On the Length of Writings in EFL:
Measuring Combining Tasks by T-unit

Keiko SATO

INTRODUCTION

According to John C. Mellon's claim, the two primary purposes of grammar are to teach "the principles of correctness, or error avoidance, in language use" and to present "what was presumed to be the full range of sentence structure available to one in the production of mature language."

Though these purposes of grammar are quite practical and thus face toward its utility, little research has been conducted objectevely on data basis. Most linguists of language learning have asserted from their empirical point of view that there is no high correlation of grammar knowledge with English writings. They do not count what natives have acquired as competence, but they just call grammar what students learn at school as part of the school curriculum.

It is evident that grammar is not independent of language use. It is not until grammar is used for interpreting sentences or producing new sentences that its purposes have been accomplished. In other words the interpretation and construction of sentences are never carried out without grammar knowledge as well as lexical knowledge and others. That is why it takes even native children more than ten years to develop full linguistic capacity. Much more for foreign language learners who start learning when they may have lost their biological endowment of our species, essential is the grammar of the language they are learning.

Along with the recent development of psycholinguistics, the assumption that children do not use incorrect grammar but experiment with various patterns in order to find the most appropriate through experience is leading these days. Foreign language learners can also be seen using the same procedure. The same learner can sometimes use it in correct form and cannot in other cases. At that stage he or she has not masterd its usage yet, just struggling toward the goal. He or she may be able to master it in the near future, through many trials. That is the same process as children naturally take when they are on the way to acquiring their own language. Selinker calls it the 'interlanguage.'

What is the concrete interlanguage of Japanese students? 'Interlanguage' is the everlasting process toward the goal of a target language. We may be able to find the characteristics at every level by cutting the process at several points. The performance ought to reflect the competence of writers.

To advance the study, the following position was taken up: that the learners must be

developing in language use as they grow older, and that they must be developing the length of sentences as they are developing the ability of combining each linguistic element for sentence expansion. For some expected counter remarks, some other research will be conducted, but our main interest is in the points above.

This study owes the method of procedure to Kellogg W. Hunt's paper (1965). He provided his own new procedure to determine mean length which shows both the length of writings and the complexity to make them longer. Before his study, mean sentence length, mean clause length and subordination had been taken for granted as measures of what L1 speakers wrote. He had some doubt as to these being good measures of sentence constructing ability. Immature children might make run-on sentences or too many and's. Such sentences should not be given high evaluation. Hunt first thought that "a sentence might be taken as whatever a student wrote between a capital letter and an end punctuation." But he was informed by his reseachers that such a definition was uncomfortable and found sentence length not to be a good measure. He thought of the combining index which moderated subordination index and clause index. He called it a 'minimal terminable unit,' T-unit for short, which is any main clause with some or no embedded clauses within it. His actual testing out of this index proved to be promising. He examined in great detail the grammatical items to lengthen T-units.

For the comparison with the results of Hunt (1965), the three mean lengths of sentence, T-unit and clause were adopted to measure what the subjects wrote. The purpose of this study is (1) to reveal how long sentence structures Japanese EFL learners produce actually, (2) to show the differences in length of what native speakers wrote and what Japanese EFL learners wrote, and (3) to investigate the causes to make the distinct difference between L1 and L2.

METHOD

Subjects

The ninth grade (the 3rd year of junior high school) seemed to be a good place to begin. They have been taught almost all the elemental grammatical items except relatives. Below the ninth grade, Japanese students can only manage to make the simplest sentences. The next grades were to be taken up by every three years: the twelfth grade (the 3rd year of senior high school) and the junior of college. But in actuality the sophomores were chosen as a good place to stop because of an appropriate number of students in a class to cooperate this study.

	grades	number of subjects	institutions
Data Source	Ninth grade	30	Nagoya YMCA advanced class
	Twelfth grade	30	Chikusa High School
	Sophomore	30	Nanzan University

The collected data has a problem. Though the informants are selected from the view of ability to produce English sentences, these three grade levels cannot be said to be of

the same level of intelligence. Hunt (1965) selected the students from each grade with their IQ between 90 and 110 as subjects. The unknown intelligence of our subjects is our most weak point.

