博士学位論文（東京外国語大学）

Doctoral Thesis (Tokyo University of Foreign Studies)

氏　　名	村越　亮治
学位の種類	博士（学術）
学位記番号	博甲第 313 号
学位授与の日付	2021 年 5 月 26 日
学位授与大学	東京外国語大学
博士学位論文題目	The Development of the Grammatical Features in Japanese Upper Secondary School Students' English Compositions: A Longitudinal Study

Name	Murakoshi, Ryoji
Name of Degree	Doctor of Philosophy (Humanities)
Degree Number	Ko-no. 313
Date	May 26, 2021
Grantor	Tokyo University of Foreign Studies, JAPAN
Title of Doctoral Thesis	The Development of the Grammatical Features in Japanese Upper Secondary School Students' English Compositions: A Longitudinal Study

The Development of the Grammatical Features in Japanese Upper Secondary School Students' English Compositions:

A Longitudinal Study

日本人高校生の英作文に見る文法特性の発達

（長期的研究）

Ryoji MURAKOSHI

Tokyo University of Foreign Studies

2020

東京外国語大学大学院

総合国際学研究科　博士後期課程　言語文化専攻

村越　亮治

Abstract

The present study reports on the usage patterns and development of grammatical competence of one group of EFL writers in Japan. The data was collected from a longitudinal examination of the incidence of attempts with grammatical features in English compositions by Japanese upper secondary school students. A different composition task was given each academic year from pedagogical reasons.

First, the criterial grammatical features that discriminate the levels of the CEFR and the other target grammatical items in government-authorized English textbooks for lower secondary school were collected and listed. These items were then electronically extracted and sorted out by the Rasch model analyses from the longitudinal sets of compositions. The result showed that almost all the extracted items were items that were designated as criterial features at the A2 level of the CEFR or items featured in lower secondary school textbooks, namely, items taught in the earlier stages of English education in Japan. It was not until in the third year of upper secondary school that the items learned in the third year of lower secondary school emerged as 'acquired' grammatical items. When looking closer at the variety of use of the items, several items seemed to be used in the same strings as formulaic expressions, some of which showed the transition to extensive use over the years. As well as the trajectories of acquisition of the grammatical items, the development of the Mean Length of Utterance (MLU), a measure of syntactic complexity, was investigated, and it was confirmed to be statistically significant. However, the MLUs at the three academic years were all found to fall into the A2 level or lower.

Then, the composition data from another group of EFL writers was analyzed in order to identify the task-dependent grammatical items. After the identification of these items, the trajectories of grammatical competence of the main subjects were revisited,

and the development of acquisition of the task-independent items was confirmed.

The MEXT's goal of English education for upper secondary school students is to acquire the CEFR B1-level proficiency by the time of graduation. Supposing that grammar knowledge is considered to be part of this 'proficiency', the fact that no B1 grammatical items were judged as 'acquired' in the students' compositions, may be a testament to the issues of teaching practice in English class in Japan. Teachers of English need to teach the context in which each grammatical item is often used, and many writing opportunities should be given in class in order for the students to try to use their linguistic knowledge at will and acquire it more steadily through trial and error. Also, different types of tasks should be given to learners so that they can try to use as various items as possible. The use of particular grammatical items could also be an indicator of writing proficiency of learners. Collecting learners' output data as much as possible and analyzing it would help, not only reconsider language activities, but also design more practical syllabi and create more effective teaching materials for teaching English in the future.

Keywords: CEFR, CEFR-J, criterial features, EFL writing, English composition, grammatical items, Japanese, secondary school, Mean Length of Utterance (MLU), Rasch model, syntactic complexity

Acknowledgements

I would like to express my deepest gratitude to Professor Masashi Negishi, my supervisor, for making a number of helpful suggestions and giving me warmhearted encouragement throughout my course of study at Tokyo University of Foreign Studies, and even after the course. My heartfelt thanks are also due to Professor Yukio Tono and Professor Asako Yoshitomi, who gave me a lot of helpful advice on this dissertation and precious opportunities to study linguistic theories and English education at the university. Also, Professor Emeritus Hideyuki Takashima gave me inspiring lectures and warm encouragement.

I am extremely grateful to Mr. Yoshihito Fujita, who helped me collect the composition data at his school. Without his cooperation, this study could not have been completed.

My appreciation also extends to the 507 students at two high schools, who agreed to offer the valuable data of their English compositions.

I am also greatly indebted to Ms. Thoa Thi Nguyen, my wonderful coworker, who kindly spared her precious time for proofreading my thesis and giving me helpful suggestions.

Lastly, I am solely responsible for all errors in this thesis.

Table of Contents

List of Tables

Page

List of Figures

Page

List of Abbreviations

Abbreviation	Meaning
CAE	Certificate in Advanced English
CEFR	Common European Framework of Reference for Languages
CEFR-J	CEFR-Based Framework for English Language Teaching in Japan
CLC	Cambridge Learner Corpus
CPE	Certificate of Proficiency in English
EFL	English as a Foreign Language
EGP	English Grammar Profile
ESOL	English for Speakers of Other Languages
FCE	First Certificate in English
IELTS	International English Language Testing System
KET	Key English Test
MEXT	Ministry of Education, Culture, Sports, Science and Technology
MLU	Mean Length of Utterance
NICT JLE	National Institute of Information and Communications Technology Japanese Learner English
NS	Not Significant
PET	Preliminary English Test
RASP	Robust Accurate Statistical Parser
RLD	Reference Level Description
STEP	Society for Testing English Proficiency

Chapter 1. Introduction

For the things we have to learn before we can do them, we learn by doing them.

– Aristotle

The only source of knowledge is experience.

– Albert Einstein

Since 2014, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan has given a nationwide English proficiency test to third-year upper secondary school students all over Japan. The result shows that the English proficiency of upper secondary school students is far behind the MEXT's goals. Prior to the test, the MEXT had announced the specific goals of English education in Japanese secondary schools. The report "Five Proposals and Specific Measures for Developing Proficiency in English for International Communication" was issued in June 2011 and presented the MEXT's goals for upper secondary school students of attaining the proficiency corresponding to Eiken[1] Grade 2 or pre-2. Following that proposal, the MEXT issued another report in September 2014, titled the "Report on the Future Improvement and Enhancement of English Education: Five Recommendations on the English Education Reform Plan Responding to the Rapid Globalization". This report set even higher target for upper secondary school graduates. In it, the MEXT recommended that 50 percent of upper secondary school graduates should score at Eiken Grade pre-2 or 2 or higher, depending on students' individual situations and career goals.

The test was administered in 2014, 2015, and 2017, in order to examine whether students' English skills in listening, speaking, reading and writing had been developed in a well-rounded way. The result of the 2014 test, which approximately 70,000 third-year

upper secondary school students took, showed that overall, a majority of the participants scored at or below the level of Grade 3, the supposed target for lower secondary school graduates. The MEXT used the Eiken grades as the standards of goals, probably because they are familiar to Japanese students, teachers and parents. Along with the Eiken grades, the levels of the Common European Framework of Reference for Languages (CEFR: Council of Europe, 2001) were held as the goals, because the CEFR is more globally used and its levels show what the learners can practically do with languages. That "can-do" policy of the CEFR matches what the MEXT expects from Japanese learners of English. In comparison, Eiken Grade 2 is roughly equivalent to B1 level of the CEFR, Grade pre-2, A2, and Grade 3, A1. Therefore, most of the participants were proven to belong to A1 level of the CEFR (Listening, 75.9%; Speaking, 87.2%; Reading, 72.7%; Writing, 86.5%) (MEXT, 2015). In the test of 2015, even though some progress was admitted, on average, more than 75 percent of the participants (another group of approximately 81,000 third-year upper secondary school students) were found to be still in the level of A1 (Listening, 71.9%; Speaking, 87.2%; Reading, 66.4%; Writing, 80.4%) (MEXT, 2016). The test of 2017 involved approximately 60,000 third-year upper secondary school students all over Japan. The result was almost the same as the previous year's test: on average, about 75 percent of the participants were judged to be in A1 level (Listening, 66.4%; Speaking, 87.1%; Reading, 66.5%; Writing 80.3%) (MEXT, 2018). All these results show that even after five years of learning English at school, many of the third-year students of upper secondary school have fallen far short of the MEXT goals and that they especially have challenges in acquiring productive skills of English, which indicates that what students learn in English class has not been effectively applied to what they can do using English.

The results on the writing portion were particularly disconcerting. It consisted of two parts of free writing: a summary of information they listened to and an argumentative

essay on a provided topic. In addition to the high percentage of A1 levels, it is very alarming that in the 2014 test, 29.2 percent of the participants scored zero or gave no response to the writing tasks, which accounts for the largest portion in the score distribution. The percentage decreased in the 2015 and 2017 tests (17.6%, 15,1%, respectively), but the inadequacy of students' writing skills still cannot be overlooked. Moreover, when we take a closer look at the details of the scoring, we can see one of the latent factors of the problem. In the argumentative essay, the skill of manipulating grammar knowledge appropriately was assessed on a scale of zero to four. It was found that in the 2014 test, 38.6 percent of the participants scored zero points (defined as no response or irrelevant content), and 36.9 percent scored one point (defined as many global errors that hinder comprehension found). In the following year on the 2015 test, even though the percentage of students who did not respond or scored zero decreased from 30.4 percent to 18.1 percent, 82.1 percent was still judged to be in A1 level. As for the use of grammar in the argumentative essay, 49.5 percent scored zero points, and 40.1 percent scored one point. These results were actually worse compared to the previous year.

The Course of Study for Lower Secondary School has provided the items of English grammar that are to be taught in the three years of lower secondary school (MEXT, 2008a, 2008b). Since most of the lower secondary school students in Japan take entrance examinations of English for upper secondary schools, where grammar knowledge is needed, all the listed items ought to be covered in English class. Similarly, in the Course of Study for Upper Secondary School, especially with regards to *English Communication I*, a required English subject for all upper secondary school students, there are suggestions of particular grammatical items to teach (MEXT, 2009a, 2009b). Many of the items, however, have already been taught in lower secondary school. Both of the courses of study for secondary schools emphasize that grammatical knowledge of English should be

instructed as a means to support communication through effective linkage with language activities, and that the instruction should center on actual use of grammatical items. However, the results of the tests mentioned above show that the input of the grammatical knowledge did not seem to be internalized enough to be utilized in more realistic tasks, and that students' general writing skills have not been sufficiently developed, either. The participants of the tests were all third-year upper secondary school students and should have learned linguistic knowledge more advanced than A1 level. However, they were not able to draw on their grammatical knowledge to complete the writing tasks, even if they could answer grammar questions on other objective tests (e.g. a fill-in-the-blank type of test).

The survey conducted in parallel with the nationwide test mentioned above found that less than half of the respondents experienced summary writing or essay writing in class in the previous academic year (2014, 38.7%; 2015, 42.5%; 2017, 47.8%), and that almost the same percentage of upper secondary school teachers of English give their students summary or essay writing activities (2014, 39.7%; 2015, 46.7%; 2017, 50.8%).[2] Although the situations of classroom practice seem to be becoming better, more improvement should still be expected. Learners need more opportunities to use their English writing skills as this is how their linguistic knowledge can be assimilated and organized into their linguistic repertoire.

---

[1]Eiken: The Practical English Proficiency Test (sponsored by the Society for Testing English Proficiency [STEP])

[2]The results of 2014 and 2015 are elicited from the data of the prefectural/municipal upper secondary school participants (students and English teachers), whereas the 2017 ones are overall results of the national and prefectural/municipal upper secondary school participants, since the detailed report of the 2017 survey is not issued.

As with the Course of Study of Japan, the Core Inventory for General English (North et al., 2010) categorizes grammatical items of English as a second language, as well as functions, vocabulary, discourse markers, and topics of language activities. One of its advantages over the Course of Study is that by the analyses of teaching materials and the insights and experiences of teachers and researchers, all the items are sorted according to the CEFR levels from A1 to C1. The selecting and sorting of the items is based on "good practice" of instructors of English all over the world. In other words, the inventory reflects actual, real-life instruction (Negishi, 2012a). The Course of Study for Lower Secondary School used to specify the grammatical items to teach according to student's year in school. However, in 1989, the MEXT decided to eliminate this rigid system and instead view all these grammar items as a whole in the hopes to make teaching these grammar points easier for teachers and publishers (MEXT, 1989). Yet, in government-authorized textbooks, they are sorted in pedagogically traditional orders by tacit agreement between teachers and textbook publishers that particular items should be covered in a particular academic year. All the learners are supposed to learn the same items at roughly the same time. Thus, all the learners are expected to acquire the same items at roughly the same pace. There seems to be no room for considering individual difference of proficiency development. Another advantage of Core Inventory is that it encourages teachers and syllabus designers to refer to an inventory of minimal, yet precise "core" items and make their own decision about what to take, not to take, or even recycle, based on their learners' needs and interests. They have the choice and autonomy when they teach learners or make the syllabi for different proficiency levels. To be fair, however, the Course of Study for Lower Secondary School was designed for compulsory education, so it is doomed to be decisive and have to provide the knowledge and skills to acquire, as well as the learning attitudes to cultivate, that all the Japanese citizens should have.

Whether the items are listed in a traditional way or based on the proficiency levels (such as the CEFR), one thing is for sure: with or without the labeling of proficiency levels or the designated timing of instructions, grammar should support the use of language skills, not exist merely as linguistic knowledge. Learners should be able to use the language at will, not just keeping it somewhere in their brain as dispersed knowledge only for answering fill-in-the-blank questions on tests.

One research project has actually been listing the grammatical items based on the English proficiency levels. The English Profile Programme is a global research program led by Cambridge University and some other organizations. They have been working to elucidate what aspects of the English language are typically acquired at each CEFR level (English Profile, 2015). They have analyzed learners' language and found the grammatical features that can be criterial to judge learners' CEFR levels, as well as the lexical criterial features. They provide the lists of "criterial features" of grammar from A2 to C2 in the CEFR (Hawkins & Buttery, 2010; Hawkins & Filipović, 2012). Their data is elicited from Cambridge Learner Corpus (CLC), which compiles English composition data from Cambridge ESOL Examinations of different levels. The whole set of criterial features of grammar is named the "English Grammar Profile", and the set of lexical features is called the "English Vocabulary Profile". Both of these are available on their web site. In addition to their potential of level discriminability, it is interesting to see that the same structure or the same vocabulary word could fall into different levels, depending on its pragmatic usage. This data-driven research may be able to reveal the state of procedural grammatical knowledge that learners of English as a second / foreign language have acquired and are able to use in a proficient manner.

Since most young Japanese learners' command of English has been proven to be quite unsatisfactory compared to the goals established in the Course of Study, it would be

meaningful to reveal how much of basic grammatical knowledge they are estimated to have acquired as a foundation for written production, and compare it to the CEFR-labeled grammatical features. Consequently, this would also be helpful to reflect on pedagogical practices with regards to English education in Japanese public schools. Also, extracting assumedly-acquired grammatical features by a statistical method and collating that information with the history of their instruction might unveil one aspect of actual situation of Japanese students' English learning. Hopefully, it would also contribute to provoke thought of more effective teaching procedures, syllabi, and means of assessments, at upper secondary schools in Japan.

In the following chapter, I will review the studies of the English Profile Programme regarding grammatical features of learner language, and visit some officially-issued lists of grammatical items. Then, the stream of second language acquisition research on grammatical features will be surveyed, in order to compare the findings in previous studies and the present study later. In Chapter 3, first, I will try the statistical extraction of grammatical items that the group of learners can use at will in their written production. Then, the calculation of statistical relationship between syntactic complexity and fluency of the learners' compositions will be attempted. Also, I will examine how much correlation there is between syntactic complexity and the learners' overall scores of writing proficiency. The next chapter is devoted to the follow-up study, where the task-dependent grammatical items will be sorted out from the pseudo-longitudinal compositions written by another groups of EFL learners. In Chapter 5, I will interpret the results of both the main study and the follow-up study, and explore pedagogical implications of the present study, as well as its limitations and future perspectives. Finally, Chapter 6 concludes by summarizing and tying together the discussions in the preceding chapters.

Chapter 2. Previous Studies

2.1 English Profile Programme and "Criterial Features"

The English Profile Programme is an ongoing collaborative research project which aims to investigate what learners of English at different levels can and cannot do using English, and how well they perform with the language at their disposal (Milanović, 2009). In other words, its goal is to establish a set of Reference Level Descriptions (RLDs) for English for all six levels of the Common European Framework of Reference (CEFR: Council of Europe, 2001) from A1 to C2 (Salamoura & Saville, 2009, 2010; Saville, 2010; Harrison, 2015a). To be more concrete, it aims to add specific grammatical and lexical details of English to the functional characterization of the CEFR levels, on the assumption that there are certain linguistic properties that are characteristic and indicative of L2 proficiency at each level (Hawkins & Buttery, 2010; Saville, 2010). The necessity of the English Profile Programme arose from the underspecification of the CEFR. It is a common framework for all languages and must be neutral and underspecified by its nature, so its "can-do" descriptors do not give details about the grammar and lexis that are characteristic of each proficiency level. Thus, teachers, examiners, or syllabus designers of a given language need to determine the linguistic features that increasing proficiency in English entails, because the ability to "do" the particular task does not show how the learner does it and with what grammatical and lexical properties (Milanović, 2009; Hawkins & Buttery, 2010). In sum, the English Profile Programme intends to establish and compile "profiles" of learners of English at different CEFR levels. It should give considerable benefit to teaching/learning, examining and publishing of English, as well as reconstructing theories of second language acquisition (Hawkins & Filipović, 2012).

The Cambridge Learner Corpus (CLC) has been used in research designed to extract the characteristic language properties of English learners. The corpus incorporates examinees' written production in the Cambridge English exams. *Figure 1* shows the exams that provide the writing data for the CLC. Taken together, they cover the CEFR levels from A2 to C2 (Hawkins & Buttery, 2009).

CEFR	Cambridge English Qualifications		IELTS
	General and higher education	Business	
A2	Key / KET		
B1	Preliminary / PET	Preliminary	IELTS
B2	First / FCE	Vantage	
C1	Advanced / CAE	Higher	
C2	Proficiency / CPE		

Figure 1. The CEFR levels and the exams that provide the writing data for the CLC

The CLC currently contains over 50 million words and is still growing each year (Harrison, 2015a). For the English Profile Programme, the CLC has been tagged for parts of speech and parsed using the Robust Accurate Statistical Parser (RASP) (Briscoe et al., 2006), which enables the researchers to obtain grammatical information of the learner language (Hawkins & Buttery, 2009). At present, the Cambridge English Profile Corpus and the Cambridge English Corpus have also been the subjects of analysis. The Cambridge English Profile Corpus collects its data from real classrooms or virtual courses specifically for English Profile Programme. It comprises of essays, course work, and spoken data produced by learners of English around the world. The Cambridge English Corpus covers more than a billion written and spoken words collected from published

resources, public speeches and broadcasting, which enables the English Profile researchers to compare the learner language with the expert or native speakers' language (Harrison, 2015a).

The characteristic linguistic features acquired in each CEFR level are called "criterial features". They are not micro-level, morphological items, but bigger grammatical chunks, although there are some lexical features involved. Those features can serve as a basis for the estimation of a learner's proficiency level, discriminating it from adjacent levels (Salamoura & Saville, 2009, 2010). Hawkins and Buttery (2010) classified the criterial features into four types: positive linguistic properties, negative linguistic properties, positive usage distributions for a correct property, and negative usage distributions for a correct property. Positive linguistic properties are correct properties of English that are acquired at a certain level and that generally persist at all higher levels, while negative linguistic properties are incorrect properties or errors that occur at that level, and with a characteristic frequency. Positive usage distributions for a correct property match the corresponding distribution for native speakers of English, while negative usage distributions for a correct property do not match the distribution of native speaking users. With regard to the diagnostic potential of criterial features, Salamoura and Saville (2010) argue that the occurrence or nonccurence of criterial features in a learner's output can diagnose their CEFR level and/or distinguish the learner from other learners whose use of the same criterial features significantly differs from that of him/her (p.109). Regarding the advantages in teaching, Hawkins and Buttery (2009) mention that if the characteristic properties of each CEFR level and of the next attainable stage in learning are defined precisely, materials and syllabuses can be calibrated with much greater precision (p.174). Hawkins and Filipović (2012) also claim that teachers can better understand what is easy and what is not easy for learners of English, blending

new linguistic viewpoints with knowledge from their experience (pp.84-85).

