes the short sentences longer with the use of 'and' and commas.

Notes

1) W. Somerset Maugham, *The Mixture as Before* (London: William Heinemann, 1940), Forward vii.

Incidentally, 'my last volume of short stories' refers to *Cosmopolitans* (1936) and the London *Times* review of it (31 March, 1936) is as follows:

"Fiction: Mr. Maugham's Mixture as Before"

The stories of *Cosmopolitans* "are all very slight and may not add greatly to the author's reputation: but if read 'one or two now and then,' as directed, they will not noticeably diminish it,"

Charles Sanders (ed), *W. Somerset Maugham—An Annotated Bibliography of Writings about him* (Illinois: Northern Illinois U.P.), p. 198.

2) Stanley Archer, *W. Somerset Maugham—A Study of the Short Fiction* (N.Y.: Twayne Publishers, 1993), p. 62.

3) Charles Sanders, p. 229.

4) See our papers, *Journal of Tezukayama College*, Nos. 25 (1988)~30 (1993).

5) See our papers, *Journal of Tezukayama College*, No. 30, March 1993.

6) Yukio Takefuta, 'Konpyu-ta no mita gendai-eigo' (*Modern English Analysed by Computer*) (Tokyo: Edyuka Shuppan, 1981), p. 180

7) As for the richness of vocabulary, we are wholly depend on Robert F. Allen, *A Stylo Statistical Study of "Adolfe"*, (Geneve-Paris: Slatkine-Champion, 1984), p. 158.

8) Allen, p. 164.

9) G. U. Yule, *The Statistical Study of Literary Vocabulary*, (Cambridge: Cambridge U. P., 1968), pp. 86-87.

10) Laurence Brander, *Somerset Maugham—A Guide* (London: Oliver & Byed, 1963), pp. 124-5.