Material for Analysis

To measure Japanese students' ability to combine the sentence elements and expand them into longer sentences at each level, the set of kernel sentences of a story was given to combine and the subjects were asked to reproduce the new forms of the original story. Adding words needed and omitting words not needed were allowed, but leaving out any information in the plot and changing the story were prohibited.

The most desirable meterial for combining is something that is comprehended perfectly, making it easier to undertake the task. Moreover, more desirable for our study is the same material as Hunt employed so that comparisons of results can be made. Therefore, a fairy-tale was used: a chicken's story. The meanings of some words were given in Japanese to ensure understanding for junior high school students. This was done to minimize errors due to misunderstanding out of unkown words. The material we used is presented in the appendix A at the end of this paper.

Method of Analysis

This quantitative study is obliged to begin with counting words in writings. For sentence length, all the words of the subjects' products and all the sentences were counted and the former was devided by the latter. Every word should have been counted, but only some strings that could hardly be understood were excluded. In Hunt (1965) also, such extraneous matters called 'garbles' were not counted. The synopsis of formulas for calculating the three lengths is as follows:

Sentence length = the number of words of the whole passage divided by sentence number

T-unit length   = the number of words of the whole passage divided by T-unit number

Clause length   = the number of word of the whole passage divided by clause number

RESULTS

Sentence Length

Sentence length will be in general a natural place with which we begin a quantitative study of writing development. We have no intention of going into the definition of sentence, but what is to be taken as a sentence must be determined beforehand for convenience. A few students under-punctuated where they were supposed to punctuate. Others sometimes forgot to begin their sentences with capital letters. The trouble with no punctuation, no capitalization and overuse of and left the decision of sentence to our judgment. However, most of the writings were well

punctuated. Unusually long were four sentences consisting of fifty-three, forty-six, thirty-one and forty words made by a twelfth-grade student. He combined all the fifty kernel sentences to form the four sentences. His average sentence length was 42.5 words. We excluded his product from data to be examined, for fear of its disturbance of statistic figures.

The averages of sentence length were 8.53 for grade 9, 11.65 for grade 12, and 10.43 for the sophomore. The distribution of each subject's mean sentence length is shown in Figure 1.

Figure 1　Distribution of Mean Sentence Length

T-unit Length

Definition of T-unit according to Hunt: one main clause with one or more than one embedded clause or without any embedded clause, namely a coordinate sentence is divided into simple sentences but a subordinate sentence remain a unit.

T-unit length may perhaps be a more satisfactory index to measure language maturity of young average students than sentence length. Most Japanese students, who begin to learn English at the age of twelve, have realized that two or more

Figure 2　Distribution of Mean T-unit Length

clauses in one sentence should be connected with some conjunctions. Only a few students have no firm mental attitude about that rule. The difference between mean sentence length and mean T-unit length ought to show the necessity of the creation of T-unit. Mean T-unit length of each grade were 7.35 for grade 9, 10.22 for grade 12, and 9.41 for the sophomore. The developmental order in T-unit length is the same as that in sentence length. T-unit length, however, shows the differences more clearly. The differnce between the ninth grade and the twelfth grade became larger, and that between the twelfth grade and the sophomore bacame smaller. We found that the ninth-graders lengthened sentences mainly by using coordinators between clauses. (See Figure 2 for the distribution of mean T-unit length of each grade.)