In order to specify positive linguistic features, which is the focus of the present study, the "10-to-1" rule was adopted. At two adjacent levels (e.g. A2 and B1), if the quantity of occurrence for a structure at the higher level exceeds that at the lower level by a ratio of 10 to 1, they place it as criterial at the higher level (Hawkins & Filipović, 2012). This analysis is possible because every piece of writing is CEFR-labeled from the outset and the occurrence of certain features found can be counted and compared between different levels. The positive linguistic features have now developed into the "English Grammar Profile (EGP)". The EGP shows how learners' grasp of English grammar progresses as their proficiency improves, in terms of a) the new structures and features which are learned at each level, and b) the ability to use structures already learned with different words or for different functions (Harrison, 2015b). Table 1 shows the positive linguistic features that are criterial for discriminating learners' CEFR levels (Hawkins & Filipović, 2012).

Table 1

Criterial grammatical features in the CEFR levels

Levels	Features
A2	Intransitive clauses (e.g. *You can go to Yilte Park.*)
	Transitive clauses (e.g. *Now I wrote a post card for you …*)
	Verbs with a finite complement clause (e.g. *I think the zoo is an interesting place.*)
	Verbs with an infinitival complement (Subject control) (e.g. *I want to buy a coat.*)
	Direct WH-questions (e.g. *Where is the park?*)
	Pronoun plus infinitive (e.g. *You can bring something to eat if you want to.*)
	Ditransitive clauses (e.g. *I can give you my guitar.*)

	Postnominal modification with *–ed* (e.g. *There are beautiful paintings painted by famous Iranian painters.*)
	Double embeddings of [of] (e.g. *I like the colours of the back of the mobile phone.*)
	Modal auxiliary verbs: may, might, can (possibility), must (obligation), should (advice) (e.g. *We must be there at7 o'clock in the morning. / You should wear old clothes because we will get dirty.*)
B1	Postnominal modification with *–ing* (e.g. *I put an advertisement asking if someone had it.*)
	Verbs with an infinitival complement (Object control) (e.g. *I ordered him to gather my men to the hall.*)
	Verbs with *–ing* complements (Object control) (e.g. *Maria saw him taking a taxi.*)
	It Extraposition with finite clauses (e.g. *It's true that I don't need a ring to make me remembering you.*)
	Verbs with a Prepositional Phrase plus finite complement clause (e.g. *He said to me he would like to come back soon.*)
	Genitive relatives (e.g. *I met a very nice boy whose name's John.*)
	Pseudoclefts type(i): WH-NP-VP (e.g. *I opened the door and what I saw was so amazing.*)
	Indirect WH-questions in finite clauses (e.g. *I can't understand what she is saying.*)
	Indirect WH-questions in infinite phrases (e.g. *I didn't know what to buy for you.*)
	Complex auxiliaries: would better, had better (e.g. *If you don't like to go with them, you had better tell them why you don't want to come.*)
	Adverbial subordinate clauses with *-ing* (following main clause) (e.g. *He was sitting there, drinking a coffee and writing something.*)
	Subject-to-Subject Raising verbs and adjectives: seem, supposed (e.g. *I was supposed to go to the English class.*)
	Subject-to-Object Raising verbs (unpassivised): expect, like, want (e.g. *I expected it to be more difficult, but it is not so hard.*)
	Tough Movement: easy (e.g. *The train station is easy to find.*)
	Double embeddings: of [-s] (e.g. *I am a big fan of the world's most famous British secret service agent.*)
	Modal auxiliary verbs: may (permission), must (necessity), should (probability) (e.g. *May I suggest that you book me in for the new accommodation? / She must be feeling so happy.*)

B2	Adverbial subordinate clauses with *-ing* (preceding main clause) (e.g. *Talking about spare time, I think we could go to the Art Museum.*)
	It Extraposition with infinitival phrases (e.g. *It would be helpful to work in your group as well.*)
	Pseudoclefts type(ii): WH-VP (e.g. *What fascinated me was that I was able to lie on the sea surface and read this newspaper.*)
	Verbs with an NP plus finite complement clause (e.g. *She told me that she had worked for summer camp for children.*)
	Secondary predications (e.g. *Just go and paint the houses yellow and blue.*)
	Subject-to-Subject Raising verbs and adjectives: appear, cease, fail, happen, prove, turn out, certain, likely, sure, unlikely (e.g. *My worries proved to be wrong. / Whenever money is involved, some problems are likely to happen.*)
	Subject-to-Object Raising verbs (unpassivised): imagine, prefer (e.g. *I would prefer my accommodation to be in log cabins.*)
	Subject-to-Object Raising verbs plus passive: expected, known, obliged, thought (e.g. *How many hours a day should I be expected to work?*)
	Tough Movement: difficult, good, hard (e.g. *The grammar and vocabulary are a bit hard to learn.*)
	Double embeddings: [of] –'s (e.g. *After this I went to a friend of mine's house where I spent one week.*)
C1	Subject-to-Subject Raising verbs and adjectives: chance (e.g. *I chanced to know your Competition from an international magazine.*)
	Subject-to-Object Raising verbs (unpassivised): believe, find, suppose, take (e.g. *I find this to be more interesting than the walking route to Lake Hawksmere.*)
	Subject-to-Object Raising verbs plus passive: assumed, discovered, felt, found, proved (e.g. *Internet is a valuable tool which can be proved to be the most important aspect in the learning process.*)
	Double embeddings: [–'s] –'s (e.g. *After spending the first day of their marriage in the bride's family's house…*)
	Modal auxiliary verbs: might (permission) (e.g. *Might I tell you what we discuss?*)

C2	Subject-to-Object Raising verbs (unpassivised): declare, presume, remember (e.g. *They declare some products to be the hits of the season.*)
	Subject-to-Object Raising verbs plus passive: presumed (e.g. *Not only meetings with people are presumed to give new experiences.*)
	Tough Movement: tough (e.g. *What she knew would be really tough to live with was the reason of his death.*)

* excerpted from Hawkins and Filipović (2012)

In the English Profile Programme, it is found that the length of a sentence learners produce can also be "criterial" for discriminating their CEFR levels, in that the more grammatical items learners can use correctly, the more complex and longer their sentences can be (Hawkins & Filipović, 2012, pp.22-23). This index is called the Mean Length of Utterance (MLU), which is the mean of the numbers of words in all the sentences judged as being complete sentences. It is shown that the MLU of A2 level is 7.9 (words), B1, 10.8, B2, 14.2, C1, 17.3, and C2, 19.0, and that one of the factors that makes sentences complex is the subordinate structure, such as the because-clause or although-clause (Hawkins & Filipović, 2012, pp.28).

The indices of sentence complexity have been compared in several studies. Wolfe-Quintero et al. (1998) conducted meta analyses of more than 100 indices of fluency, complexity and accuracy of produced language used in 40 studies, in order to find the indices that can best indicate linguistic development. Concerning complexity, they claimed that grammatical complexity means that a wide variety of both basic and sophisticated structures are available and can be accessed quickly, whereas a lack of complexity means that only a narrow range of basic structures are available or can be accessed (p.107). Polio (2001) questioned the validity of complexity, arguing that more

complex sentences may not necessarily be indicative of quality, and that at an advanced level, too many complex sentences may be a problem. However, Ferris (1994) showed that there was positive relationship between syntactic complexity and L2 writing ability level. Many researchers have used average length of T-unit (Hunt, 1965, 1970) as a measure of syntactic complexity. For example, it was positively related to L2 grade level in Yau & Belanger (1985) and language development in Lim (1983). A comparison of the means of average length of T-unit across 40 studies in Wolfe-Quintero et al. (1998) shows that there is a range from 6.0 words per T-unit for the lowest level learners to 23.0 for the most advanced, with words per T-unit increasing in a linear relationship with proficiency. However, they prioritized another complexity measure using T-unit, average number of clauses per T-unit, as it correlated best with development. Meanwhile, Lu (2011) maintains that in order to indicate the development of complexity, the indices using clause, such as the mean length of clause, were most effective, but he also admits the usefulness of the mean length of sentence. Bardovi-Harlig (1992) argues the indices using sentences are more effective, because even the nonreversible coordinate structure (e.g. *I got home and I opened the front door.* / **I opened the front door and I got home.*) could be divided into two units in indices that use T-units as a measure, thereby decreasing the complexity of language unfairly.

(e.g.) *I got home and I opened the front door.* - one sentence (mean: 9.0)

- two T-units (mean: 4.5)

In Ferris (1994), which examined the correlations between holistic ratings for writing and 28 indices, the length of sentence discriminated the learners' proficiency levels reasonably enough, though the results were not statistically significant (p.416).

In Murakoshi (2015), I pursued the development of the MLU, examining the longitudinal data of the English compositions written by the Japanese upper secondary

school students. The text types of the compositions were narrative (the first year), narrative (the second year) and argumentative (the third year). The development of the MLU showed statistical significance, but the average did not reach the B1 level criterion even in the third year of upper secondary school. On closer look, the development from the second year to the third was more rapid than that from the first year to the second. This might be because of the difference of the text type of the compositions. Beers and Nagy (2009) argue that the complexity of learners' written production was higher in persuasive essays than in narrative essays. Lu (2011) also mentions that argumentative compositions produce more complex structures than narrative. Furthermore, Newton and Kennedy (1996) refer to the use of conjunctions that is more frequent in argumentative tasks. Then, I gave particular attention to the use of subordinate structures as an enhancing factor of the MLU. While the use of some structures, such as that-clause, because-clause and relative-clause, showed the progressive development, it was assumed that the use of some structures, such as when-clause and if-clause, was influenced by the task requirement. Among those above, many students tried to use because-clauses through the three years, but there were many errors found in its use, especially in early times. In the first year, 75 percent of the students tried to use this structure, but 90 percent of them made the same error as "*I like my sister*. **Because she is kind to me*" (Murakoshi, 2012a). As for the coordinate structures, I adopted the classification in Bardovi-Harlig (1992). If the structure was not irreversible, it was dealt as one sentence (e.g. *I got home and I opened the front door*. / *I visited Fred but he was out*.). If the structure was reversible, it was dealt as two sentences, which caused less MLUs (e.g. *I like soccer and I like baseball*.). The excessive use of coordinate conjunctions (more than two), which might produce longer sentence in vain, was not found in the data (e.g. *I went shopping and I bought a shirt and I ate a hamburger*.).

2.2 Teaching and Learning of Grammatical Items

2.2.1 The Core Inventory for General English

Compared to the work on the criterial features in the English Profile Programme, which is based on learner corpora, the Core Inventory project shows the CEFR-labeled grammatical items that were drawn from popular course books and the syllabuses of excellent teachers, not based on data-driven analyses. In other words, the Core Inventory is based on experience and consensus, not data-based analysis of the language that learners actually use (North et al., 2010). Negishi (2012a) refers to this comparison, saying that the Core Inventory reflects the reality of teaching English, and the English Profile Programme, the results of learning English. He points out that there are some items that are supposed to be instructed in early stage but actually acquired much later, while others are classified in the same level in the Core Inventory and the English Profile.

2.2.2 The Course of Study in Japan

In Japan, the MEXT issues the Course of Study, where particular items of English grammar are listed to be instructed in lower and upper secondary schools (MEXT, 2008a, 2008b). The English teachers are required to teach all of them, and the textbook publishers have to cover those items in their government-authorized publications. On the contrary, the aim of the Core Inventory is not to tell teachers what to teach, but to provide a simple overview of the apparent consensus on what constitutes the most important content for teaching and learning at each CEFR level. The users of the Core Inventory are encouraged to analyze the learners' needs in order to establish the basis for actual teaching, both in curriculum level and in classroom level (North et al., 2010). It should be fruitful to see the gaps between the grammatical items officially taught in schools and the ones actually acquired and used at will in the Japanese educational context, and to think about how to

fill in the gaps, with a side glance at the instructional items in the Core Inventory, the findings of the English Profile Programme, and the insights from other research projects.

2.2.3 The CEFR-J Grammar Profile

The CEFR-J[3] project, which was substantially launched in 2008, was aimed at creating an adapted version of the CEFR for English language teaching and learning at primary, secondary and tertiary education in Japan (Tono, 2012, 2013, 2017). It was motivated by the preceding research project, which proposed the application of the CEFR to the situation of Japan (Koike, 2008). However, since approximately 80% of Japanese learners of English are considered to belong to A-levels even after they have received 10 years of English instruction at secondary and tertiary levels (Negishi et al., 2012), the researchers needed to specify the A-levels in even greater detail in order to properly diagnose the language abilities of this majority of beginners (Tono, 2017). As a result, the CEFR-J divides A1 into three levels (A1.1, A1.2, and A1.3) and A2 into two levels (A2.1 and A2.2). Moreover, B1 and B2 are respectively divided into B1.1 and B1.2, B2.1 and B2.2. What is more, with the introduction of English education to elementary schools, pre-A1 level is added below A1 level (Tono, 2013). The CEFR-J has been influencing the English education system in Japan, and some material / test developers have already utilized this framework for their products.

As a part of the CEFR-J project, the criterial grammatical items have been extracted from several corpora and labeled, a set of which is called "the CEFR-J Grammar Profile[4]." It is basically a list of the CEFR-J-labeled grammatical categories or items as instructional input. The researchers of this project created and analyzed three corpora. In order to create a corpus of input language, they collected the CEFR-labeled ELT course books used in the UK since none of the authorized textbooks in Japan are CEFR-labeled. They even

used two corpora of output language. One was JEFLL Corpus, a collection of composition written by Japanese lower and upper secondary school students. The other was NICT JLE Corpus, a collection of spoken data from Japanese learners of English. The main resource was the ELT course book corpus because the aim of this part of the project was to propose a way to present grammatical items as instructional input in order to design a syllabus based on the CEFR-J. The other output corpora were supplementarily used to examine whether Japanese learners of English can manipulate those items or to extract characteristic errors in each level (Ishii & Tono, 2016; Tono, 2016).

2.3 Development of Particular Grammatical Features in Learner Language

The term "development" is used for describing some processes of learner language acquisition, and such processes often show certain patterns. Ellis (2008) explains that in research of second language acquisition, "developmental pattern" is a cover term for the general regularities evident in the acquisition, such as an order of acquisition and a sequence of acquisition. The former represents which target-language feature is acquired before another. The latter concerns how a particular linguistic feature is acquired over time. In early research of second language acquisition, many researchers focused on the order of acquisition of different linguistic features, such as morphemes. The morpheme studies investigated the order of acquisition of grammatical functors (i.e. articles) and inflectional features (i.e. plural -s). For example, Krashen (1982) summarized the order of acquisition of grammatical morphemes as in *Figure 2*.

---

[3]The CEFR-J (version 1.1) is available at: http://www.cefr-j.org /download.html#cefr

[4]The CEFR-J Grammar Profile is available at: http://cefr-j.org/sympo2018/pdf/CEFRJGP_GRAMMATICAL_ITEM_LIST.pdf#back

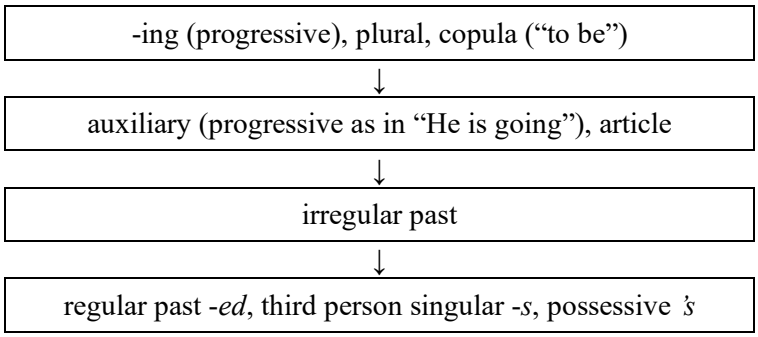


Figure 2. Order of acquisition of grammatical morpheme

In more recent years, tense and aspect have been attracting the researchers' interests. Basically, the studies concerning development of tense and aspect try to investigate the order of different forms describing different tense and aspect, involving the acquisition of explicit verbal morphology. It is noteworthy that the researchers began to turn their eyes toward not only forms but also functions. Among them, Klein (1995) conducted a longitudinal study on an L2 learner of English and identified the order of English tense-aspect morphology emerged as follows:

1. Third person -s and present tense copula
2. Irregular past tense forms and Verb-ing
3. Present perfect forms
4.Regular past tense forms
5. Future with “shall” or “will”
6. Past perfect forms

*excerpted from Ellis (2008)

Bardovi-Harlig (2002) focused on the use of the future expressions "will" and "going to" in students' speech and writing, and found that "will" emerges quite early and that "going to" tends to be used as the formula "*I am going to write about …*" first, and then used with different persons, numbers, or verbs. These findings suggest that the formulaic use of language may develop into constructions and become part of learners' grammatical knowledge that they can retrieve at will when necessary. The present study will also observe the use of "will" and "going to" and compare the results with these findings.

Meanwhile, some researchers got interested in the acquisition of particular syntactic structures. One of those structures is negative sentences. In Lightbown and Spada (2013), the developmental sequence of negation is summarized in four stages, as shown in *Figure 3*.

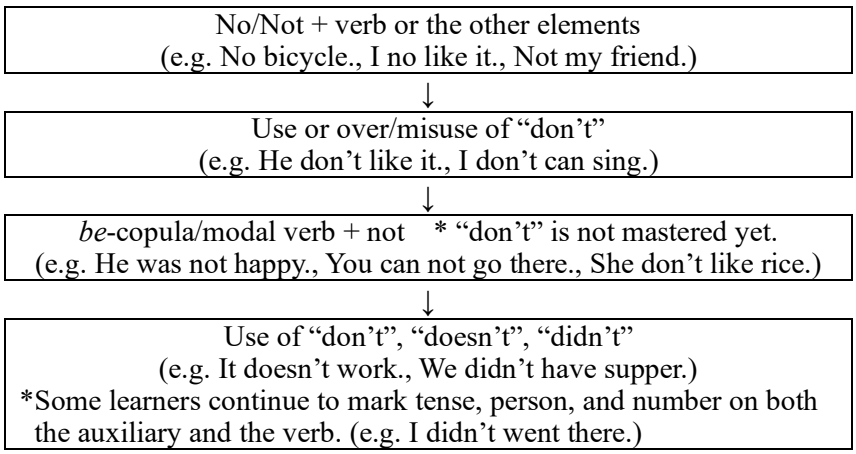


Figure 3. Developmental sequence of negation

The developmental stage of question forms has also been a focus of the research. With reference to Pienemann, Johnston and Brindley (1988), Lightbown and Spada (2013) summarized the sequence in six stages (*Figure 4*). In the beginning, learners do not usually produce the accurate form unless they are using an unanalyzed phrase, or one

chunk of language (Brown & Larson-Hall, 2012). N. Ellis (2002) claims that a developmental sequence of language follows a route from formulae through "low-scope" or slot-and-frame patterns to constructions.

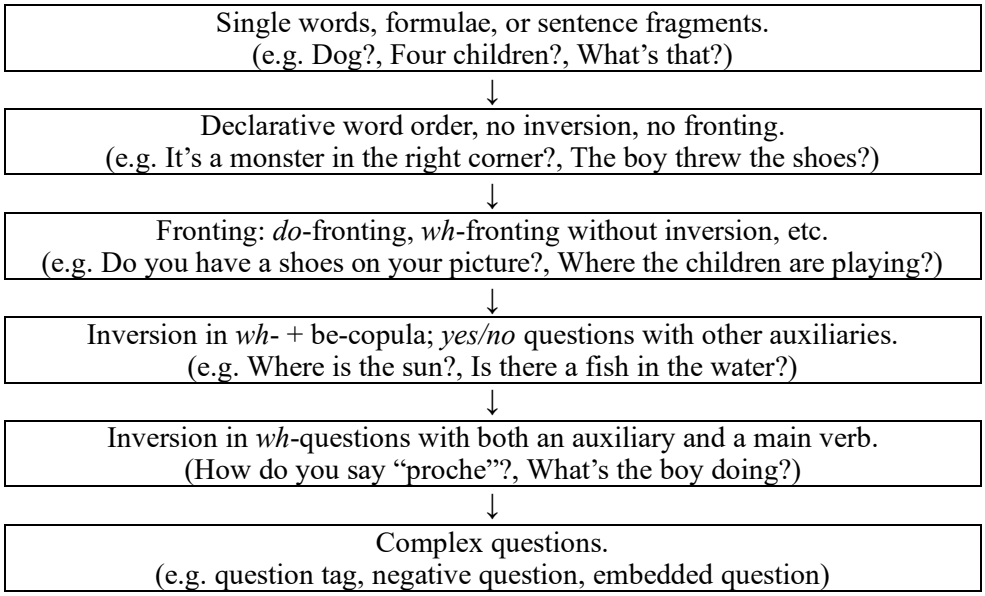


Figure 4. Developmental sequence of question

The acquisition of relative clauses was also proven to show the developmental sequence. Doughty (1991) refers to the accessibility hierarchy, which is the observed pattern of acquisition for relative clauses. Table 2 shows the hierarchy. The structures at the top are easier to access than those at the bottom.