Clause Length

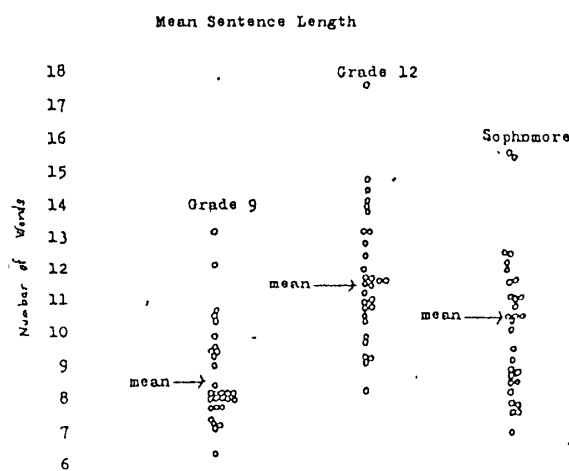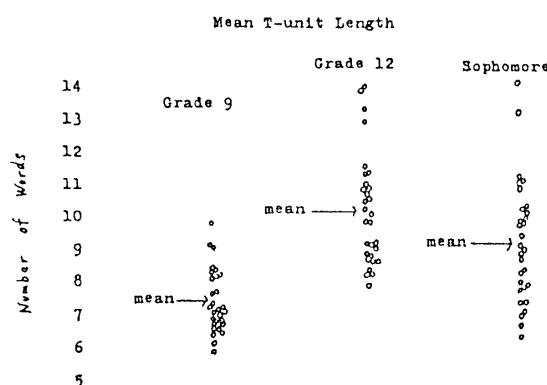In counting the number of clauses, a clause was taken to be a structure with a

subject and a finite verb at least. Coordinated subjects and coordinated verbs were regarded as what merely lengthened the clause. Clause length shows the simple sentence length without depending on any subordinators.

To get the mean clause length, a student's total number of words was divided by his or her total number of clauses. In this way the mean clause length was found for each student and each grade; 6.67 words for the ninth grade, 7.53 for the twelfth and 7.48 for the sophomore. There was a clear differ-
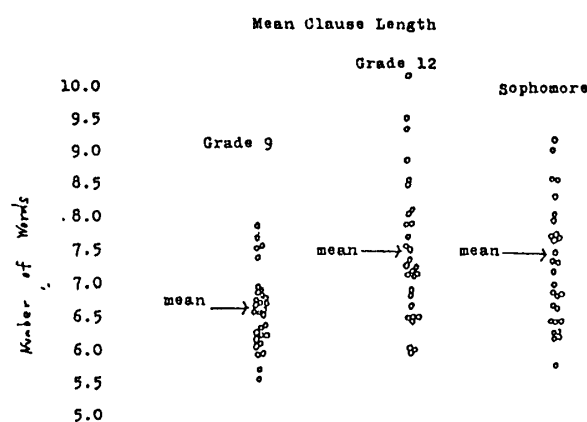

Figure 3 Distribution of Mean Clause Length

ence between the ninth and the twelfth, but little or no significant difference between the twelfth and the sophomore. If we count as one the fractions of more than .5 inclusive and cut away the rest, mean clause length for both grades will be 7.5 words. From the twelfth toward the sophomore the students can be said to have had no growth in lengthening their clauses.

Comparison of Three Lengths

Hunt (1965) investigated the three mean lengths of nine articles from the Harper's and nine from the Atlantic. He called those contributors "super adults" and decided their product as a target that undeveloped language learners should aim at. Table 1 presents the results of the mean lengths of three grade levels and the super adults Hunt studied with our reaults of Japanese three grade levels.

Table 1 Development of Three Mean Lengths

	J:JH3	J:SH3	J:C2	A:G4	A:G8	A:G12	Super Adult
Mean Sentence Length	8.71 (35%)	11.66 (47%)	10.44 (42%)	13.5 (55%)	15.9 (64%)	16.9 (68%)	24.7 words (100%)
Mean T-unit Length	7.35 (36%)	10.22 (50%)	9.41 (46%)	8.6 (42%)	11.5 (57%)	14.4 (71%)	20.3 words (100%)
Mean Clause Length	6.67 (58%)	7.53 (65%)	7 48 (65%)	6.6 (57%)	8.1 (70%)	8.6 (75%)	11.5 words (100%)

Note· The groups classified were as follows: J-JH3 for Japanese Junior High 3rd grade, J-SH3 for Japanese Senior High 3rd grade, J-C2 for Japanese College 2nd year, A-G4, A-G8 and A-G12 for American students of the 4th, 8th and 12th grades, and SuperAdults mentioned above.