Table 2

Accessibility hierarchy for relative clauses in English

Part of speech	**Relative clauses**
Subject	The girl who was sick went home.
Direct object	The story that I read was long.
Indirect object	The man who[m] Susan gave the present to was happy.
Object of preposition	I found the book that John was talking about.
Possessive	I know the woman whose father is visiting.
Object of comparison	The person that Susan is taller than is Mary.

When looking at the structural connections between main clauses and relative clauses, they are categorized in *Figure 5*, in the order of their frequencies (Yule, 1998). The O-S combination at the top is the most frequent, and the S-O at the bottom is the least common. According to Yule (1998), the relative clause in final position, which is by far the most common, is used to introduce new information. In medial position, the relative clause is used to make a connection with given information, modifying the subject. Yule (1998) also explains that second language learners are generally much more successful at learning to use the relative clauses in final position (in the O-S and O-O structures) than those in medial position (in the S-S and S-O structures). Some other researchers support this idea, especially concerning ESL learners. Celce-Murcia and Larsen-Freeman (1999) suggest that the O-S type is the easiest, and the O-O and S-S type follow, and the S-O is the most difficult, based on the difference of frequencies in language production found in several studies, such as Schumann (1980) and Wong (1991). Celce-Murcia and Larsen-Freeman (1999) also mention that it is assumed that Japanese learners have some difficulty in producing or comprehending English relative clauses intrinsically. They point out the reason that English relative clauses follow the head noun, or antecedent, while in Japanese, as well as Chinese and Korean, the relative clauses occur before the head noun. The present study will also scrutinize the data for the accessibility hierarchy and positions of relative clauses.

Relative clause in final position
O-S *I met **a man who** has a small cat.*
O-O *I also met **the woman that** he wants to marry.*
Relative clause in medial position
S-S ***The man who** has the small cat likes the woman.*
S-O ***The woman that** the man likes has a large do*

*excerpted from Yule (1998)

Figure 5. Positions of relative clauses

One of the common methods for identifying these developmental patterns of second language acquisition is "obligatory occasion analysis". The important procedure of the analysis is to identify obligatory occasions for the use of particular language features in samples of naturally occurring learner language production. Then the percentage of accurate use of each feature is calculated by judging whether the feature has been supplied in all the occasions where it is required to be produced. In obligatory occasion analysis, the level of acquisition is usually set at 80-90 percent of accuracy, so 10-20 percent of overuse or misuse can be overlooked. Another common method is known as "target-like use analysis". As the term "target-like" indicates, the analysis defines acquisition as mastering not only when to use a particular feature but also when not to use it (Ellis, 2008).

2.4 General Patterns of Grammatical Development of Learner Language

Some studies have been taking a broader view of the grammatical development of learner language. Pienemann and the other researchers found a clear developmental pattern of the acquisition of German word order rules by speakers of other languages

(Pienemann et al., 1988). Later, the cognitive processing operations underlying the rules were incorporated into Processability Theory.

Processability Theory is a theory of second language development in general. The theory utilizes Lexical Functional Grammar (LFG) (Bresnan, 2016) in order to explain why the utterance of a particular linguistic structure is possible. LFG encodes syntactic properties primarily in the lexicon, which makes it particularly suitable for the study of developing learner grammars, because LFG affords a formal account of linguistic dynamics present in developing learner grammars (Pienemann & Keßler, 2012). Processability Theory claims that at any stage of development, the learner can produce and comprehend only those second language linguistic forms that the current state of the language processor can handle (Pienemann, 1998, 2003, 2005, 2007, 2008; Pienemann & Keßler, 2012). This means that a certain grammatical item that has been learned by or exposed to learners should not emerge in production or serve in comprehension if their language processors have not been developed enough and ready for activation. In other words, L2 learners can produce only those linguistic forms for which they have acquired the necessary processing procedures, and learners cannot produce structures that are beyond their level of processing, because output is constrained by processability. The core of this theory is a universal processability hierarchy, which is based on the notion of grammatical information within and between the phrases of a sentence. The hierarchy is a developmental sequence of the competence at matching grammatical information to assemble phrases or sentences (Pienemann, 2007). The original processability hierarchy shown in Pienemann (1998) is as follows:

1. No procedure (e.g. producing a simple word such as *yes*).
2. Category procedure (e.g. adding a past-tense morpheme to a verb).
3. Noun phrase procedure (e.g. matching plurality as in "*two kids*").

4. Verb phrase procedure (e.g. moving an adverb out of the verb phrase to the front of a sentence “I went yesterday / Yesterday I went.”).
5. Sentence procedure (e.g. subject-verb agreement).
6. Subordinate clause procedure (e.g. use of subjunctive in subordinate clause triggered by information in a main clause.)

This hierarchy is implicationally ordered, meaning that every procedure is a necessary prerequisite for the next procedure, and the hierarchy reflects the time-course of language generation. Observed stages of development are a direct result of the stage of processing in which learners find themselves. The learners cannot skip a stage in the developmental sequence when they are not developmentally ready to learn. Thus, instruction cannot change the natural developmental course, although learners may happen to be able to produce more advanced forms on tests, given just after instruction, for example, or in very restricted pedagogical exercise in class (Pienemann, 1998).

The research based on Processability Theory tried to confirm whether the learners can manipulate a particular phrase or sentence structure using different words or morphemes, in order to exclude the use of formulae and chunks and ascertain the systematic use of linguistic features. For that, multiple opportunities to use a particular linguistic feature need to be provided for the learners intentionally. Those opportunities are called “obligatory contexts”, and the researchers examine the suppliance of a feature in its obligatory context (Pienemann, 2007). The rate of suppliance that decides “acquisition” in the obligatory context depends on the studies or researchers. The number of obligatory contexts was five in Andersen (1978); four in Meisel et al. (1981), Pienemann (1998), and Zhang (2004); three in Zhang (2005) and Rahkonen & Håkansson (2008). There are some studies that require the existence of two examples with different

lexical and different morphological forms (Mansouri, 2005; Dyson, 2009, 2010). On the other hand, some studies examine only the appearance of the very first, single example (Hammarberg, 1996; Glahn et al., 2001). Zhang (2004) points out that many acquisition criteria that have been used in the history of language acquisition research are accuracy-based (e.g. comparing L2 performance against the target language norm). In contrast, Processability Theory adopts the emergence criterion, aiming at identifying the point at which a grammatical item makes its first systematic and productive appearance in an L2. Pallotti (2007) claims that emergence is understood as referring to unplanned and unmonitored use of the structure, which most likely involves implicit knowledge, and is the minimum amount of evidence necessary to state that a structure is there. Likewise, Baten et al. (2015) argue that only when the procedural skills are available in the language processor will the corresponding linguistic structures appear, evidently, to a minimal extent at first.

Most of the studies drawing on the Processability Theory has primarily dealt with spoken data, and very few studies have analyzed written data. However, Rahkonen and Håkansson (2008) support the application of this theory to writing. They claim that, as in the case of speech, learners' attention directed to the linguistic forms should be similarly limited in written production while they try to create content.

Partly inspired by the theory, the present study will examine formulaicness of the use of some grammatical items found in the data.

2.5 Definition of "Acquisition" and "Development" of Grammatical Features in the Present Study

Having reviewed some representative studies on developmental patterns of second

language acquisition, the present study takes its own view as to the acquisition and development of grammatical features that appeared in the students' writings. In this study, at least one-time correct use of a particular grammatical feature is dealt with as its "acquisition", even if there is another example of the same feature that is erroneous. Thus, multiple obligatory occasions or contexts where a particular grammatical feature is used are not taken into consideration, and target-like use of it is not really confirmed, either, because the misuse or overuse of an item is overlooked if at least one example of its successful use is found. Also, the present study defines "development" of grammatical features as the increase of "acquired" grammatical items emerging in the students' writing that are taught as listed target grammar in the textbooks. Therefore, in some individual cases, fully stable use of a particular grammatical item might not be guaranteed. In addition, there would be a gap of proficiency among the individuals. However, it should be worth examining how much of the instructed grammatical items the group of learners have learned enough to manage to use, because that might depict the whole picture of transition from learned grammar to manipulatable grammar, and that would also indicate how successful the classroom instruction was in terms of developing productive skills of English. As we have observed, the analyses of many studies on acquisition or development of linguistic features tend to be narrow-scoped, and see the staged development of interlanguages at the micro level or in a fine-grained manner. However, the principles of developmental patterns are not easily introduced into pedagogical syllabi, at least because only a small number of linguistic features have been investigated in terms of development (Lightbown & Spada, 2013). Many school teachers would be more interested in how much of all the target grammar they taught has been acquired and used by their students, than the developmental status of an individual piece of linguistic feature. Therefore, in this study, "acquisition" or "development" is defined as how much of

students' grammatical knowledge has become manipulatable at a specific point of learning, that is each academic year, even if each piece of grammatical item is used only one time. The following three studies adopted the same definition of acquisition or development.

2.6 Examining the Emergence of Grammatical Features in Japanese EFL Writing

With regard to examining the emergence of grammatical features in Japanese EFL learners' language production, Negishi (2012b) examined the emergence of the CEFR-labeled grammatical items in university students' compositions. He gave the participants (900 university students) three types of writing tasks: an email, a story, and an argumentative essay. Their writing was scored based on the checklist of 53 CEFR-labeled grammatical items, drawn from English Profile Programme research, with some necessary modifications. The study intended to ascertain the usefulness of the checklist as an alternative device for assessing EFL learners' writing, for example, in place of the traditionally used error-counting method. The participants were expected to use different grammatical items in the different writing tasks, but were not intentionally provided multiple obligatory contexts where any particular items were likely to appear. At least one-time correct emergence of a certain grammatical item was interpreted as an evidence that the learner can use it for him/herself, even if a non-target-like use of it was also found. The emergence of each item was dealt with as though it was a correct answer of a test, and "1" was put into the checklist. No emergence or incorrect use of an item was dealt with as an incorrect answer, and "0" was put into the checklist. The 1/0 table was used for the Rasch model analysis, and the item difficulties (the difficulties of the grammatical items for the learners to use in this case) and the person abilities were calculated in order

to examine the discriminative potential of items and the correlation between the person abilities and the human holistic ratings. The results showed that 42 linguistic features were effective in discriminating this level of English learners and that the participants' abilities correlate moderately with the human ratings, ranging from CEFR A2 to $B2^{+}$.

Following Negishi (2012b), Murakoshi (2012b, 2013) examined the compositions written by Japanese public upper secondary school students, using a checklist that comprises of the CEFR-labeled grammatical items from the EPP research and those enumerated in the Course of Study for lower secondary school. Thus, the list included a few morphological items such as third person singular -s and the past tense of a verb (-ed), since they are itemized as target grammar in the authorized textbooks edited in accordance with the Course of Study.

In Murakoshi (2012b), I made the checklist by combining the one used in Negishi (2012b) with the list of target grammatical items in the government-authorized textbooks for three years of lower secondary school. The number of the items added up to 116. Then, I examined the use of each listed item in the compositions written by the Japanese first-year upper secondary school students. Using the 1/0 checklist, I conducted the Rasch model analysis, and came up with the tentative conclusion that students who were at the beginning of the first year of upper secondary school, were presumed to be able to use only 24 basic grammatical items at their will. These were the grammatical items they had learned in the first and the second year of lower secondary school. However, I had some concerns about this result. The writing task that the students worked on was a narrative ("The Key Person in My Life"), so there might have been some grammatical items that were unlikely to emerge, such as interrogative or imperative sentences. Another concern is that the checklist for this study contained not a few grammatical items that the students had not learned at the time of the research. Naturally, none of the students could use those

items correctly. As a result, this drawback might have affected the calculation of the item difficulties in the Rasch model analysis.

In Murakoshi (2013), I conducted an investigation on the development of the grammatical features of the same students during the three years by using the Rasch model analysis, following Murakoshi (2012b). For this study, I developed a new checklist that comprised of 87 items, with some minute classifications put together. I found that the Rasch-extracted manipulatable items increased in a relatively accumulative manner year by year, and that it was not until the third year of upper secondary school that the grammatical items they learned in the third year of lower secondary school emerged in their compositions as manipulatable linguistic knowledge. However, I had two concerns about this study. One was, again, that the checklist still contained some grammatical items that they had never learned that potentially could have affected the Rasch calculation of the item difficulties. The other concern was that the topic of each writing task given to the students in each academic year was different ("The Key Person in My Life", "My School Life of Last Year", "Should High School Students Have a Part-time Job?") and that might have affected the emergence of grammatical items in their compositions. In other words, the choice of topic could have influenced the types of grammatical items that the students chose to use. Although these two studies have a few defects, it could be said that there might be a huge gap between the language exposed to the students as input and the language used in their output. The grammatical items that were more difficult would have needed to wait a longer time to be used at the students' will after they had been instructed in the classrooms. In contrast, all the grammatical items that emerged in their production and were judged as being manipulatable, were easier to learn or earlier-instructed and presumably were processable in their language systems.

2.7 Declarative and Procedural Knowledge of Language Learners

The difference between declarative and procedural linguistic knowledge is one of the distinctions that is often a focus of the SLA research. Declarative memory underlies the facts learned and stored, and the stored information can be retrieved when necessary. Thus, declarative knowledge can be defined as knowledge *about* something. Procedural knowledge, on the other hand, is concerned with motor and cognitive skills that involve sequencing information, such as producing language. While declarative knowledge can be retrieved when called upon, procedural knowledge is rather inaccessible (Gass & Selinker, 2008). In other words, declarative knowledge is knowledge THAT something is, and procedural knowledge is knowledge HOW to do something (DeKeyser, 2009).

Another well-known distinction is the difference between explicit and implicit knowledge. Explicit knowledge is knowledge that one is aware of, that one has conscious access to. Implicit knowledge is outside awareness, and therefore cannot be verbalized, only inferred indirectly from behavior (DeKeyser, 2009, p.121). In the eyes of most researchers, second language acquisition is equated with the development of implicit knowledge of the language (R. Ellis, 2006). This distinction is often regarded in the same light as the declarative-procedural distinction. However, DeKeyser (2009) claims that explicit is not exactly the same as declarative, and implicit not exactly the same as procedural. Considering linguistic competence in the Chomskyan sense, for example, declarative knowledge is not necessarily explicit because it is not necessarily accessible to awareness. On the other hand, procedural knowledge is not necessarily implicit, because it can be the result of proceduralization and (partial) automatization of declarative knowledge, and still allow or require a certain degree of conscious access when being used (p.121). Therefore, he maintains that the declarative-procedural-automatized

distinction is the most useful from a developmental perspective (p.126). R. Ellis (2005a) argues that we cannot be sure that learners do not access at least some explicit knowledge, especially when leaners are working on writing task—one of the spontaneous production tasks, which he thinks are probably the best means eliciting implicit knowledge.

Many researchers have been interested in and investigating the role of practice to help convert declarative knowledge to procedural knowledge and then to automatization (Lightbown & Spada, 2013). N. Ellis (2011) explains that particularly in initial stages, controlled use of declarative knowledge, with explicit knowledge coaching and practice, guides the proceduralization and eventual automatization of language processing. DeKeyser (2007) mentioned that the process of proceduralization and automatization can be advanced if the declarative knowledge and a task that requires the knowledge are available to the learners, and that the leaners need to experience a large amount of practice in order for the process to proceed. Similarly, R. Ellis (2005b) claims that declarative grammatical knowledge can support practice of a behavior long enough so that procedural memory can begin to take over with enough repetition. For practice in the foreign language classroom, Ortega (2007) has proposed three principles: practice should be interactive, practice should be meaningful, and there should be a focus on task-essential forms. Also, Lightbown and Spada (2013) propose that teachers should be more aware of language features that are just beginning to emerge in the second language development of their students and provide some guided instruction in the use of these forms (p.197). That proposal is based on the assumption that students may produce certain structures after they have been taught them in class, but cease to use them later because they are not fully integrated into their interlanguage systems (p.177).

Provided that we can apply the researchers' views as to declarative and procedural knowledge shown above to the present study, it may be reasonable to define grammatical

knowledge instructed in class as declarative, and grammatical items correctly used by the subjects in their compositions as proceduralized knowledge. In Chapter 3, prior to the results section, how the students had experienced the instructions of declarative grammatical knowledge will be reviewed. In Chapter 5, it will be discussed how teachers should provide their students with the instruction and practice in class in order for the proceduralization of grammatical knowledge to advance.

Chapter 3. The Present Study

In order to examine more rigidly the Japanese upper secondary school students' mastery of manipulatable grammatical knowledge and the gap between the language exposed to them as input and the language used in their output, I decided to carry out a similar study to Murakoshi (2012b, 2013) by an improved method. Also, I tried to examine the development of the syntactic complexity of learners' sentences more deeply, following Murakoshi (2015). Accordingly, I set up the following research questions:

1. When checking the use of grammatical items that are labeled as A2 to B1, or learned in lower secondary school, how much of those items can be regarded as "being used at will" in particular writing tasks given to the Japanese upper secondary school students?
2. How is the syntactic complexity of sentences (MLU) related to the fluency of the students' compositions and their overall writing proficiency?

3.1 Participants

The participants of this study were 209 Japanese EFL learners at a public upper secondary school. They all had finished three-year English courses at lower secondary school (approximately, 300 hours), and continued to learn English in upper secondary school for another three years, under the curricula based on the former Course of Study (enforced in 2003). It was assumed that many of them had had limited opportunities of writing at least paragraph-level free compositions partly because they did not need to write compositions in their entrance examinations for upper secondary school. In the first year of upper secondary school, they studied "English I", a 140-hour integrated skills

English course, and “Oral Communication I”, a 70-hour course focusing on speaking and listening. In English I class, they learned grammatical items targeted in each unit of the textbook, most of which were the ones they had already learned in lower secondary school (e.g. present tense, past tense, future expressions, prepositional phrase, comparison, modal auxiliary verbs, S+V+C=that-clause, S+V+O=that-clause, subordinations: if; when; because, passive, to-infinitives, nominal modification with *-ing*/*-ed*, gerund, present perfect, indirect WH-question, WH-+to-infinitive, S+*tell/want*+O+to-infinitive, relative pronoun: except whose, it is – [for –] to …). Only a few grammatical items were newly taught in the latter of academic year (e.g. S+V+O+that-clause, S+V+O+WH-clause, it is – that …, relative adverb, present perfect continuous, adverbial clause with *-ing*, nominal relative: what, modal+passive, past subjunctive). There, they also began to learn how to write a structured English passage, from a paragraph to a short essay. They wrote paragraphs or short essays as a formal assignment five or six times in a year. In addition, they were required to keep a journal in English at least three days a week and submit it once a week to the assistant language teachers (ALTs) for enhancing writing fluency and developing rapport with them. The ALTs, both of whom were from the United States of America, were asked to give feedback (comment) mainly on the content, not much on the use of grammar or vocabulary so that the students could not be demotivated to write by too much error correction. In the second year, all of them studied “English II”, a sequel to English I, for about 140 hours, and “International Understanding”, a content-based language course, for about 70 hours a year. In English II, they learned some more new grammatical items targeted in each textbook unit (e.g. bare infinitive, S+V+O+participle, non-defining relative clause, past perfect, past perfect continuous, past perfect subjunctive, perfect modal, relative pronoun: whose, perfect infinitive). They also continued to practice writing short essays in English II class and worked on five or

six essay writings in a year as a formal assignment. Some of the students took “Oral Communication II”, an elective course (approximately, 100 hours) in addition to the required courses. In the third year, the students had to take “Reading” and “Writing” courses for 140 and 100 hours a year, respectively. In the writing class, they learned how to build an argument in English as preparation for debating. Also, in the reading class, they wrote a few argumentative essays as an assignment, based on the content of textbook passage they read. Some of them took an additional English course as an elective focusing on preparation for entrance examinations for university as well (approximately, 70 hours).