With mean T-unit, Japanese students of a senior high school and a college reached about half of that of super adults (50%,46% respectively). They exceeded only

American fourth graders in T-unit length probably due to their greater clause length, but came up to none of the native groups in sentence length. The reason that the fourth graders of native speakers wrote much longer sentences than any of Japanese students is considered that American young students overused the coordination between clauses. That was why Hunt create the new index "T-unit." Here also, an evidence is found that T-unit is the better index in lower levels than sentence length.

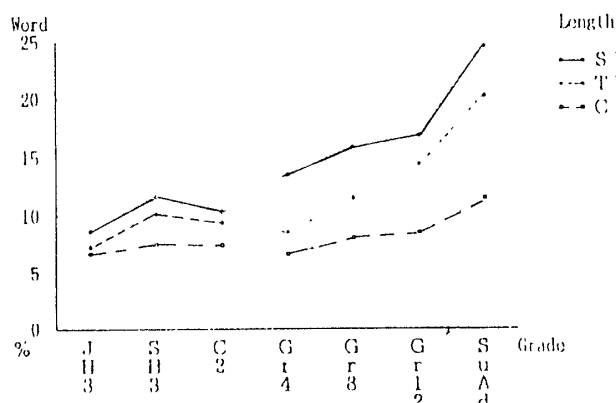The development of each group in the three lengths is presented in Figure 4.

Figure 4　Development of Three Mean Lengths

There is no outstanding growth of clause length except with super aduls. We can find the developmental process of sentence and T-unit lengths except between Japanese Senior High 3rd graders and College students. Especially from the 4th graders to super adults there found a very smooth growing line. What constitutes T-unit? What is the difference of T-unit inner structure between writings by native speakers of English and by Japanese learners of English?

Number of Clauses within a T-unit

The more use of coordination and subordination creates longer T-unit and longer sentence. How many clauses did our Japanese subjects put into one T-unit? Table 2 and Figure 5 show the results achieved.

Table 2　Percentage of Clauses Used within T-unit

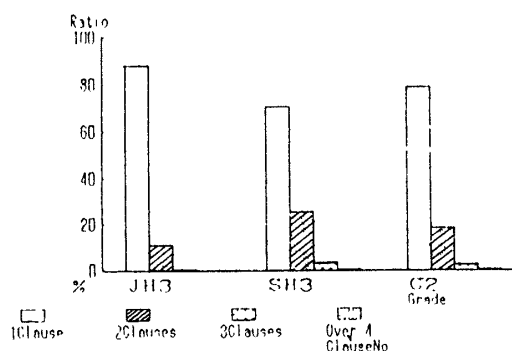|	JH3	SH3	C2
1 clause	88 1	70 5	78.8
2 clauses	11 1	25 8	18 5
3 clauses	.8	3 7	2 7
4 or more	none	1 case	1 case

Figure 5　Number of Clauses Used within T-unit

We find here how large clusters of clauses each grader prefers to write. The preference of single-clause T-units, 2-clause T-units, 3-clause ones, or more-than-3-cluase ones is given in percentage for each grade in Table 2. An overwhelming

majority of T-units written by Japanese learners was one clause (70% to nearly 90%). But as the students grow older, they come to use 2-clause T-units more often until they finish their high school education. One out of four or five persons uses 2-clause T-units in senior high and college levels. A few of the older students can even make 3- or 4-clause T-units, but not so often. It seems quite difficult for Japanese students to construct such complicated units.

Depth of clause subordination might be a good index of maturity for advanced learners. It cannot be, however, made use of by Japanese students, who rarely write such elaborate sentences. It is true that multi-clause T-units contribute to the lengthening of sentences, but the number of instances is too small to indicate the developmental tendency statistically.

Distribution of Short and Long T-units

Everyone will probably suppose that younger students write shorter T-units. But who knows how short their sentences are? What kind of short sentences do they use most so that they make their sentences shorter than older students do?

Getting the frequencies of T-units of all lengths was attempted before dividing them into three types of length; short, middle-length, and long T-units. Appendix B gives a full picture of how many T-units from the shortest to the longest were written by the students in each grade. Following the naming by Hunt, we call 1-8 word units "short", 9-20 "middle-length", and more than 20 words "long", for the later comparison with the native students' products. Table 3 presents the percentage of the use of three categorised lengths for each level, and Figure 6 presents the number of students using each length (from 6 to 14 words long) for each level.

Table 3 Percentage of Use
Three Kinds of T-unit Length

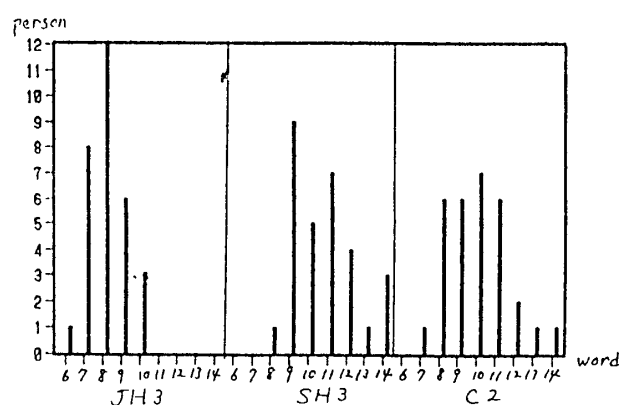|	JH3	SH3	C2
Short T-unit (1−8words)	68.0	45.0	51.5
Middle T-unit (9−20words)	31.5	51.0	46.0
Long T-unit (over 20 words)	.5	4.0	2.5

Figure 6   Number of Users
Various Lengths of T-unit

There is no great difference between senior high and college students in all three kinds of T-unit length. The junior high students prefer short units to middle-length ones (68% to 31.5% out of all their own T-units), but in the occurrences of middle-length units there is little difference from those by older students (248, 308, 291

respectively. See Appendix B for details). Middle-length units are apparently easy for Japanese students to write. Or rather we might as well call more-than-16-word T-units "long", judging from the number of occurrences by the Japanese students. On the contrary, the occurrences of short units by the ninth graders were twice as many as those by the senior high school students. The younger students wrote about three times as many 5-word units as the older. Their use of a great many short units increases the total number of all their units, resulting in their small mean T-unit length. As for the senior high and the college students, we can find only a little difference as mentioned earlier. They wrote as many short units as middle-length ones (45%, 51% for SH3, 51.5%, 46.0% for C2, respectively). This small difference seems to result from the difference in their individual tastes of sentence combining. In other words the college students may not always deteriorate in writing skills, but may not grow so much in point of sentence expansion. In the occurrences of long units also, the senior high was above the college students. We will mention about this and the comparison with native speakers later in the section of discussion.

The Japanese students do not prefer to use long units. The ratio of short units to long ones is, dramatically, 68 to .5 for the junior high, 45 to 4 for the senior high, and 51.5 to 2.5 for the sophomore. More-than-23 word T-units seem to be superlong for Japanese students. From these facts we can say that short T-units are characteristic of younger students in the early stage of learning English.

Two Ways of Lengthening T-units

Now we can see whether the language development to lengthen T-units is characterized more prominently by the tendency of the students to produce longer clauses or to produce a larger proportion of subordinate clauses. The ratio of clauses per T-unit is made by the number of all clauses divided by the number of T-units, or main clauses. The result shows one main clause plus X subordinate clause in ratio. It provides an arithmetical bridge for relating clause length to T-unit length. The mean clause length multiplied by the mean number of clause per T-unit equals the mean T-unit length. We can say from the decimal, for example, that 10 percent of the time the junior high students added a subordinate clause to a main clause. The clause to T- unit length factors are presented in Table 4.