In terms of proficiency in English, they were considered as intermediate EFL learners in upper secondary school. Their average total scores on the GTEC for STUDENTS®[5], an English proficiency test for Japanese lower and upper secondary school students, went from 383.3 in the first year (Reading 145.0, Listening 150.9, Writing 87.4), to 417.0 in the second year (Reading 151.8, Listening 161.8, Writing 103.4), to 435.2 in their third year (Reading 161.8, Listening 163.5, Writing 109.9). The test was out of 660 points. The national average score of 2011-13 was 393.4. Thus, it may be reasonable to say that they were ranked in the medium level, at least among all test takers in the country. Their school was designated as a pilot school on a national project, called “Super English Language High School (SELHi)”, commissioned and funded by the Ministry of Education, Culture, Sports, Science and Technology (MEXT). The students all agreed on the use of their anonymous data for the SELHi project and further research.

[5]GTEC for STUDENTS® is an English test, administered by Benesse Corporation, that measures the three skills of listening, reading and writing of lower and upper secondary school students.

3.2 Data Elicitation

I decided to analyze the same set of writing tasks used in Murakoshi (2012b) and Murakoshi (2013) again but in a more sophisticated way so that I could make up for the shortcomings of those studies. This longitudinal data set consists of three writing tasks written by the 209 Japanese upper secondary school learners of English mentioned above, one for each year. The students were asked to write an essay in English at the beginning of each academic year in order to measure the effectiveness of teaching and learning of writing skills for the SELHi project of the school. In the first year, just after entering the school the students wrote an essay on the topic "The Key Person in My Life" (Composition A). In the second year, the topic was "My School Life of Last Year" (Composition B). In the third year, the topic was "Should High School Students Have a Part-time Job?" (Composition C). Each of them was a 40-minute timed essay. The students had to plan and handwrite their drafts as a form of a writing test within that time. The average numbers of words in these compositions are 94.7 in the first year, 143.4 in the second year, and 112.7 in the third year. All their handwritten drafts were converted into electronic data by manual typing and saved as text files (.txt.), which enabled quicker and more accurate detection of grammatical items and some other analyses. Errors, including misspelled words, were typed as they were, and contracted forms were separated with space (e.g. can't => can /'t).

As the data showing the students' overall writing proficiency, the score of writing section of GTEC for STUDENTS test was adopted. The GTEC writing test that they took in those days required them to write their opinion on a given topic in 20 minutes and was graded out of 160 points. The criteria of assessment were opinion, reason, vocabulary, grammar, and organization. The students took the test at almost the same time as the

composition test in three academic years. Even though the students took the standardized private writing test, I felt more interest in analyzing the composition that they wrote as a classroom activity incorporated within the teaching syllabus. However, the compositions were assessed with a three-grade rubric. The assessment points were fluency (range of word counts), organization (essay structure: introduction-body-conclusion), and accuracy (range of percentage of error-free T-unit). Since the scores were given in three grades (A, B, C), it seemed unsuitable to compare them with the MLU values, even if the grades were converted to ordinal scales (e.g. 3, 2, 1). That is why I decided to adopt their GTEC scores as the index of overall writing proficiency to be compared with their MLU values.

3.3 Method

3.3.1 The Rasch Model Analyses of the Emergence of Learned Grammatical Items Examined with the Revised Checklist

In Murakoshi (2012b, 2013), I used a checklist which comprised 87 grammatical items in total. Those were all the items labeled with the CEFR levels in Hawkins and Filipović (2012) and some other items taught in lower secondary school. The items listed in Table 3 are the "target" grammatical items in government-authorized textbooks for lower secondary school. As already mentioned, they include from morpheme-level items to sentence structures. Some of them are also listed as the CEFR-labeled criterial features in Table 1. In Table 3, those CEFR-labeled items are accompanied by their description as the criterial features and their levels (some of them are listed in the same way as the criterial features). Also, the figures in the column "LSS" show the academic years of lower secondary school when those items were taught in class. Some items were taught in different years in different textbooks, so the coverage in four or more books out of six

decided the academic years when they were taught.

Table 3

The "target" grammatical items taught in lower secondary school

Item	LSS	Item	LSS
S+V (Intransitive clauses: A2)	1 yr.	could/would you -?	2 yr.
S+V+O (Transitive clauses: A2)	1 yr.	will you -?	2 yr.
be-copula (present)	1 yr.	shall I -?	2 yr.
verb (base form)	1 yr.	infinitive (purpose)	2 yr.
direct WH-question (direct WH-questions: A2)	1 yr.	infinitive (obj.) (verbs with an infinitival complement: A2)	2 yr.
verb (past)	1 yr.	infinitive (complement)	2 yr.
S+*be*+C(adj.)	1 yr.	infinitive (adj.)	2 yr.
verb (3rd person singular)	1 yr.	infinitive (reason of feeling)	2 yr.
personal pronoun (as obj.)	1 yr.	if (conj.: adverbial)	2 yr.
present progressive	1 yr.	when (conj.)	2 yr.
can (ability)	1 yr.	because (conj.)	2 yr.
S+V+O+O (ditransitive clauses: A2)	2 yr.	comparative	2 yr.
S+V+that- (Verbs with a finite complement clause: A2)	2 yr.	superlative	2 yr.
something+to- (pronoun plus infinitive: A2)	2 yr.	as – as	2 yr.
may/can/might (possibility) (modal auxiliary verbs: may, might, can: A2)	2 yr.	passive	3 yr.
must (obligation) (modal auxiliary verbs: must: A2)	2 yr.	NP+-*ed* (Postnominal modification with –*ed*: A2)	3 yr.
should (advice) (modal auxiliary verbs: should: A2)	2 yr.	NP+-*ing* (Postnominal modification with –*ing*: B1)	3 yr.
may (permission) (modal auxiliary verbs: may: B1)	2 yr.	it … that (*It* extraposition with finite clauses: B1)	3 yr.
must (necessity) (modal auxiliary verbs: must: B1)	2 yr.	indirect WH-question (indirect WH-questions in finite clauses: B1)	3 yr.
be-copula (past)	2 yr.	WH-+to- (indirect WH-questions in infinite phrases: B1)	3 yr.
S+*look*/*feel*/*become*+C(adj.)	2 yr.	want/like/expect+O+to- (subject-to-object raising verbs (unpassivised): B1)	3 yr.
past progressive	2 yr.	it … to- (*It* extraposition with infinitival phrases: B2)	3 yr.
there is/are	2 yr.	S+V+O+C(adj.) (Secondary predications: B2)	3 yr.
can (request)	2 yr.	it is … for- to-	3 yr.

can (permission)	2 yr.	gerund (obj.)	3 yr.
be going to-	2 yr.	gerund (subj.)	3 yr.
will	2 yr.	present perfect	3 yr.
have to-	2 yr.	relative pronoun (except whose)	3 yr.

In general, these grammatical items are taught rather explicitly: teachers explain the items usually in their first language (Japanese), and students deductively learn them. Their inductive learning outside the classroom cannot be expected so much.

In Murakoshi (2012b, 2013), not surprisingly, the items ranked in the CEFR B2 or higher levels were almost never used in the students' writings in all the academic years, because the students had never learned those items and didn't know how to use them. My concern was that the absence of users of those items could have caused inaccurate calculation of their item difficulties. Also, as I mentioned above, the MEXT's goal of English proficiency of upper secondary school students is the CEFR B1 level. Therefore, in the present study, I decided to eliminate those items from the checklist in order to calculate the item difficulties and the learners' abilities more accurately in the following statistical analyses, which would lead to investigating the trajectory of the students' acquisition of the learned grammatical items more precisely. The item eliminated from the list were as follows:

- Adverbial subordinate clauses with *-ing* (preceding main clause) <B2>

  e.g. *Talking about spare time, I think we could go to the Art Museum.*

- Pseudoclefts type(ii): WH-VP <B2>

  e.g. *What fascinated me was that I was able to lie on the sea surface and read this newspaper.*

- Verbs with an NP plus finite complement clause <B2>

  e.g. *She told me that she had worked for summer camp for children.*

- Subject-to-Subject Raising verbs and adjectives: appear, cease, fail, happen, prove, turn out, certain, likely, sure, unlikely <B2>

  e.g. *My worries proved to be wrong.*

  Whenever money is involved, some problems are likely to happen.

- Subject-to-Object Raising verbs (unpassivised): imagine, prefer <B2>

  e.g. *I would prefer my accommodation to be in log cabins.*

- Subject-to-Object Raising verbs plus passive: expected, known, obliged, thought <B2>

  e.g. *How many hours a day should I be expected to work?*

- Tough Movement: difficult, good, hard <B2>

  e.g. *The grammar and vocabulary are a bit hard to learn.*

- Double embeddings: [of] –'s <B2>

  e.g. *After this I went to a friend of mine's house where I spent one week.*

- Subject-to-Subject Raising verbs and adjectives: chance <C1>

  e.g. *I chanced to know your Competition from an international magazine.*

- Subject-to-Object Raising verbs (unpassivised): believe, find, suppose, take <C1>

  e.g. *I find this to be more interesting than the walking route to Lake Hawksmere.*

- Subject-to-Object Raising verbs plus passive: assumed, discovered, felt, found, proved <C1> e.g. *Internet is a valuable tool which can be proved to be the most important aspect in the learning process.*

- Double embeddings: [–'s] –'s <C1>

  e.g. *After spending the first day of their marriage in the bride's family's house…*

- Modal auxiliary verbs: might (permission) <C1>

  e.g. *Might I tell you what we discuss?*

- Subject-to-Object Raising verbs (unpassivised): declare, presume, remember <C2>

  e.g. *They declare some products to be the hits of the season.*

- Subject-to-Object Raising verbs plus passive: presumed <C2>

  e.g. *Not only meetings with people are presumed to give new experiences.*

- Tough Movement: tough <C2>

  e.g. *What she knew would be really tough to live with was the reason of his death.*

In addition to those above, "*It* Extraposition with infinitival phrases" and "Secondary predication" (S+V+O+C=adjective) are also labeled as the B2 features in Hawkins and Filipović (2012). However, they were explicitly taught as target grammar in Japanese lower secondary schools, so I decided to keep them in the revised checklist. That left me with a total of 71 items to check (Table 4).

Table 4

Revised checklist of the grammatical items

1	S+V	25	S is easy to-	49	will
2	S+V+O	26	of [-'s]-	50	have to-
3	S+V+O+O	27	may (permission)	51	could/would you -?
4	S+V+that	28	must (necessity)	52	will you -?
5	direct WH-question	29	should (probability)	53	shall I -?
6	something+to-	30	it … to-**: B2**	54	infinitive (purpose)
7	NP+-*ed*	31	S+V+O+C(adj.)**: B2**	55	infinitive (obj.)
8	NP+-*ing*	32	*be*-copula (present)	56	infinitive (complement)
9	double embedding: of-[of-]	33	verb (base form)	57	infinitive (adj.)
10	may/can/might (possibility)	34	*be*-copula (past)	58	infinitive (reason of feeling)
11	must (obligation)	35	verb (past)	59	it is … for- to-

12	should (advice)	36	S+*be*+C(adj.)	60	if (conj.: adverbial)
13	S+V+O+to-	37	S+*look/feel/become*+C(adj.)	61	when (conj.)
14	S+V+O+-*ing*	38	verb (3rd person singular)	62	because (conj.)
15	it … that	39	personal pronoun (obj.)	63	gerund (obj.)
16	S+V+to-+that	40	present progressive	64	gerund (subj.)
17	whose (relative pronoun)	41	past progressive	65	comparative
18	pseudocleft (what+S+V)	42	S+V+O+C(noun)	66	superlative
19	indirect WH-question	43	tag question	67	as – as
20	WH-+to-	44	there is/are	68	passive
21	would rather/had better	45	can (ability)	69	present perfect
22	adverbial clause with -*ing* (following main clause)	46	can (request)	70	relative pronoun (except whose)
23	seem/be supposed+to-	47	can (permission)	71	prepositional phrase (adj.)
24	want/like/expect+O+to-	48	be going to-		

The items "S+V+O+to-" (No.13) and "want/like/expect+O+to-" (No.24) are seemingly similar, but they are different. The structure "S+V+O+to-" is explained as "verbs with an infinitival complement (object control)" (e.g. *I ordered him | to gather my men to the hall.*), while "want/like/expect+O+to-" has the structure "subject-to-object raising verbs" (e.g. *I want <u>you</u> to [∅] help me.*). Another note is that the item "passive" does not cover "adjectival passives". (e.g. *I was surprised at the news.* / *He is interested in Japanese castles*.). These examples are regarded as "S+*be*+C (adjective)". Although there have been many studies about the developmental sequence of question sentence (e.g. Pienemann et al., 1988; Pienemann, 1998), the structures of *yes/no* question (e.g. Is this …? / Do you …?) were decided not to be in the list independently, because they are taught with "*be*-copula" or "verb" as a package and are not CEFR-labeled in Hawkins and Filipović (2012). For the same reason, the negative structures were also decided not to be

included in the checklist. As for relative pronoun, its developmental stages have also been investigated and its different use has been proven to be developmentally hierarchical as shown in Doughty (1991). However, in Murakoshi (2012b, 2013), the number of users of any relative pronoun was not large enough to be divided depending on the usage, so I decided to keep that list item as it was ("relative pronoun" together).

Checking the use of grammatical items was done basically by visual inspection (with eyes), but beforehand, rough extraction of the items was conducted by the computer software "AntConc". AntConc is a corpus analysis toolkit for concordancing and text analysis, developed by Dr. Lawrence Anthony (*Figure 6*). Misspelling of the words was overlooked to some extent, but the items in incomplete sentences (e.g. missing subject, critical word disorder, unintelligible sequence of words) were excluded from analyses and not counted as being used. Therefore, in the first step of item extraction by AntConc, some possible candidates of overlooked misspelling (observed and listed while typing) were added to the correct form in the search window, as seen at the bottom of *Figure 5* ("because | becouse | becose | because"). Only one correct example of the item in the whole passage determined that the item was manipulatable for its user, even if there was an incorrect example of the same item in another sentence.

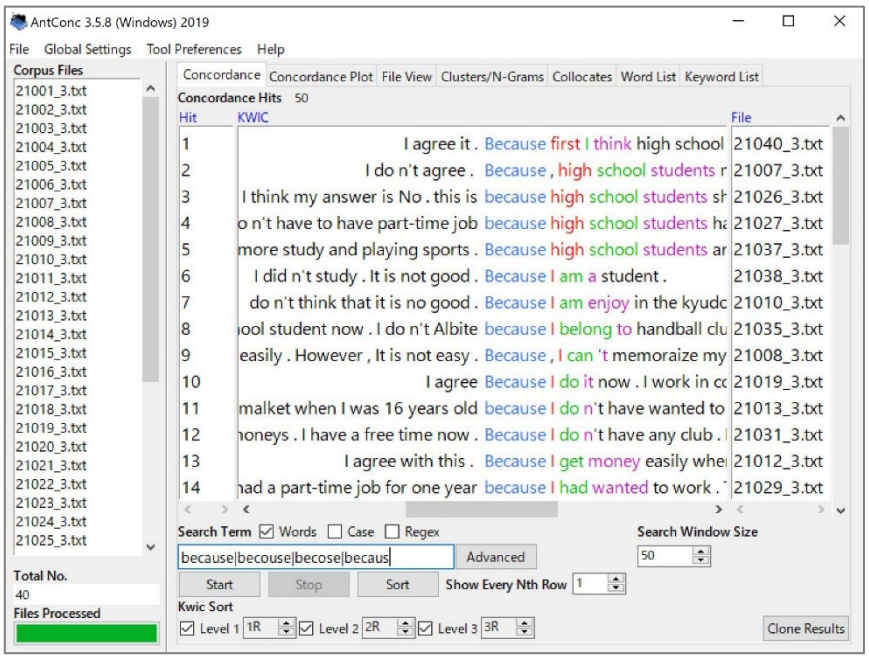


Figure 6. Sample screen of analysis using AntConc

Using the revised checklist with all the data input, I conducted the same Rasch analyses that I did in Murakoshi (2012b, 2013) on the 1/0 dichotomous data of the emergence of the listed grammatical items (emergence: 1, no emergence or incorrect use: 0) elicited from the longitudinal data of English compositions written by the 209 students in three years.

The Rasch model is a unidimensional measurement model that deals with dichotomous data; that is, examinees' positive reactions (correct answers) or negative reactions (incorrect answers) to test items, in order to calculate the relationship between the ability of examinees and the difficulty of test items (Apple, 2013; Yu, 2013). This model is often regarded as the simplest variety of the Item Response Theory (IRT) models. However, the Rasch model has developed into a unique tool (Toyoda, 2002), and should not be regarded as an IRT model, in that it tries to fit the data into a model for obtaining more objective results, while other IRT models try to search for models that can best

explain the data (Shizuka, 2007).

The reasons the Rasch model is preferred for use in different disciplines are that its low number of parameters makes an estimation easier (Toyoda, 2002), that the interpretation of the result is easier (Ohtomo, 1996), and that it does not require normal distribution of the ability of examinees as an essential prerequisite (Shizuka, 2007). The Rasch model can deal with smaller numbers of examinees, compared to two or three-parameter IRT models (Table 5). Therefore, in terms of desirable sample size too, it seemed appropriate to adopt the Rasch model analysis here since there are 209 subjects in the present study.

Table 5

Desirable numbers of subjects in the Rasch and IRT model analyses

	Rasch	2-parameter IRT	3-parameter IRT
Ohtomo, 1996	100-200	200-400	1000-2000
Toyoda, 2002	100 and over	300 and over	1000 and over

The general expression of the Rasch model is as follows (Shizuka, 2007):

$$\boldsymbol{P}\left(\boldsymbol{X_{ni}} = \boldsymbol{x_{ni}} \mid \boldsymbol{\theta_n}, \boldsymbol{\delta_i}\right) = \frac{\boldsymbol{exp}[\boldsymbol{x_{ni}}(\boldsymbol{\theta_n} - \boldsymbol{\delta_i})]}{\boldsymbol{1} + \boldsymbol{exp}(\boldsymbol{\theta_n} - \boldsymbol{\delta_i})} \qquad (\boldsymbol{1})$$

θ_n is person ability, δ_i stands for item difficulty, and $exp(\theta_n - \delta_i)$ means raising e (natural logarithm: $\fallingdotseq 2.718$) to the power of $(\theta_n - \delta_i)$. If $X_{ni} = 1$, which means a correct answer (in this study, at least one-time correct use of a particular grammatical item), its probability is calculated by the following formula:

$$\boldsymbol{P}\left(\boldsymbol{X_{ni}} = \boldsymbol{1} \mid \boldsymbol{\theta_n}, \boldsymbol{\delta_i}\right) = \frac{\boldsymbol{exp}(\boldsymbol{\theta_n} - \boldsymbol{\delta_i})}{\boldsymbol{1} + \boldsymbol{exp}(\boldsymbol{\theta_n} - \boldsymbol{\delta_i})} \qquad (\boldsymbol{2})$$

If X_{ni} =0, which means an incorrect answer (in this study, absence or any incorrect use of a particular grammatical item), its probability is calculated by the following formula, transformed from (1) above:

$$\boldsymbol{P\,(X_{ni}=0\mid\theta_n,\delta_i)=\frac{1}{1+exp(\theta_n-\delta_i)}} \qquad (3)$$

The joint probability of a pattern of 1/0 data set is calculated by the product of (2) or (3) above. For example, if the response pattern { 0, 1, 1 } is given, the calculation of its joint probability goes as follows:

$$\boldsymbol{P\begin{pmatrix}X=0\\1\\1\end{pmatrix}=\frac{1}{1+exp(\theta_n-\delta_i)}\times\frac{exp(\theta_n-\delta_i)}{1+exp(\theta_n-\delta_i)}\times\frac{exp(\theta_n-\delta_i)}{1+exp(\theta_n-\delta_i)}}$$

The final stage of the Rasch calculation is to estimate the most likely $\boldsymbol{\theta_n}$ and $\boldsymbol{\delta_i}$ that maximize the joint probability. This procedure is called "maximum likelihood estimation". The Rasch analysis software called RASCAL (Assessment Systems, 1988), which was employed in this study, estimates unknown person ability and item difficulty at the same time, using the maximum joint probability, so its procedure is called "joint maximum likelihood estimation" (Ohtomo, 1996). Note that maximum likelihood estimation cannot estimate the ability of the person answering all the items correctly or all the items incorrectly (Takahashi, 2002).

Most Rasch analysis programs create an item-person map, which enables the comparison the ability of examinees and the difficulty of test items aligned on the same scale (*Figure 7*).

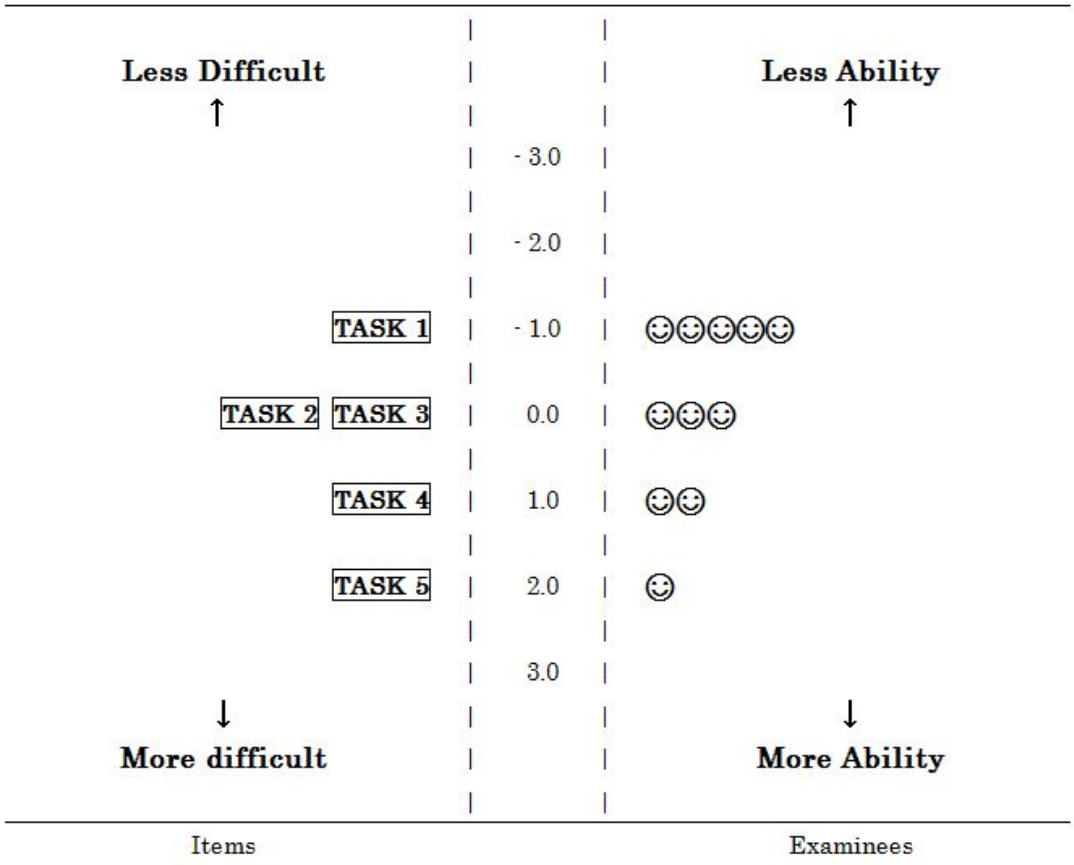


Figure 7. Item-Person Map (sample)

In *Figure 7* above, for example, three persons (☺×3) and two items (TASK2, TASK 3) are aligned at the point of 0.0. This means that those three persons can be considered to have the ability to achieve TASK 2 and TASK 3, with a probability of 50 percent. In the Rasch model, when the item difficulty and the person ability are the same, the percentage of correct answers is set at 50 percent (Baker, 1997; Bond & Fox, 2015; Ohtomo, 1996). In other words, an item of a certain amount of difficulty is one that a person having the same amount of ability can answer correctly with 50-percent success rate (Shizuka, 2007). Therefore, it can also be estimated that the three persons at the 0.0 point (or logit) have the ability to achieve the easier TASK 1 with more probability, because the order and the difference of item difficulty are the same for all the examinees

(Bond & Fox, 2015; Shizuka, 2007).

The Rasch model analysis is of great use for this study. As shown in *Figure 7*, the Item-Person Map in the Rasch model analysis aligns the difficulty of items and the ability of persons on the same scale base on the logit values. This is beneficial to identify the items within the range of learners' ability, as the manipulatable grammatical items. The manipulatable, or "acquired" grammatical items for this group of learners were defined as those that at least one person in the group correctly used on his/her own. That is, the cut-off line was drawn at the point where the person with the highest ability was located, and the matching item(s) at that point and the less difficult items were defined as the acquired for those learners with a success rate of 50 percent or more.

The issue of fit analysis, which reveals whether the data harmonizes with the model or not, will not be discussed here because the present objective is not to create a better set of test items, but to explore the relationship between the difficulty of the grammatical items and the ability of the subjects.

3.3.2 Examining the Syntactic Complexity of Students' Compositions: Its Relationship with Fluency and Overall Writing Proficiency

In Murakoshi (2015), the development of the mean length of utterance (MLU) of the student's writings was found from the first to the third academic year. However, glancing over the MLUs and the total numbers of words of the compositions, I noticed that there were some compositions in which the MLUs were relatively high but did not have so many words. In the present study, following Murakoshi (2015), first, I confirmed again that the development of the MLUs was statistically significant and examined whether the development of the total numbers of words was also statistically significant. Then, I looked into the relationship between the MLUs and the total numbers of words of

the compositions written in each year, by calculating the correlation coefficient (Pearson's product-moment correlation coefficient). Also, I investigated the relationship between the students' MLUs and their score of writing section of GTEC for STUDENTS test, also by looking into the correlation coefficient (Pearson's product-moment correlation coefficient).

3.4 Results

3.4.1 The Rasch Model Analyses of the Emergence of the Listed Grammatical Items in a Particular Writing Task in Every Academic Year

3.4.1.1 The Acquired Grammatical Items Extracted from Composition A (Narrative Essay) Written in the First Academic Year

First of all, using the data of Composition A, written in the first academic year, I examined the distributions of the item difficulties and the students' abilities from the output of the RASCAL program. While the item difficulties spread over a wide range, the students' abilities are concentrated in a narrow area of less abilities, that is, easier items (*Figure 8*). The persons with the highest ability are found at the item difficulty of -0.4, so the grammatical items whose difficulties are -0.4 or less (located at the top in the figure) can be judged as the ones that the students distinctively used in their narrative compositions in the first year. Then, using another output of RASCAL, I tried to identify what the items were. In *Figure 9*, the items with difficulties of -0.4 or less are No. 32 to No. 31, 18 items in total. They are "*be*-copula (present)", "verb (base form)", "S+V+O", "S+*be*+C (adjective)", "personal pronoun (as object)", "S+V", "verb (past)", "S+V+that-clause", "when (conjunction)", "*be*-copula (past)", "infinitive (as object)", "S+V+O+O",

"verb (3rd person singular)", "will", "can (ability)", "present progressive", "prepositional phrase (adjectival)", and "S+V+O+C (adjective)".

Meanwhile, there are nine grammatical items that showed the highest difficulty, 2.696 (*Figure 9*). They were "indirect WH-question", "seem/be supposed+to-", "as – as", "NP+-*ing*", "should (advice)", "it … that", "NP+-*ed*", "can (permission)", and "infinitive (complement)". However, the fact is that 18 grammatical items were deleted in the calculation of the item difficulties, because no students used those items in their compositions (See Appendix A for the frequency of each item, and B.1 for the whole output of the Rasch analysis).

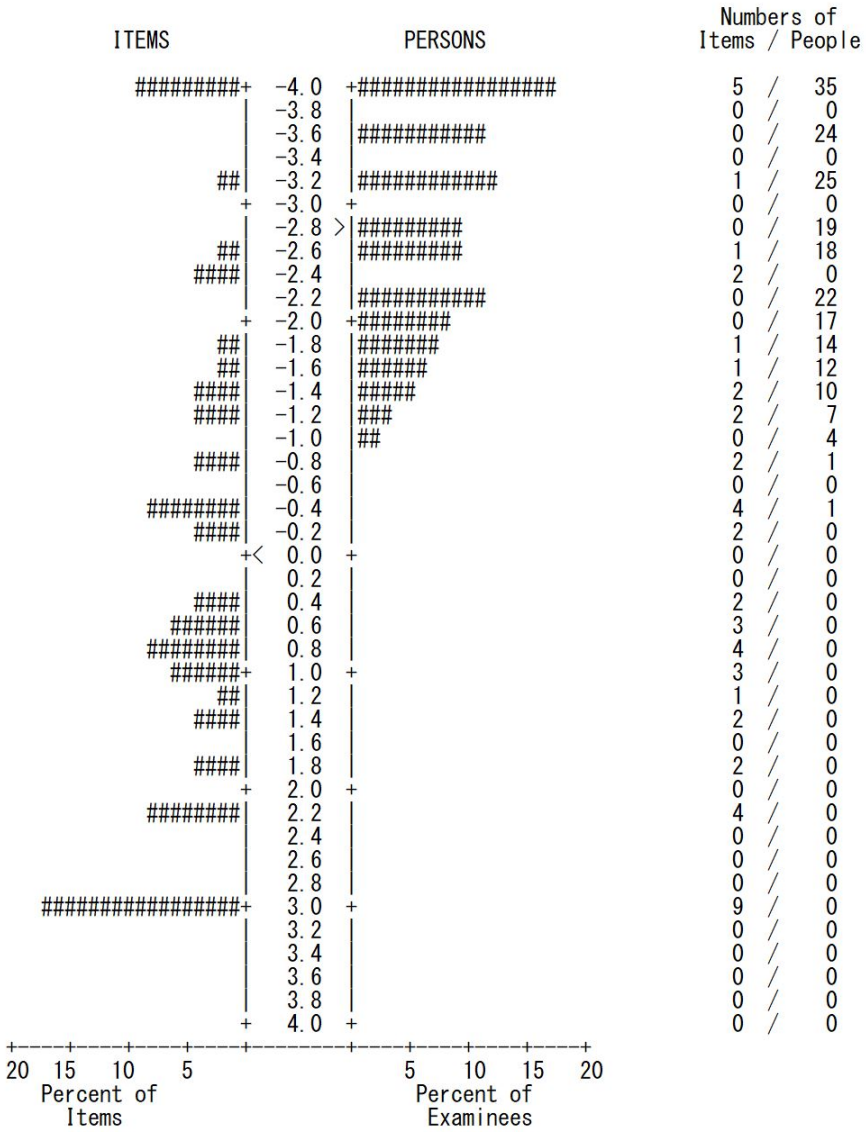


Figure 8. Item-Person Map (the first-year: Composition A)

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
32	-7.440	0.557	3.260	15	32
33	-6.393	0.357	5.328	15	42
2	-5.599	0.266	13.366	15	49
36	-5.099	0.226	17.093	15	54
39	-4.807	0.208	9.036	15	56
1	-3.158	0.156	10.250	15	71
35	-2.621	0.153	4.243	15	76
4	-2.390	0.154	28.259	15	78
61	-2.390	0.154	21.191	15	78
34	-1.865	0.160	14.570	15	83
55	-1.606	0.166	17.073	15	85
3	-1.497	0.169	22.563	15	86
38	-1.441	0.171	20.890	15	87
49	-1.234	0.178	11.361	15	89
45	-1.140	0.182	12.928	15	90
40	-0.866	0.194	27.866	15	92
71	-0.828	0.196	11.031	15	92
31	-0.453	0.219	11.255	15	96
68	-0.356	0.226	10.573	15	97
60	-0.305	0.230	10.038	15	97
10	-0.305	0.230	16.930	15	97
63	-0.252	0.234	18.353	15	98
37	-0.197	0.239	17.231	15	98
5	0.433	0.303	13.366	15	104
11	0.433	0.303	11.854	15	104
66	0.527	0.315	18.722	15	105
69	0.527	0.315	20.208	15	105
50	0.630	0.329	11.275	15	106
59	0.742	0.345	6.201	15	107
65	0.742	0.345	18.039	15	107
58	0.866	0.363	16.343	15	108
44	0.866	0.363	15.056	15	108
64	1.006	0.386	20.223	15	109
30	1.006	0.386	6.476	15	109
62	1.006	0.386	8.244	15	109
57	1.166	0.415	14.065	15	111
41	1.354	0.452	18.600	15	112
70	1.354	0.452	16.373	15	112
42	1.873	0.575	6.490	15	117
54	1.873	0.575	30.417	15	117
20	2.280	0.699	12.826	15	121
6	2.280	0.699	6.926	15	121
48	2.280	0.699	40.611	15	121
9	2.280	0.699	6.806	15	121
56	2.969	0.977	7.807	15	127
47	2.969	0.977	17.307	15	127
7	2.969	0.977	42.131	15	127
15	2.969	0.977	7.807	15	127
12	2.969	0.977	18.728	15	127
8	2.969	0.977	31.226	15	127
67	2.969	0.977	12.866	15	127
23	2.969	0.977	8.027	15	127
19	2.969	0.977	9.030	15	127

Figure 9. Order of the item difficulties (the first-year: Composition A)

3.4.1.2 The Acquired Grammatical Items Extracted from Composition B (Narrative Essay) Written in the Second Academic Year

In the same way, I looked into the distributions of the item difficulties and the students' abilities from the data of Composition B, written in the second academic year. As in the case of Composition A, the item difficulties are widespread, but the students' abilities are located in a narrow area of lower abilities (*Figure 10*). This time, the persons with the highest ability are found at the item difficulty of -0.2. Thus, the grammatical items whose difficulties are -0.2 or lower can be judged as the ones that the students distinctively used in Composition B in the second year. In *Figure 11*, the items with difficulties of -0.2 or lower are No. 33 to No. 37, 17 items in total. Those items are "verb (base form)", "S+V+O", "*be*-copula (present)", "S+*be*+C (adjective)", "verb (past)", "S+V", "personal pronoun (as object)", "*be*-copula (past)", "when (conjunction)", "S+V+that-clause", "prepositional phrase (adjectival)", "infinitive (as object)", "can (ability)", "may/can/might (possibility)", "will", "because (conjunction)", and "S+look+C (adjective)".

There were four items that showed the highest difficulty, 3.919 (*Figure 11*). They were "can (request)", "NP+-*ed*", "want/like/expect+O+to-", and "can (permission)". However, again, 18 items were not used by any students, so they were deleted in calculating the item difficulties (See Appendix A for the frequency of each item, and B.2 for the whole output of the Rasch analysis).

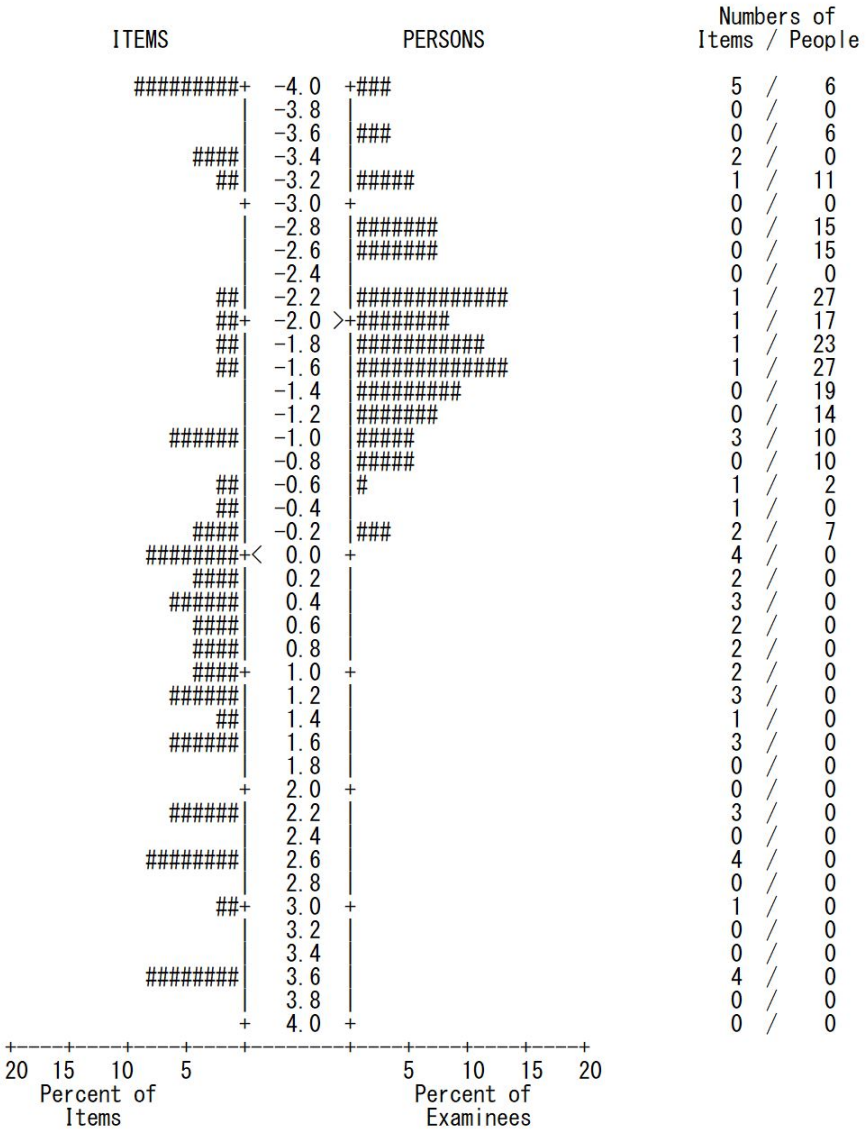


Figure 10. Item-Person Map (the second-year: Composition B)

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
33	-6.895	0.672	4.427	15	37
2	-6.485	0.555	5.033	15	41
32	-5.336	0.336	5.619	15	51
36	-5.221	0.321	11.992	15	52
35	-4.432	0.239	11.059	15	60
1	-3.447	0.180	9.145	15	69
39	-3.447	0.180	20.566	15	69
34	-3.113	0.168	12.135	15	72
61	-2.100	0.149	16.446	15	81
4	-1.926	0.149	17.734	15	82
71	-1.752	0.149	9.916	15	84
55	-1.665	0.149	11.250	15	85
45	-1.097	0.157	13.222	15	90
10	-0.949	0.161	13.401	15	91
49	-0.923	0.162	14.186	15	92
62	-0.656	0.170	12.061	15	94
37	-0.354	0.183	17.426	15	97
44	-0.183	0.191	23.857	15	98
40	-0.110	0.195	19.832	15	99
65	-0.034	0.200	7.825	15	100
38	0.006	0.202	10.978	15	100
31	0.006	0.202	13.879	15	100
3	0.088	0.207	17.106	15	101
69	0.175	0.213	9.711	15	102
58	0.220	0.216	19.005	15	102
63	0.365	0.227	41.032	15	103
41	0.470	0.236	22.154	15	104
11	0.470	0.236	24.037	15	104
66	0.527	0.241	27.589	15	105
5	0.646	0.251	16.660	15	106
60	0.710	0.258	17.487	15	106
70	0.778	0.264	15.359	15	107
54	0.925	0.280	7.501	15	108
68	1.006	0.289	13.913	15	109
50	1.186	0.312	13.272	15	111
12	1.186	0.312	12.134	15	111
30	1.186	0.312	12.975	15	111
57	1.399	0.341	18.261	15	113
64	1.522	0.360	12.194	15	114
42	1.522	0.360	11.591	15	114
59	1.522	0.360	11.395	15	114
19	2.235	0.498	13.191	15	120
22	2.235	0.498	7.358	15	120
18	2.235	0.498	22.672	15	120
67	2.525	0.571	11.614	15	123
56	2.525	0.571	12.084	15	123
20	2.525	0.571	35.157	15	123
48	2.525	0.571	21.742	15	123
27	2.931	0.694	20.168	15	127
47	3.619	0.970	9.808	15	133
24	3.619	0.970	14.261	15	133
7	3.619	0.970	7.152	15	133
46	3.619	0.970	7.152	15	133

Figure 11. Order of the item difficulties (the second-year: Composition B)

3.4.1.3 The Acquired Grammatical Items Extracted from Composition C (Argumentative Essay) Written in the Third Academic Year

Finally, I scrutinized the distributions of the item difficulties and the students' abilities from the data of Composition C that the students wrote in their third academic year. The general view of the distributions is similar to the previous cases: the item difficulties range widely, while the students' abilities converged closely in the area of less abilities (*Figure 12*). Here, the persons with the highest ability are located at the item difficulty of 0.0. *Figure 13* shows that the items with difficulties of 0.0 and less are No. 2 to No. 50, 28 items. They are "S+V+O", "S+V", "*be*-copula (present)", "S+*be*+C (adjective)", "S+V+that-clause", "may/can/might (possibility)", "should (advice)", "personal pronoun (as object)", "if (conjunction: adverbial)", "can (ability)", "infinitive (as object)", "verb (past)", "will", "when (conjunction)", "prepositional phrase (adjectival)", "must (obligation)", "because (conjunction)", "it … to-" , "comparative", "gerund (as object)", "gerund (as subject)", "*be*-copula (past)", "infinitive (adjectival)", "it is … for – to-", "present progressive", "relative pronoun (except *whose*)", "there is/are", and "have to-". Actually, "verb (base form)" should be added to these items. This item was deleted in the calculation of item difficulties, because all the students used it. Therefore, the "acquired" grammatical items are 29 items in total.