Table 4　Factors Related to T-unit

	Mean length of clauses	Ratio of clauses per T-unit		Mean length of T-units	
JH3	6 67 words	X	1 10	=	7 35 words
SH3	7.53 words	X	1 36	=	10.22 words
C2	7.48 words	X	1 25	=	9.41 words

Distribution of T-units among Multi-T-unit Sentences

Besides subordination, coordination is also one of the factors to expand sentences. Another direction we can take for more detailed studies is to analyze to see how often each grade coordinate two, three or more T-units. The high frequency of such cannot be expected for Japanese students merely because they do not join even two T-units so often. The information is presented in percentage in Table 5. Majority of Japanese students wrote sentences with only one T-unit in them. One out of four or five joins two T-units in sentences, but there is no difference between the junior high and the senior high. It seems that the junior high students often use cordination, the senior high students make much use of both coordination and subordination, and the college students reduce the use of coordination.

Table 5 Distribution of T-units
among Multi-T-unit Sentences

	1 T-unit	2 T-units	3 T-units	4 Tunits
JH3	75.0 %	23.7 %	.8 %	.5 %
SH3	76.1 %	22.2 %	1.0 %	.7 %
C2	83.6 %	13.7 %	2.7 %	none

Table 6 presents T-unit-to-sentence length factors similar to Table 4 for comparison.

Table 6　Factors Related to Sentence

	Mean length of T-units	Ratio of T-units per sentence	Mean length sentence
JH3	7.35 words	X 1.16	= 8.53 words
SH3	10.22 words	X 1.14	= 11.65 words
C2	9.41 words	X 1.11	= 10.43 words

The descripancy between the senior high school students and the college students in most components and partial comparison with native speakers are discussed in the following section.

DISCUSSION

To summarize the questions raised from our data, (1) roughly all the three lengths indicate the growth of maturity, though not in the natural upward line, but between the high school level and the college level, there is some reverse change or some regression in a sense just in length, and (2) Japanese students reach far away from the ideal length they should aim at.

How should we make an explanation of the regression from the highest level of high school to the college level? Some admmitted this tendency because high school students are forced to orient themselves toward syntactic structures.[1] Others supported this

apparent regression by asserting that college students may shift their attention to the content or style.[2] It is proved that the three kinds of lengths involved in syntactic complexity develop remarkably in high school education.(in press)[3] The apparent descripancy may be supposed not to be a real regression but rather a form of development.

Then why do Japanese students stop or reduce the speed of developing the lengths of writing units in the course of progressing to the aimed lengths? Even the 12th graders of native language reach about 70% of aimed lengths. The intelligent maturity of Japanese students aged 18 to 20 must be about the same at least as that of American students aged 18. Japanese students' inability to make long writing units may be attributed to their language deficits — lack of knowledge about how to combine multiple concepts in one T-unit or one sentence. Here we can see the limitation of the formal grammar instruction in school education. Grammar is involved in all the four skills — hearing, reading, speaking, and writing. However, some students of foreign languages do not know how to use the formal grammar they have acquired. They need to learn how they should use their knowledge in different situations of performance. Even if they have their own linguistic competence, they may not learn the strategic competence.[4] It may be the teacher's responsibility to give students some suggestion of ideal length and ways of combining concepts in school. With the knowledge of desired length, they should turn their attention to the content, style, or effectiveness.

Do Japanese learners not develop the length of writing units after leaving school education? Here is a small data from 3 graduates' master's theses and 3 college teachers' papers written in English.