In this analysis, five items showed the highest difficulty, 3.460 (*Figure 13*). Those were "can (request)", "S+V+O+-*ing*", "could/would you -?", "something+to-", and "tag question". In this composition, 14 items were not used by any students, so did not show item difficulties. On the contrary, the item "verb (base form)" was used by all the students, so it did not show item difficulty, either (See Appendix A for the frequency of each item, and B.3 for the whole output of the Rasch analysis).

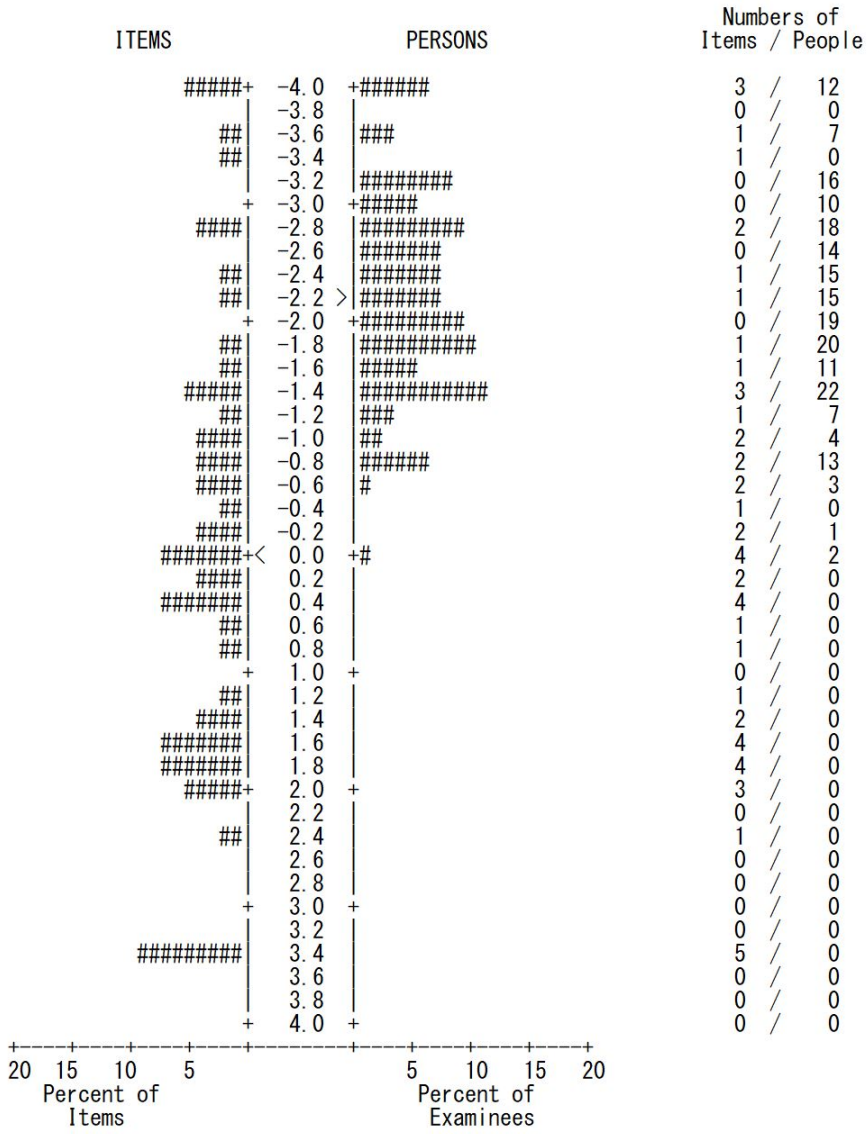


Figure 12. Item-Person Map (the third-year: Composition C)

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
2	-5.454	0.315	6.754	19	50
1	-4.803	0.249	15.819	19	56
32	-4.740	0.244	13.928	19	57
36	-3.558	0.176	13.875	19	68
4	-3.497	0.174	37.856	19	68
10	-2.848	0.157	30.237	19	74
12	-2.848	0.157	15.267	19	74
39	-2.313	0.151	12.057	19	79
60	-2.112	0.151	27.633	19	81
45	-1.753	0.152	20.197	19	84
55	-1.522	0.155	19.228	19	86
35	-1.498	0.156	17.222	19	86
49	-1.451	0.156	20.130	19	87
61	-1.451	0.156	29.092	19	87
71	-1.280	0.160	13.717	19	88
11	-1.074	0.165	52.089	19	90
62	-0.966	0.168	24.307	19	91
30	-0.882	0.171	16.044	19	92
65	-0.795	0.174	19.041	19	93
63	-0.610	0.181	13.203	19	94
64	-0.578	0.183	13.895	19	95
34	-0.443	0.189	22.551	19	96
57	-0.297	0.197	21.220	19	97
59	-0.219	0.201	20.327	19	98
40	-0.095	0.209	32.979	19	99
70	-0.095	0.209	30.405	19	99
44	-0.052	0.212	26.473	19	100
50	-0.052	0.212	13.946	19	100
38	0.135	0.225	16.280	19	101
68	0.239	0.233	10.100	19	102
3	0.350	0.242	18.760	19	103
66	0.350	0.242	43.225	19	103
54	0.471	0.253	24.692	19	104
37	0.471	0.253	15.318	19	104
69	0.677	0.274	26.783	19	106
20	0.753	0.282	18.699	19	107
18	1.119	0.327	22.363	19	110
58	1.356	0.362	11.375	19	112
19	1.495	0.385	24.566	19	114
47	1.656	0.413	9.332	19	115
27	1.656	0.413	46.674	19	115
13	1.656	0.413	24.196	19	115
31	1.656	0.413	23.442	19	115
15	1.844	0.450	11.762	19	117
24	1.844	0.450	26.694	19	117
5	1.844	0.450	10.808	19	117
48	1.844	0.450	14.230	19	117
41	2.072	0.500	23.010	19	119
56	2.072	0.500	6.588	19	119
22	2.072	0.500	17.274	19	119
67	2.363	0.573	20.833	19	122
43	3.460	0.974	11.108	19	131
6	3.460	0.974	39.434	19	131
51	3.460	0.974	25.754	19	131
14	3.460	0.974	19.277	19	131
46	3.460	0.974	41.897	19	131

Figure 13. Order of the item difficulties (the third-year: Composition C)

3.4.2 Development of Sentence Complexity of the Students' Compositions and Its Relationship with Fluency and Overall Writing Proficiency

3.4.2.1 Development of the MLUs, the Total Word Counts, and the Scores of GTEC Writing Test

I counted the total numbers of words in Compositions A, B, and C, written by the main subjects in each academic year. The total word count is defined as the index of fluency of the students' writing in this study. Also, I calculated the MLUs, or average length of sentences, as the index of sentence complexity of writing production. Then, I investigated the trajectories of the word counts and the MLUs for the three years, as well as the scores of GTEC writing test. Table 6, 7 and 8 show the descriptive statistics of both sets of data (the individual data are shown in Appendix C).

Table 6

Descriptive statistics of the MLUs

	N	*Mean*	*Median*	*SD*	*Min.*	*Max.*
1st year (Composition A)	209	6.9	6.8	1.26	4.6	12.5
2nd year (Composition B)	209	7.7	7.5	1.50	4.9	13.6
3rd year (Composition C)	209	8.8	8.5	2.17	4.1	17.7

Table 7

Descriptive statistics of the word counts

	N	*Mean*	*Median*	*SD*	*Min.*	*Max.*
1st year (Composition A)	209	94.7	90.0	45.41	20	250
2nd year (Composition B)	209	143.4	138.0	47.84	29	304
3rd year (Composition C)	209	112.7	112.0	50.82	20	295

Table 8

Descriptive statistics of the scores of GTEC writing test

	N	*Mean*	*Median*	*SD*	*Min.*	*Max.*
1st year	209	87.4	88.0	11.69	13	120
2nd year	209	103.4	105.0	11.32	50	138
3rd year	209	109.9	110.0	11.94	32	136

According to Hawkins and Filipović (2012), the MLUs can differentiate the CEFR levels: A2, 7.9; B1, 10.8; B2, 14.2; C1, 17.3; C2, 19.0. Although the students' MLUs increased every academic year, they went from below A2 up to the upper A2. Meanwhile, the total word counts increased from the first year to second year, but decreased from the second year to the third year. The scores of GTEC writing test went up every academic year, but the increase from the second year to the third year was relatively small, compared to the progress from the first year to the second year. Throughout the three years, there were large gaps between the highest scores and the lowest scores. Also, in the third year, even though the average score became higher, both the highest and lowest scores were lower than the second year. These are probably because of the difference of the students' motivation. Not every student was enthusiastic about taking the test, or studying English itself. In addition, in the third academic year, there was a tendency among the students to put energy into only what was required for the entrance examinations to universities. Almost all of them seemed to feel that they did not need writing skills particularly for the entrance examinations.

In order to confirm whether the differences (increase or decrease) in the three sets of data were significant or not, I decided to carry out several statistical tests. First, I needed to examine the normality of each data for choosing the appropriate test for the analyses. Table 9 shows the results of Kolmogorov-Smirnov test, a test of normality. It

proved that almost all the data were not normally distributed (the null hypothesis is "the data is normally distributed"), so I chose the Friedman test, a nonparametric test, instead of the repeated measures ANOVA, for investigating the statistical significance of the differences of the MLUs, the word counts and the scores of GTEC writing test between the academic years.

Table 9

Test of normality (Kolmogorov-Smirnov test)

	Statistic	*df*	*Sig.*	Normality
MLU: 1st year	.079	209	.00	No
MLU: 2nd year	.081	209	.00	No
MLU: 3rd year	.090	209	.00	No
Word count: 1st year	.087	209	.00	No
Word count: 2nd year	.064	209	.04	No
Word count: 3rd year	.059	209	.08	Yes
GTEC writing: 1st year	.119	209	.00	No
GTEC writing: 2nd year	.101	209	.00	No
GTEC writing: 3rd year	.119	209	.00	No

Table 10

Results of Friedman test (w/ multiple comparisons)

		Friedman			2nd-1st		3rd-1st		3rd-2nd	
	N	χ^2	*df*	*Sig.*	*z*	*Sig.*	*z*	*Sig.*	*z*	*Sig.*
MLU	209	125.90	2	.00*	-6.22	.00*	-10.23	.00*	-6.97	.00*
Effect size						*r* = .43		*r* = .71		*r* = .48
Word count	209	116.85	2	.00*	-10.54	.00*	-4.46	.00*	-7.54	.00*
Effect size						*r* = .73		*r* = .31		*r* = .52
GTEC writing	209	297.56	2	.00*	-11.71	.00*	-12.17	.00*	-9.14	.00*
Effect size						*r* = .81		*r* = .84		*r* = .63

The result of Friedman test on the MLU data showed that the development of MLUs for the three academic years was statistically significant. The post-hoc multiple comparisons (Wilcoxon signed-rank test) also showed statistical significance, with medium to large effect sizes. When comparing the difference in the total word counts between the years, we were also able to see a statistically significant difference, including the decrease from the second year to the third year. The effect size of the post-hoc analyses was from medium to large, as well. Also, the development of GTEC writing scores was found to be statistically significant, endorsed by large effect sizes (Table 10).

3.4.2.2 Relationship between the MLUs and the Total Word Counts

In order to examine the relationship between sentence complexity and the fluency of students' compositions, I calculated the Pearson's product-moment correlation coefficient. The strength of the correlation coefficients showed a modest but gradual development over the three academic years. When looking at the data of the first year (*Figure 14*), there was no relationship between the MLUs and the word counts ($r = .05$, $p = .51$). The data in *Figure 15* for the second year showed a weak relationship between the MLUs and the word counts ($r = .30$, $p = .00$). Then, in the third year, shown in *Figure 16*, the MLUs became moderately correlated to the total word counts ($r = .48$, $p = .00$).

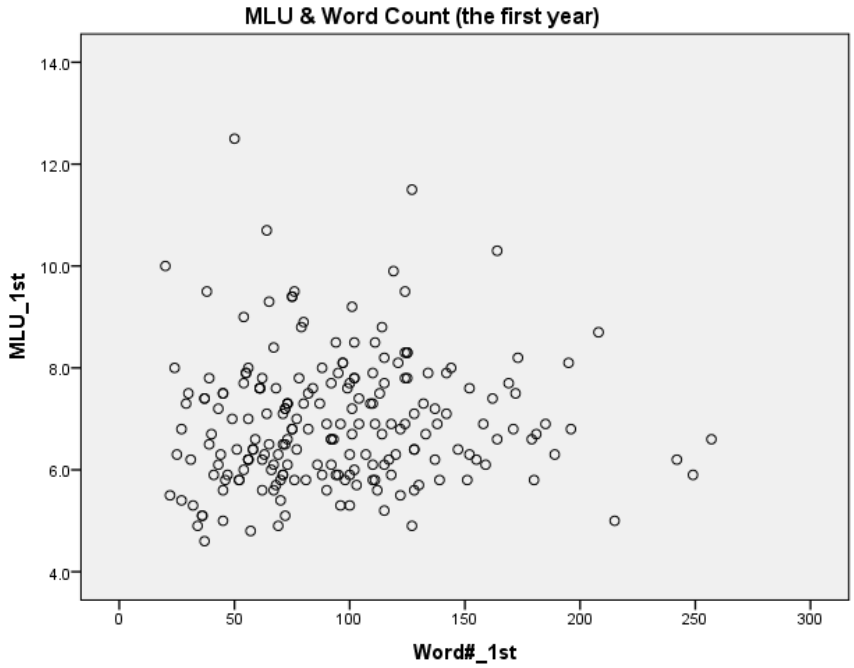


Figure 14. Scatter plot of the MLU and the word count: the first year

(r = .05, p = .51)

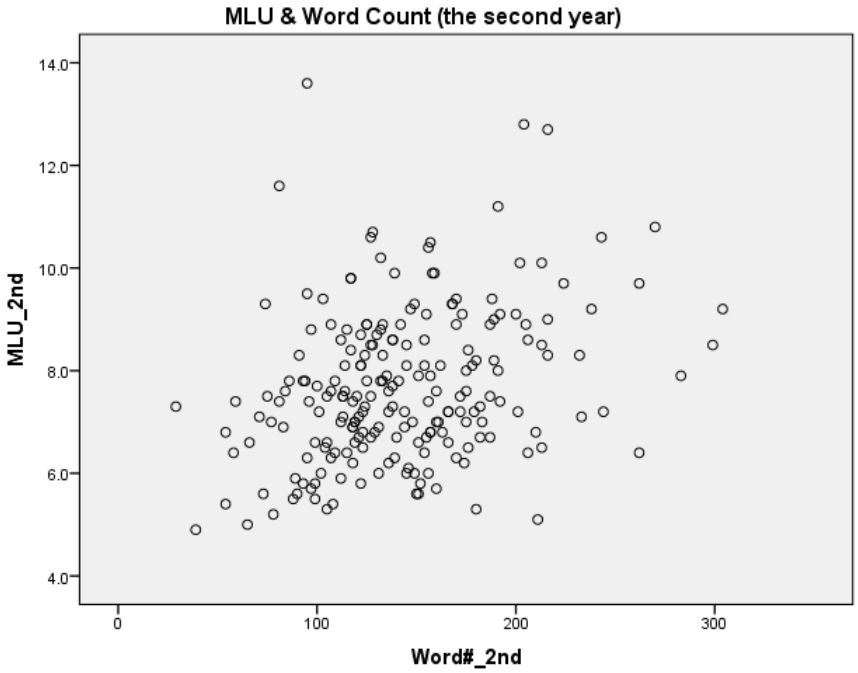


Figure 15. Scatter plot of the MLU and the word count: the second year

(r = .30*, p = .00)

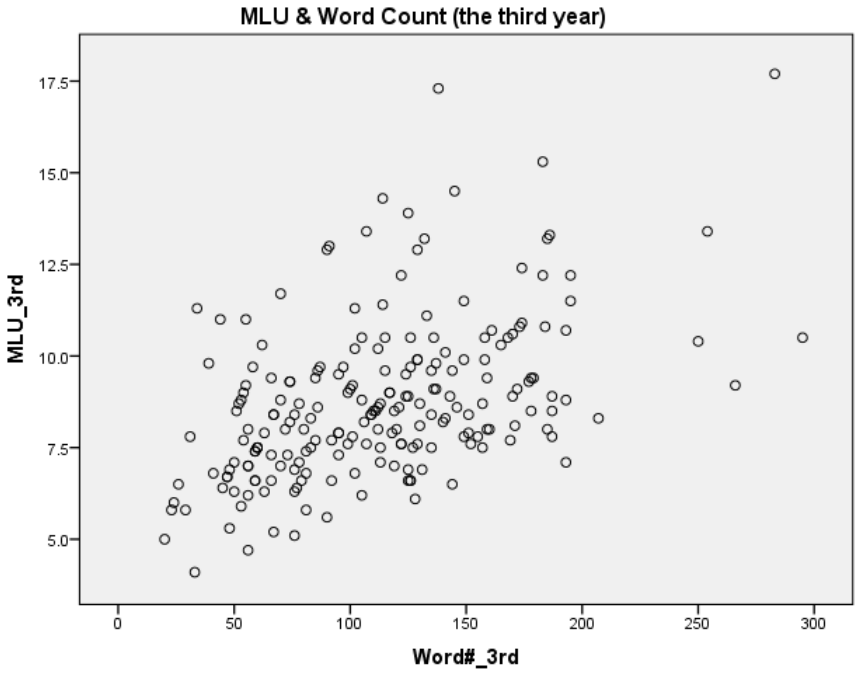


Figure 16. Scatter plot of the MLU and the word count: the third year

(r = .48*, p = .00)

3.4.2.3 Relationship between the MLUs and the Scores of GTEC Writing Test

As mentioned, these two sets of data were elicited from different pieces of writing. Thus, it might be considered that their relationship would not speak so eloquently of developmental aspects of the students' writing skills. However, if both sets of data equally reflect the students' writing ability in each academic year, it is worthwhile examining their relationship and its transition over three years.

Figure 17 shows that in the first academic year, there was no correlation between the students' MLUs and the GTEC writing scores. (r = - .04, p = .61). When looking at the data of the second year, as shown in *Figure 18*, there was still no relationship found between the MLUs and the GTEC scores (r = .05, p = .48). In the third year, as shown in *Figure 19*, there finally appeared weak correlation between the students' MLU values and the scores of GTEC writing test (r = .28*, p = .00).

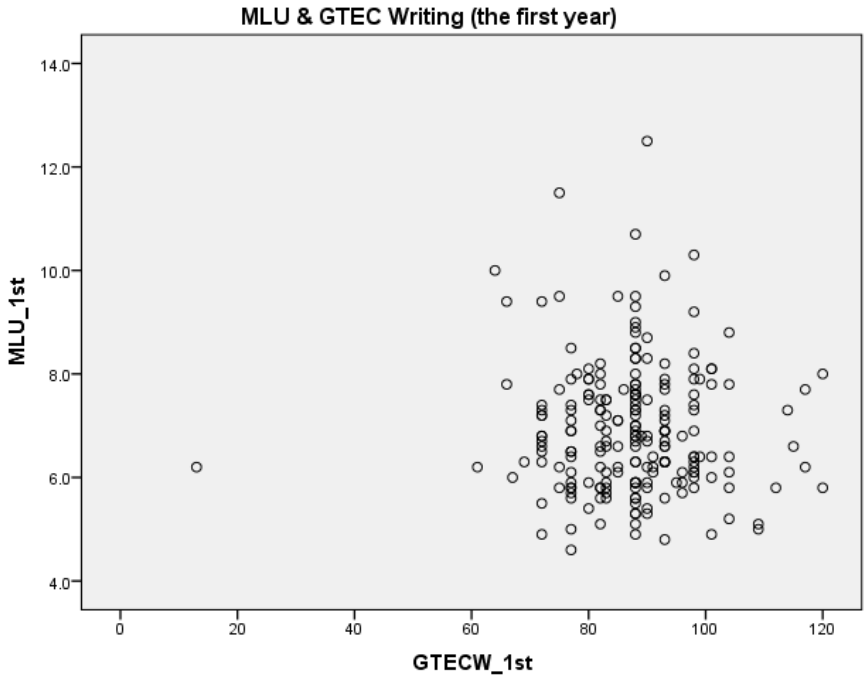


Figure 17. Scatter plot of the MLU and the score of GTEC writing: the first year

(r = - .04, p = .61)

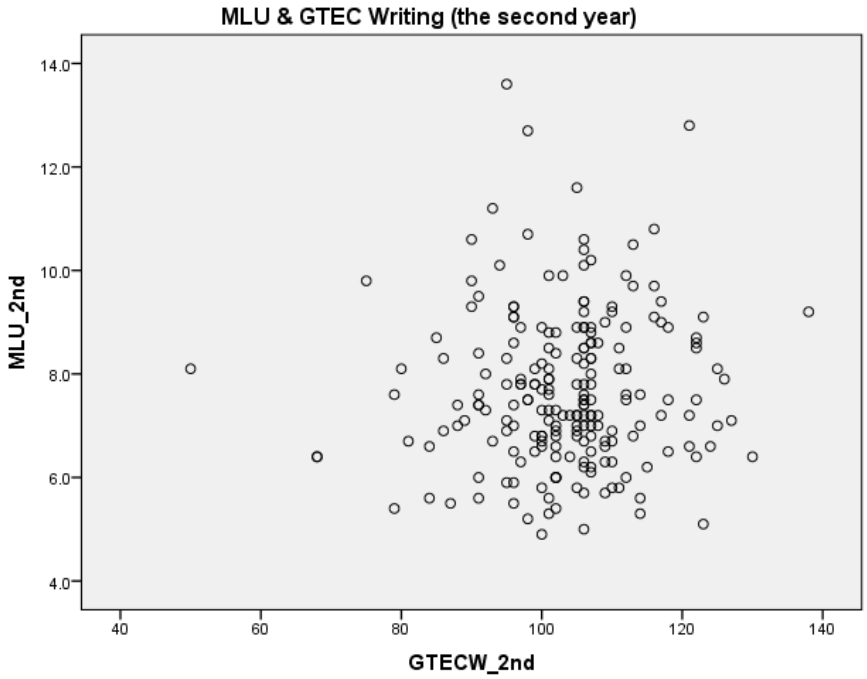


Figure 18. Scatter plot of the MLU and the score of GTEC writing: the second year

(r = .05, p = .48)

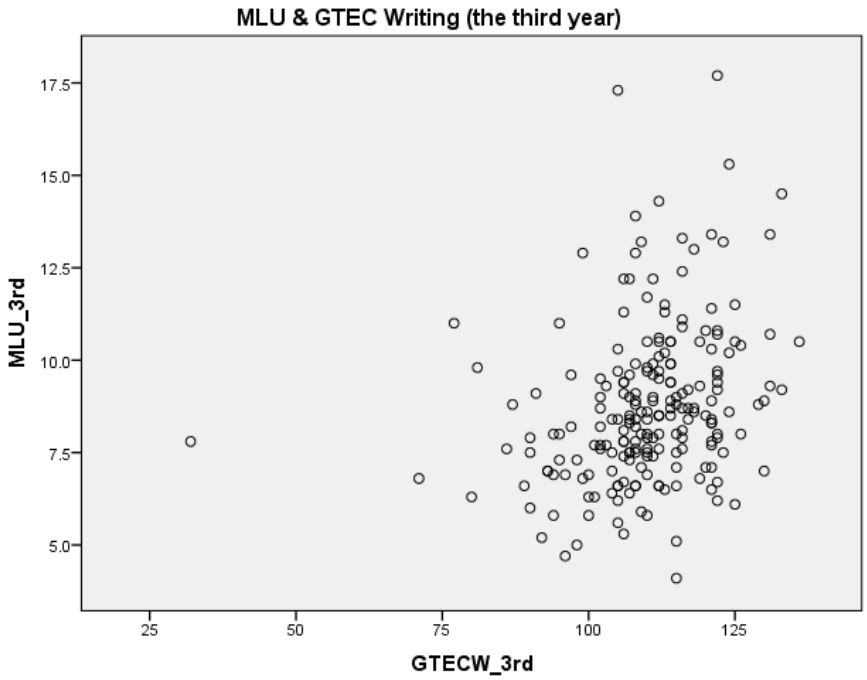


Figure 19. Scatter plot of the MLU and the score of GTEC writing: the third year

(r = .28*, p = .00)

3.4.3 Samples of a Student's Compositions

I picked one student and examined the characteristics of his three compositions. He was ranked first among the 209 students for the total score of the GTEC test in the third year. He steadily improved the total scores: 458 in the first year (the average was 383.3), 550 in the second year (the average was 417.0), and 619 in the third year (the average was 435.2). The present study aims to examine the subjects' acquisitional status of grammatical items as a group of learners, and these samples would not represent performance of the whole group. However, it will be informative to examine an individual's development of use of the listed grammatical items and observe some other items that are not included in the checklist.

3.4.3.1 Sample of the First-Year Composition (Composition A)

The topic was “The Key Person in My Life”, and the text type was narrative. The total number of words was 115, and the MLU was 8.2. When calculating the MLU, incorrect sentences were eliminated (Hawkins & Filipović, 2012). In the following sample, the isolated subordinate clause “*Because My friend and I played together sometimes (*sic*)” was considered to be an error and was not counted as a sentence in the calculation of the MLU.

> I think that my friends are very important to me. Because My friends and I play together sometimes. It’s too fun. And my best friend’s parents are very kind. I have over 100 friends in Japan, but I don't have friends in foreign countries. I want many friends in foreign countries. I wish I can get any friends in such countries.
>
> One week ago, I played with my best friend. The day was so exciting. We went shopping and ate “Gyudon” in MA***YA. It’s a significant memory of my memories.
>
> To be friends with many people is so good, I think. So, my important persons are my friends. I like my friends.
>
> (part of a proper noun is replaced by asterisks)

There were 10 grammatical items of the checklist found in this sample: “S+V+that”, “verb (base form)”, “*be*-copula (present)”, “S+*be*+C (adjective)”, “personal pronoun (object)”, “S+V+O”, “may/can/might (possibility)”, “S+V”, “verb (past)”, “*be*-copula (past)”. Sometimes one sentence has more than one checklist item. For example, the first sentence “*I think that my friends are very important to me.*” contains “S+V+that”, “verb

(base form)", "*be*-copula", "S+*be*+C (adjective)", and "personal pronoun (object)". When looking at subordinate structures, only "S+V+that" was found in "I think that …" and "I wish (that) …".

When looking at other linguistic features than the listed items, first of all, as the only one transition signal[6], "so" is used at the beginning of a sentence. Although the use in that position is sometimes considered to be inappropriate, especially in writing, it was judged as being correct here because it is allowed in such text type as narrative. As so-called (adverbial) intensifiers, "very", "too", and "so" are used. It is assumed that he has had some variety of expressions for intensification. It is also worth notice that one infinitive is used in the sentence-initial position as a subject. He seems to have been able to use different parts of speech as a subject: nouns, pronouns, and nominal infinitives.

As a sentence complexity index, the present study sticks to the MLU because it is attributed with the CEFR levels. However, it would not hurt us to try some other indices, which are average length of T-unit, frequently used as an index of complexity, and average number of clauses per T-unit, which is recommended in Wolfe-Quintero et al. (1998). As for this sample, the average length of T-unit was 7.1, and the average number of clauses per T-unit was 1.1.

[6]There are several categorizations of transition signal. Here, the list created by Western University (Canada) was referenced. It is available at: http://writing.uwo.ca/undergrads/handouts_new/Transition Signals.pdf

3.4.3.2 Sample of the Second-Year Composition (Composition B)

The topic of composition written in the second year was “My School Life of Last Year”, and the text type was narrative. The word count rapidly increased up to 206, and the MLU grew 8.6.

> I think that I was happy to spend a year in Y*** because I enjoyed my school life with new friends. At first, I had no friend. However, I could join and enjoy the class. In May 2006 we had a sports festival. We were helped by second and third grader student. Then I thought that “we also should help them.” After that we could help them and I thought “I'm very happy” because we could help each other. This is one of the good things.
>
> Second thing is school festival. Decorating the class was very fun. After we did it, the class became a wonderful pizza stand. We sold and sold pizza and juice. Finally, they were sold out. We felt very happy. After the festival I talked my friend T***. He said “It was great to become them sold out. We did our best! " His words made me very glad. I never forget the festivals I enjoyed in the first grade. There are wonderful memories. Friends in Y*** are also very wonderful. I was lucky to meet them in Y***.
>
> I’m happy to spend a year with my friends. I’m looking forward to spend the rest time with my best friends.
>
> (part of a proper noun is replaced by asterisks)

The listed grammatical items recognized in this sample leaped up to 21 items, from 10 in the first year. They were “S+V+that”, “verb (base form)”, “*be*-copula (past)”, “S+*be*+C (adjective)”, “infinitive (reason of feeling)”, “because (conjunction)”,

“S+V+O”, “verb (past)”, “can (ability)”, “passive”, “should (advice)”, “may/can/might (possibility)”, “personal pronoun (object)”, “*be*-copula (present)”, “gerund (subject)”, “it … to-”, “S+V+O+C (adjective)”, “relative pronoun (zero relative pronoun)”, “there is/are”, “prepositional phrase (adjectival)”, and “present progressive”. It is worthy of attention that he used a gerund as another form of sentential subject. In addition to “S+V+that”, two subordinate structures, “because (conjunction)”, and “zero relative pronoun”, were found in this sample. This student used “because” incorrectly in the first year, but used it correctly in two sentences here. The clause with zero relative pronoun, which is sometimes called “contact clause”, is learned in the third year of lower secondary school, often as the introductory example of relative structure. The “infinitive” to show reason of feeling was repeatedly used in the same or similar structure (“I am happy/lucky to …). Another infinitive was used embedded in “it is … to-” structure, which is one of the B2-labeled features listed in Hawkins and Filipović (2012). As another B2-labled feature, “S+V+O+C (adjective)” was also found. Furthermore, “passive”, learned in the second year of lower secondary school and listed as a B2 item in the Core Inventory (North et al., 2010), was used correctly here.

As transition signals to indicate logical sequence, “at first”, “second …”, and “finally” are used. Also, “after that” and “after S+V” are used to show the time order of events. In addition, “however” is used to introduce an opposite idea. This time, as an intensifier, only “very” was used five times. This may be because his desire to deliver more message might have decreased attention to adding variety to the expression.

When calculating the indices of sentence complexity, the average length of T-unit was 8.2 words, and the average number of clauses per T-unit was 1.3.

3.4.3.3 Sample of the Third-Year Composition (Composition C)

The topic of Composition C was “Should High School Students Have a Part-time Job?”, and the text type was argumentative. The word count decreased to 183 words, from 206 in the second year. However, the MLU jumped up to 15.3, from 8.6 in the second year.

> I think that high school students should work after school.
>
> I think so because they should help the economy of their family if they can work after school and working after school will be the preparation of their lives in society in the future. So, I think high school students should work after school if they can. However, not all high school students can work after school. There are many students who join the club activities such as Soccer, Baseball and Handball. Their activities end at six thirty almost every day. So, they don't have much time to work after school. I think I should work after school or in my free time. I want to try to work but I don't have much time. So, I decided to work when I become the college student. Students that don't have much time to work now should work when they become college students like me in my opinion. Of course, my main opinion is that high school students should work after school if they have much time and help their family.

In this sample, 19 grammatical items were found, a little fewer than 21 in the second year. They are “S+V+that”, “verb (base form)”, “should (advice)”, “S+V”, “because (conjunction)”, “prepositional phrase (adjectival)”, “if (conjunction: adverbial)”, “may/can/might (possibility)”, “gerund (subject)”, “will”, “there is/are”, “relative pronoun”, “S+V+O”, “infinitive (adjectival)”, “infinitive (object)”, “verb (past)”, “when

(conjunction)", "personal pronoun (object)", and "*be*-copula (present)". It is noteworthy that the student used relative pronouns in two sentences: both are in the subject position, but different words (*who* and *that*). Including the relative structure, five types of subordinate structures were used, which might have helped the leap of the MLU. The because structure was used, continuing from the second academic year, in a more complex sentence, combined with "if" clause. Thus, the item "because" seems to have been manipulatable to this student. Also, the different uses of infinitives were found here: adjectival infinitive and infinitive as an object. The latter was used following different verbs, as in "*want* to-", "*try* to-", "*decide* to-".

As transition signals, "so" is used repeatedly, and "however" introduces the counterargument to be handled later. Also, it is notable that "of course" is used elaborately in order to bring the logic back to the main assertion. While there are no adverbial intensifiers used, one downtoner, "almost" was found. In this argumentative essay, the writer might have used it in order for his opinion not to sound too assertive.

As for the complexity indices, the average length of T-unit was 14.1 words, and the average number of clauses per T-unit was 2.0.

Chapter 4. The Follow-Up Study

Idealistically, in order to assess the development of grammatical knowledge of English that learners can use, the same writing task should be given to them at the different stages of learning, such as in the first, second, and third years of upper secondary school. However, since the main subjects in this study worked on the writing tasks under the step-by-step pedagogical syllabi in an ordinary, not experimental classroom environment, the different writing tasks were provided in each academic year. Suspectedly those different tasks might have induced the learners to use particular task-dependent grammatical features. In other words, it was not certain whether a certain grammatical item emerged as a sort of criterial feature in the learner's development of proficiency, or happened to be used due to the requirement from a particular task. Thus, it was essential to identify which features were task-independent, for observing more actual status of the learners' manipulatable grammatical knowledge. In order to identify the task-independent grammatical items that are used in the three different compositions, I needed the set of data from the same three compositions written at the same time by the learners of the same ages. The research question of this follow-up study is as follows:

> When giving the same set of writing tasks at the same time to upper secondary school students with similar proficiency as the main subjects of the present study, which grammatical items used in their compositions can be judged to be influenced by the task difference?

4.1 Participants

The participants of this follow-up study were the students of each academic year at

another upper secondary school (108 first-year, 99 second-year, 91 third-year students). Their average scores on the GTEC for STUDENTS® are similar to those of the main subjects of this study (the average scores of the first-year students are 392.6 and 383.3, respectively, out of 660), so they can be considered to have similar proficiency in English.

4.2 Data Elicitation

With the cooperation of another upper secondary school, I asked their 298 students to work on the same three writing tasks at the same time (within a week). In order to avoid order effects, the students in each academic year were divided into six groups and worked on the three writing tasks in different order for counterbalancing. After receiving all the compositions, I examined their use of the grammatical items, using the same checklist as used in the main study.

4.3 Method

First, I conducted Cochran's Q test on each set of the 1/0 dichotomous data of the grammar use in the three tasks in order to confirm the statistical significance of the difference between the tasks. If the difference of use of a grammatical item between the three writing tasks is statistically significant, the item is considered to be task-dependent. Then, I examined the frequencies of the items characteristically used in the different tasks. In other words, "the frequencies" mean "the numbers of the students who used the particular item at least once".

4.4 Results

4.4.1 Frequencies of Grammatical Items Used in the Three Compositions

Each of the students at another upper secondary school wrote all the three essays at the same time. Table 11 shows the frequencies of the use of all the listed grammatical items. Here, the frequency means the number of the students who used a particular item at least once in their composition.

Table 11

Frequencies of grammatical items used in Compositions A, B, and C

	Composition	S+V	S+V+O	S+V+O+O	S+V+that	direct WH-clause
1st year (*N*=108)	A	60	105	18	53	1
	B	71	102	5	33	0
	C	77	99	6	63	4
2nd year (*N*=99)	A	77	99	27	52	2
	B	77	98	9	52	0
	C	89	96	13	66	3
3rd year (*N*=91)	A	77	91	21	58	1
	B	77	90	10	57	0
	C	84	91	14	75	3

	Composition	something + to-	NP + *-ed*	NP + *-ing*	double embedding: of-[of-]	may/can/might (possibility)
1st year (*N*=108)	A	1	2	0	0	19
	B	0	0	0	0	31
	C	1	0	0	0	49
2nd year (*N*=99)	A	2	2	0	0	36
	B	0	1	0	0	41
	C	0	0	1	0	66
3rd year (*N*=91)	A	1	0	1	0	23
	B	0	2	0	1	34
	C	1	1	1	0	50

	Composition	must (obligation)	should (advice)	S+V+O+to-	S+V+O+-*ing*	it … that
1st year (*N*=108)	A	5	5	4	0	1
	B	6	4	1	0	1
	C	23	46	0	0	0
2nd year (*N*=99)	A	5	3	3	0	1
	B	16	13	1	2	1
	C	20	56	1	0	1
3rd year (*N*=91)	A	6	7	3	0	9
	B	16	13	3	0	6
	C	22	48	2	0	10

	Composition	S+V+to-+that	whose (pron.)	pseudocleft I (what+S+V)	indirect WH-question	WH-+to-
1st year (*N*=108)	A	0	0	0	0	1
	B	0	0	0	1	1
	C	0	0	0	2	9
2nd year (*N*=99)	A	0	0	0	3	5
	B	0	1	0	1	5
	C	0	0	1	4	11
3rd year (*N*=91)	A	0	2	0	2	2
	B	0	0	0	5	4
	C	0	0	4	4	12

	Composition	would rather/ had better	adverbial -*ing* (following main clause)	seem/ be supposed +to-	want/ like/ expect +O +to-	S is easy to-
1st year (*N*=108)	A	0	0	0	3	0
	B	0	0	0	1	0
	C	1	0	0	0	0
2nd year (*N*=99)	A	0	1	0	3	0
	B	0	1	0	1	0
	C	1	3	0	0	1
3rd year (*N*=91)	A	0	2	0	2	0
	B	0	0	0	0	0
	C	0	0	0	2	0

	Composition	of [-'s] -	may (permission)	must (necessity)	should (probability)	it … to- **:B2**
1st year (*N*=108)	A	0	0	0	0	0
	B	0	0	0	0	3
	C	0	0	0	0	5
2nd year (*N*=99)	A	0	0	1	0	2
	B	0	0	0	0	4
	C	0	1	3	0	9
3rd year (*N*=91)	A	0	0	2	0	3
	B	0	0	0	0	6
	C	0	1	1	0	10

	Composition	S+V+O+C (adj.) **:B2**	*be*-copula (present)	verb (base form)	*be*-copula (past)	verb (past)
1st year (*N*=108)	A	23	106	106	35	39
	B	11	87	94	79	90
	C	7	96	105	3	6
2nd year (*N*=99)	A	19	98	99	43	52
	B	11	79	95	68	75
	C	2	88	99	10	16
3rd year (*N*=91)	A	15	91	90	47	58
	B	10	83	89	70	82
	C	8	90	89	12	24

	Composition	S+*be*+C(adj.)	S+look+C(adj.)	verb (3rd person singular)	personal pronoun (object)	present progressive
1st year (*N*=108)	A	93	15	41	80	23
	B	89	13	6	44	6
	C	82	2	6	35	14
2nd year (*N*=99)	A	89	18	46	76	31
	B	83	11	15	60	7
	C	73	11	14	43	6
3rd year (*N*=91)	A	84	13	44	82	20
	B	81	14	8	57	7
	C	78	2	26	41	14

	Composition	past progressive	S+V+O+C (noun)	tag question	there is/are	can (ability)
1st year (*N*=108)	A	4	1	0	6	26
	B	14	1	0	13	23
	C	0	0	0	12	13
2nd year (*N*=99)	A	8	2	0	10	28
	B	11	3	0	16	53
	C	4	0	1	15	31
3rd year (*N*=91)	A	4	2	0	12	33
	B	14	1	0	24	46
	C	1	0	0	18	26

	Composition	can (request)	can (permission)	be going to-	will	have to-
1st year (*N*=108)	A	0	0	3	27	4
	B	0	1	2	22	14
	C	0	3	1	24	36
2nd year (*N*=99)	A	0	0	5	29	5
	B	0	0	8	28	10
	C	1	0	2	39	33
3rd year (*N*=91)	A	0	0	5	26	16
	B	0	0	9	24	31
	C	0	5	4	46	45

	Composition	could/would you -?	will you -?	shall I -?	infinitive (purpose)	infinitive (as object)
1st year (*N*=108)	A	0	0	0	5	42
	B	0	0	0	8	34
	C	0	0	0	14	21
2nd year (*N*=99)	A	0	0	0	8	44
	B	0	0	0	10	49
	C	0	0	0	17	30
3rd year (*N*=91)	A	0	0	0	10	43
	B	0	0	0	13	48
	C	0	0	0	17	40

	Composition	infinitive (as complement)	infinitive (adjectival)	infinitive (reason of feeling)	it is … for- to-	if (conj.: adverbial)
1st year (*N*=108)	A	0	2	4	14	18
	B	2	3	6	22	11
	C	2	19	2	22	46
2nd year (*N*=99)	A	2	4	8	6	20
	B	4	9	11	19	9
	C	1	22	7	30	51
3rd year (*N*=91)	A	1	5	7	12	24
	B	3	13	3	14	23
	C	4	26	4	17	48

	Composition	when (conj.)	because (conj.)	gerund (as object)	gerund (as subject)	comparative
1st year (*N*=108)	A	41	20	15	0	9
	B	29	28	24	6	13
	C	8	22	14	30	19
2nd year (*N*=99)	A	43	16	19	5	9
	B	35	12	29	12	22
	C	29	17	19	22	17
3rd year (*N*=91)	A	52	27	18	5	9
	B	31	37	13	5	15
	C	26	36	20	24	14

	Composition	superlative	as - as	passive	present perfect	relative pronoun (except whose)
1st year (*N*=108)	A	14	2	18	11	8
	B	12	1	7	15	5
	C	14	0	6	2	8
2nd year (*N*=99)	A	10	1	19	33	24
	B	8	0	14	27	12
	C	13	0	20	12	21
3rd year (*N*=91)	A	14	1	30	24	27
	B	9	1	14	22	18
	C	10	1	16	9	29

	Composition	prepositional Phrase (adj.)
1st year (*N*=108)	A	13
	B	29
	C	11
2nd year (*N*=99)	A	30
	B	42
	C	20
3rd year (*N*=91)	A	16
	B	44
	C	21

4.4.2 Statistical differentiation of Task-Dependent Grammatical Items

As the results of Cochran's Q test on the 1/0 data elicited from the sets of three compositions written by the students of another upper secondary school, some items showed statistically-significant difference of the grammar use, while the others did not. The analysis was done for each set of the data from the students in the same academic year. I judged that the item was "task-dependent" if it showed statistically-significant difference of use in all the three tasks in two or three academic years out of three (though each effect size[7] was not large). If none of the students in more than one academic year used a particular item in every task, the item was excluded from the classification process of task dependence. An item was also excluded from the classification as being "impossible to judge" if it was judged as "task-dependent" in one academic year, and "task-independent" in another academic year, but none of the student used it in the other academic year (Table 12).