Table 7　Three Mean Lengths of Japanese Super Adults

	G-A	G-B	G-C	CT-A	CT-B	CT-C
Mean sentence length	19 55	18 83	25.00	25 14	16 90	25.69
Mean T-unit length	19 55	17 38	22.50	25 14	15 36	25 69
Mean clause length	11 32	10 76	11.84	14 67	12.07	16.70

Surprisingly their mean lengths correspond to those of native speakers. The elite remarkably have developed their writings even in length. Only length does not indicate the writing maturity. However, to say the least, we can maintain that they have acquired the ability of writing using large units.

SUMMARY and CONCLUSION

One of the English language learners' long-range goals as good writers is with the proper length of writing units. The intermediate levels show the dynamic development of syntactic structures used, but could not achieve even half of the goal in length. The

instruction to let students be aware of the proper length of T-unit or sentence may be effective. The verification by such an experimental study remains to be done in the near future. It is to be regretted that, because of the limitation of space, the explanation of conplexity related to the lengths of three units also remains to be done in the next opportunity.

Note

* This is a revised version of Chapter 2 and 3 of the master's thesis presented to the Faculty of the Department of English, Nanzan University, in 1983.
1. Some attendants showed the approving opinions on the backward change in my presentation session at the 30th Annual Convention of the Japanese Association of College English Teachers (JACET) held at Hokkaido University, on August 24,1991
2. Others expressed themselves on the possibility that the college students take acount of style at the small, local meeting held monthly at Nagoya University, on September 21, 1991.
3. Tōkai English Review No. 3, Tokai English Literature Society, 1992.
4. This idea came from the implication of Ms.Sasaki's presentation held at Nagoya University, on October 19, 1991.

References

Gaies, S. "T-unit Analysis in Second Language Research. Applications Problems, and Limitations" *TESOL Quarterly* 14(1), 53-60, 1980.

Hirano, K. "Research on T-unit Mewsures in EFL" 『上越教育大学研究紀要』 8, 67-77, 1980.

平野絹枝「言語能力の客観的指標の妥当性－日本人EFL大学生の場合－」『上越教育大学研究紀要』9, 65-77, 1990.

Mellon, J. C. *Transformational Sentence-Combining.* NCTE, Urbana, Illinois, 1969.

Hunt, K. *Grammatical Structures Written at Three Grade Levels.* NCTE Urbana, Illinois, 1965.
    "Measuring Sentence Development." *JACET BULLETIN* 5, 1974.
    "Teaching Syntactic Maturity " *Language Acquisition*, 英潮社, 1980.
    "How Little Sentences Grow into Big Ones" 小川洋通訳(1972) 「小さな文はどのようにして大きな文になるか」『応用変形文法』, 大修館書店, 1970.

Selinker, Larry. "Interlanguage." *New Frontiers in Second Language Learning*, Rowley, Mass.: Newbury House, 1974.

富田裕一「T-unitを用いた高校生の自由英作文能力の測定－総合的英語能力との関連」*STEP BULLETIN* 2, 14-27, 1990

Zamel, V. "Re-evaluating Sentence-Combining Practice." *TESOL Quarterly* 14, 81-90, 1980.

Appendix A

Direction: 次の英文のストーリーを読んでください．あまりうまく書けていないと思いませんか．ストーリーをよく研究して，もっとよい方法で書きなおしてください．文の形を変えたり,必要と思う語を加えたり,不必要と思う語を削除したりしてもかまいません．しかしストーリーの重要な部分を取り除いたり，ストーリーを変えたりしてはいけません．あなた自身の文で書きなおしてください．

Story: A man lived in a farmhouse He was old. He lived alone. The house was small. The house was on a mountain The mountain was high. The house was on the top. He grew grain He ate the

vegetables. He ate the grain. One day he was pulling weeds He saw something. A chicken was eating his grain The grain was new He caught the chicken. He put her in a pen The pen was under his window. He planned something. He would eat the chicken for breakfast The next morning came. It was early. A sound woke the man. He looked out the window. He saw the chicken. He saw an egg The chicken cackled The man thought something. He would eat the egg for breakfast He talked to the chicken Time passed He thought something He could feed her in the morning He could feed her at night Maybe she would lay two eggs every morning He fed the chicken more grain. She got fat. She got lazy She slept all the time. She laid no eggs The man got angry He had blamed the chicken. He killed her He ate her for breakfast. He had no eggs He talked to no one No one talked to him.
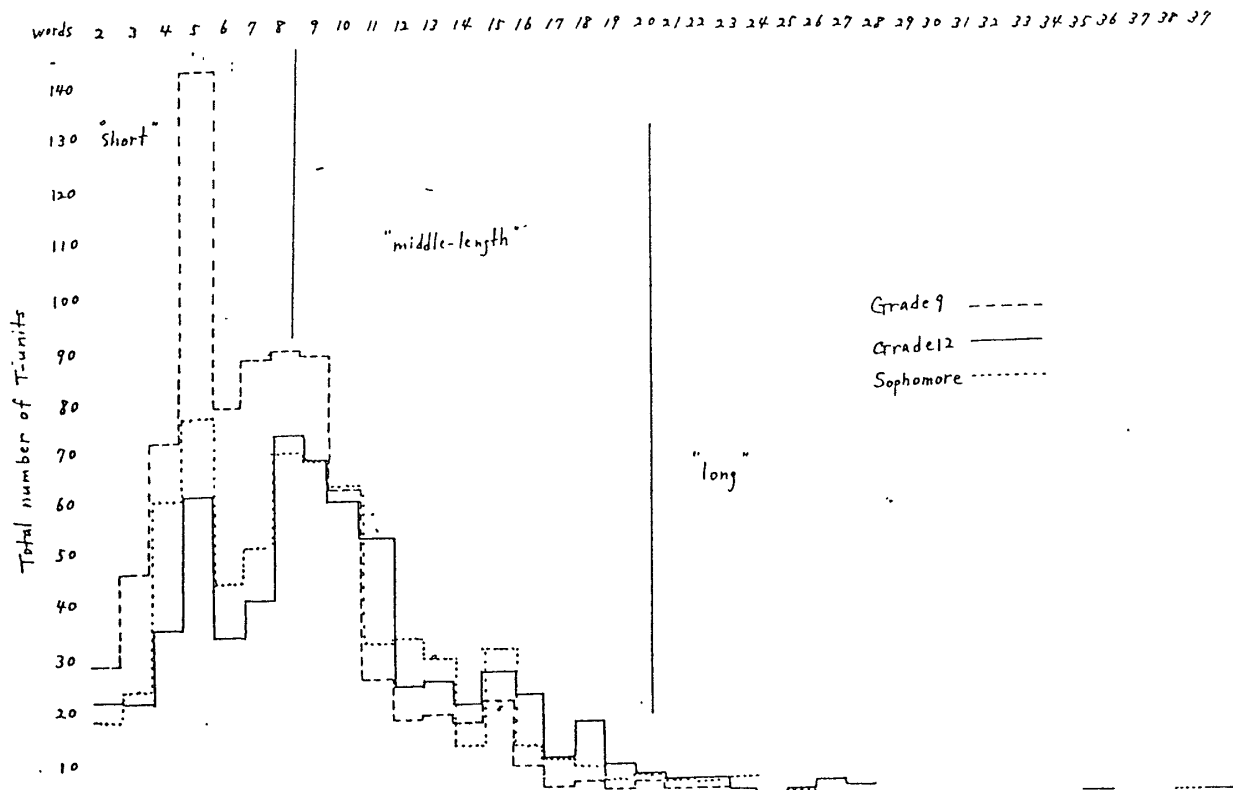
Note: grain 穀物　weed 雑草，草　pen 囲い　cackle くわっくわっと鳴く　lay (卵)を生む blame 責める

Appendix B

Figure 7　Occurrences of Each-Length T-units