Table 12

Results of Cochran's Q test

		S+V	S+V+O	S+V+O+O	S+V+that	direct WH-clause
	Cochran's *Q*	7.825	3.375	12.560	20.588	5.200
1st year	*p* (α= .05)	.02*	.19	.00*	.00*	.07
(*N*=108)	Effect size	.04	.02	.06	.10	.02
	Judge	Dependent	Independent	Dependent	Dependent	Independent
	Cochran's *Q*	7.200	3.500	13.744	6.644	2.800
2nd year	*p* (α=.05)	.03*	.17	.00*	.04*	.01
(*N*=99)	Effect size	.04	.02	.07	.03	.01
	Judge	Dependent	Independent	Dependent	Dependent	Independent
	Cochran's *Q*	3.379	2.000	5.027	13.644	3.500
3rd year	*p* (α=.05)	.19	.37	.08	.00	.17
(*N*=91)	Effect size	.02	.01	.03	.08	.02
	Judge	Independent	Independent	Independent	Dependent	Independent
		something + to-	NP + *-ed*	NP + *-ing*	double embedding: of-[of-]	may/can/might (possibility)
	Cochran's Q	1.000	4.000	-	-	24.429
1st year	*p* (α=.05)	.61	.14	-	-	.00*
(*N*=108)	Effect size	.01	.02	-	-	.11
	Judge	Independent	Independent	-	-	Dependent
	Cochran's Q	4.000	2.000	2.000	-	24.219
2nd year	*p* (α=.05)	.14	.37	.37	-	.00*
(*N*=99)	Effect size	.02	.01	.01	-	.12
	Judge	Independent	Independent	Independent	-	Dependent
	Cochran's Q	1.000	2.000	1.000	2.000	19.750
3rd year	*p* (α=.05)	.61	.37	.61	.37	.00*
(*N*=91)	Effect size	.01	.01	.01	.01	.11
	Judge	Independent	Independent	Independent	Independent	Dependent

		must (obligation)	should (advice)	S+V+O+to-	S+V+O+-*ing*	it … that
	Cochran's Q	21.929	71.792	5.200	-	1.000
1st year	*p* (α=.05)	.00*	.00*	.07	-	.61
(*N*=108)	Effect size	.10	.33	.02	-	.01
	Judge	Dependent	Dependent	Independent	-	Independent
	Cochran's Q	12.067	78.000	1.600	4.000	.000
2nd year	*p* (α=.05)	.00*	.00*	.45	.14	1.00
(*N*=99)	Effect size	.06	.39	.01	.02	.00
	Judge	Dependent	Dependent	Independent	Independent	Independent
	Cochran's Q	11.200	52.536	.250	-	1.238
3rd year	*p* (α=.05)	.00*	.00*	.88	-	.54
(*N*=91)	Effect size	.06	.29	.00	-	.01
	Judge	Dependent	Dependent	Independent	-	Independent

		S+V+to-+that	whose (pron.)	pseudocleft I (what+S+V)	indirect WH-question	WH-+to-
	Cochran's Q	-	-	-	2.000	14.222
1st year	*p* (α=.05)	-	-	-	.37	.00*
(*N*=108)	Effect size	-	-	-	.01	.07
	Judge	-	-	-	Independent	Dependent
	Cochran's Q	-	2.000	2.000	2.000	4.000
2nd year	*p* (α=.05)	-	.37	.37	.37	.14
(*N*=99)	Effect size	-	.01	.01	.01	.02
	Judge	-	Independent	Independent	Independent	Independent
	Cochran's Q	-	4.000	8.000	1.273	10.500
3rd year	*p* (α=.05)	-	.14	.02*	.53	.01*
(*N*=91)	Effect size	-	.02	.04	.01	.06
	Judge	-	Independent	Dependent	Independent	Dependent

		would rather/ had better	adverbial *-ing* (following main clause)	seem/ be supposed +to-	want/ like/ expect +O +to-	S is easy to-
	Cochran's Q	2.000	-	-	3.500	-
1st year	*p* (α=.05)	.37	-	-	.17	-
(*N*=108)	Effect size	.01	-	-	.02	-
	Judge	Independent	-	-	Independent	-
	Cochran's Q	2.000	1.600	-	3.500	2.000
2nd year	*p* (α=.05)	.37	.45	-	.18	.37
(*N*=99)	Effect size	.01	.01	-	.02	.01
	Judge	Independent	Independent	-	Independent	Independent
	Cochran's Q	-	4.000	-	2.000	-
3rd year	*p* (α=.05)	-	.14	-	.37	-
(*N*=91)	Effect size	-	.02	-	.01	-
	Judge	-	Independent	-	Independent	-

		of [-'s] -	may (permission)	must (necessity)	should (probability)	it … to- **:B2**
	Cochran's Q	-	-	-	-	7.600
1st year	*p* (α=.05)	-	-	-	-	.02*
(*N*=108)	Effect size	-	-	-	-	.04
	Judge	-	-	-	-	Dependent
	Cochran's Q	-	2.000	3.500	-	9.385
2nd year	*p* (α=.05)	-	.37	.17	-	.01*
(*N*=99)	Effect size	-	.01	.02	-	.05
	Judge	-	Independent	Independent	-	Dependent
	Cochran's Q	-	2.000	2.000	-	4.625
3rd year	*p* (α=.05)	-	.37	.37	-	.10
(*N*=91)	Effect size	-	.01	.01	-	.03
	Judge	-	Independent	Independent	-	Independent

		S+V+O+C (adj.) **:B2**	*be*-copula (present)	verb (base form)	*be*-copula (past)	verb (past)
	Cochran's Q	14.857	18.690	16.625	105.253	119.400
1st year	*p* (α=.05)	.00*	.00*	.00*	.00*	.00*
(*N*=108)	Effect size	.07	.09	.08	.49	.55
	Judge	Dependent	Dependent	Dependent	Dependent	Dependent
	Cochran's Q	15.846	20.846	8.000	68.622	68.909
2nd year	*p* (α=.05)	.00*	.00*	.02*	.00*	.00*
(*N*=99)	Effect size	.08	.11	.04	.35	.35
	Judge	Dependent	Dependent	Dependent	Dependent	Dependent
	Cochran's Q	2.690	12.667	.400	73.114	79.625
3rd year	*p* (α=.05)	.26	.00*	.82	.00*	.00*
(*N*=91)	Effect size	.02	.07	.00	.40	.44
	Judge	Independent	Dependent	Independent	Dependent	Dependent
		S+*be*+C(adj.)	S+look/feel /become+C (adj.)	verb (3rd person singular)	personal pronoun (object)	present progressive
	Cochran's Q	4.769	12.783	56.977	55.770	11.421
1st year	*p* (α=.05)	.09	.00*	.00*	.00*	.00*
(*N*=108)	Effect size	.02	.06	.26	.26	.05
	Judge	Independent	Dependent	Dependent	Dependent	Dependent
	Cochran's Q	9.333	3.063	38.192	35.522	31.632
2nd year	*p* (α=.05)	.01*	.22	.00*	.00*	.00*
(*N*=99)	Effect size	.05	.02	.19	.18	.16
	Judge	Dependent	Independent	Dependent	Dependent	Dependent
	Cochran's Q	2.348	10.640	38.880	44.947	8.467
3rd year	*p* (α=.05)	.31	.01*	.00*	.00*	.02*
(*N*=91)	Effect size	.01	.06	.21	.25	.05
	Judge	Independent	Dependent	Dependent	Dependent	Dependent

		past progressive	S+V+O+C (noun)	tag question	there is/are	can (ability)
	Cochran's Q	17.333	1.000	-	5.059	5.792
1st year	*p* (α=.05)	.00*	.61	-	.08	.06
(*N*=108)	Effect size	.08	.01	-	.02	.03
	Judge	Dependent	Independent	-	Independent	Independent
	Cochran's Q	3.895	3.500	2.000	2.214	17.746
2nd year	*p* (α=.05)	.14	.17	.37	.33	.00*
(*N*=99)	Effect size	.02	.02	.01	.01	.09
	Judge	Independent	Independent	Independent	Independent	Dependent
	Cochran's Q	17.375	2.000	-	6.750	10.655
3rd year	*p* (α=.05)	.00*	.37	-	.03*	.01*
(*N*=91)	Effect size	.10	.01	-	.04	.06
	Judge	Dependent	Independent	-	Dependent	Dependent

		can (request)	can (permission)	be going to-	will	have to-
	Cochran's Q	-	3.500	1.200	.884	44.667
1st year	*p* (α=.05)	-	.17	.55	.64	.00*
(*N*=108)	Effect size	-	.02	.01	.00	.21
	Judge	-	Independent	Independent	Independent	Dependent
	Cochran's Q	2.000	-	4.154	4.440	37.167
2nd year	*p* (α=.05)	.37	-	.13	.11	.00*
(*N*=99)	Effect size	.01	-	.02	.02	.12
	Judge	Independent	-	Independent	Independent	Dependent
	Cochran's Q	-	10.000	2.625	18.500	24.745
3rd year	*p* (α=.05)	-	.01*	.27	.00*	.00*
(*N*=91)	Effect size	-	.06	.01	.10	.14
	Judge	-	Dependent	Independent	Dependent	Dependent

		could/would you -?	will you -?	shall I -?	infinitive (purpose)	infinitive (as object)
	Cochran's Q	-	-	-	6.632	11.424
1st year	*p* (α=.05)	-	-	-	.04*	.00*
(*N*=108)	Effect size	-	-	-	.03	.05
	Judge	-	-	-	Dependent	Dependent
	Cochran's Q	-	-	-	4.621	9.238
2nd year	*p* (α=.05)	-	-	-	.10	.01*
(*N*=99)	Effect size	-	-	-	.02	.05
	Judge	-	-	-	Independent	Dependent
	Cochran's Q	-	-	-	2.467	1.441
3rd year	*p* (α=.05)	-	-	-	.29	.49
(*N*=91)	Effect size	-	-	-	.01	.01
	Judge	-	-	-	Independent	Independent

		infinitive (as complement)	infinitive (adjectival)	infinitive (reason of feeling)	it is … for- to-	if (conj.: adverbial)
	Cochran's Q	2.000	24.818	2.182	3.765	38.830
1st year	*p* (α=.05)	.37	.00*	.34	.15	.00*
(*N*=108)	Effect size	.01	.12	.01	.02	.18
	Judge	Independent	Dependent	Independent	Independent	Dependent
	Cochran's Q	2.000	17.862	1.000	24.743	52.704
2nd year	*p* (α=.05)	.37	.00*	.61	.00*	.00*
(*N*=99)	Effect size	.01	.09	.01	.13	.27
	Judge	Independent	Dependent	Independent	Dependent	Dependent
	Cochran's Q	2.000	21.742	2.364	1.520	21.855
3rd year	*p* (α=.05)	.37	.00*	.31	.47	.00*
(*N*=91)	Effect size	.01	.12	.01	.01	.12
	Judge	Independent	Dependent	Independent	Independent	Dependent

		when (conj.)	because (conj.)	gerund (as object)	gerund (as subject)	comparative
	Cochran's Q	31.000	2.889	4.919	48.774	4.471
1st year	*p* (α=.05)	.00*	.24	.09	.00*	.11
(*N*=108)	Effect size	.14	.01	.02	.23	.02
	Judge	Dependent	Independent	Independent	Dependent	Independent
	Cochran's Q	5.103	1.556	5.000	17.520	7.167
2nd year	*p* (α=.05)	.08	.46	.08	.00*	.03*
(*N*=99)	Effect size	.03	.01	.03	.09	.04
	Judge	Independent	Independent	Independent	Dependent	Dependent
	Cochran's Q	23.306	4.439	2.167	24.897	2.067
3rd year	*p* (α=.05)	.00*	.11	.34	.00*	.36
(*N*=91)	Effect size	.13	.02	.01	.14	.01
	Judge	Dependent	Independent	Independent	Dependent	Independent
		superlative	as - as	passive	present perfect	relative pronoun (except whose)
	Cochran's Q	.250	2.000	10.640	12.667	1.200
1st year	*p* (α=.05)	.88	.37	.01*	.00*	.55
(*N*=108)	Effect size	.00	.01	.05	.06	.01
	Judge	Independent	Independent	Dependent	Dependent	Independent
	Cochran's Q	1.267	2.000	1.676	15.955	8.069
2nd year	*p* (α=.05)	.53	.37	.43	.00*	.02*
(*N*=99)	Effect size	.01	.01	.01	.08	.04
	Judge	Independent	Independent	Independent	Dependent	Dependent
	Cochran's Q	1.750	.000	12.000	9.707	5.722
3rd year	*p* (α=.05)	.42	1.00	.00*	.01*	.06
(*N*=91)	Effect size	.01	.00	.07	.05	.03
	Judge	Independent	Independent	Dependent	Dependent	Independent

		prepositional Phrase (adj.)
1st year (*N*=108)	Cochran's Q	16.686
	p (α=.05)	.00*
	Effect size	.08
	Judge	Dependent
2nd year (*N*=99)	Cochran's Q	14.000
	p (α=.05)	.00*
	Effect size	.07
	Judge	Dependent
3rd year (*N*=91)	Cochran's Q	24.327
	p (α=.05)	.00*
	Effect size	.13
	Judge	Dependent

4.4.3 Classification of the Task-Dependent / -Independent Grammatical Items

The final classification is shown below. Table 13 shows the grammatical items that are judged as "task-dependent" (28 items), and the items in Table 14 are judged as the "task-independent" grammatical items (30 items). However, 13 items were excluded from this classification because they were not used by any students in two or three academic years, or because the judgment of task dependence was tied between two academic years, with no users in the other academic year.

Table 13

Task-dependent grammatical items (28 items)

1	S+V	33	verb (base form)	55	infinitive (obj.)
3	S+V+O+O	34	*be*-copula (past)	57	infinitive (adj.)
4	S+V+that	35	verb (past)	60	if (conj.: adverbial)
10	may/can/might (possibility)	37	S+*look/feel/become*+C(adj.)	61	when (conj.)
11	must (obligation)	38	verb (3rd person singular)	64	gerund (subj.)
12	should (advice)	39	personal pronoun (obj.)	68	passive
20	WH-+to-	40	present progressive	69	present perfect
30	it … to- **:B2**	41	past progressive	71	prepositional phrase (adj.)
31	S+V+O+C(adj.)**: B2**	45	can (ability)		
32	*be*-copula (present)	50	have to-		

Table 14

Task-independent grammatical items (30 items)

2	S+V+O	21	would rather/had better	54	infinitive (purpose)
5	direct WH-question	22	adverbial clause with *-ing* (following main clause)	56	infinitive (complement)
6	something+to-	24	want/like/expect+O+to-	58	infinitive (reason of feeling)
7	NP+*-ed*	27	may (permission)	59	it is … for- to-
8	NP+*-ing*	28	must (necessity)	62	because (conj.)
13	S+V+O+to-	36	S+be+C(adj.)	63	gerund (obj.)
15	it … that	42	S+V+O+C(noun)	65	comparative
17	whose (relative pronoun)	44	there is/are	66	superlative
18	pseudocleftI(what+S+V)	48	be going to-	67	as – as
19	indirect WH-question	49	will	70	relative pronoun (except whose)

The excluded items are as follows:

of-[of-], “S+V+O+-ing”, “S+V+to-+that”, seem/be supposed+to-, S is easy to-, of [-'s]-, should (probability), tag question, can (request), can (permission), could/would you -?, will you -?, shall I -?

[7]Although the effect size for Cochran’s Q test is not discussed so often in articles, Rovai et al. (2014) introduced Kendall’s *W* and Cramér’s *V*, as the index. Here, I adopted Kendall’s *W*, because it can be calculated using the same settings on SPSS as in Cochran’s Q test. The test on the items that a small number of students used showed very weak effect size.

4.4.4 Frequencies of the Items Characteristically Used in Particular Tasks

Observed in three different compositions written by the same individuals, the use of several grammatical items displayed interesting characteristics. That is, more students used particular grammatical items in a particular task than the other two. Such items were different in the three different tasks. The interpretation of the results will be attempted later, in Chapter 5.

Table 15 shows high or relatively high-frequency grammatical items in Composition A (narrative essay, "The Key Person in My Life"). All the items listed here were proven to be task-dependent by Cochran's Q analysis. The values in the brackets are the percentage of students who used the item. The grammatical items that showed characteristically high frequencies in this task were "S+V+O+O", "S+V+O+C (adj.)", "third-person singular verb", "objective personal pronoun", and "present progressive".

Table 15

High-frequency items in Composition A (in comparison with B and C)

		S+V+O+O	S+V+O+C (C=adj.)	3rd person singular verb	personal pronoun (object)	present progressive
1st year (*N*=108)	A	18 (16.7)	23 (21.3)	41 (38.0)	80 (74.1)	23 (21.3)
	B	5 (4.6)	11 (10.2)	6 (5.6)	44 (40.7)	6 (5.6)
	C	6 (5.6)	7 (6.5)	6 (5.6)	35 (32.4)	14 (13.0)
2nd year (*N*=99)	A	27 (27.3)	19 (19.2)	46 (46.5)	76 (76.8)	31 (31.3)
	B	9 (9.1)	11 (11.1)	15 (15.2)	60 (60.6)	7 (7.1)
	C	13 (13.1)	2 (2.0)	14 (14.1)	43 (43.4)	6 (6.1)
3rd year (*N*=91)	A	21 (23.1)	15 (16.5)	44 (48.4)	82 (90.1)	20 (22.0)
	B	10 (11.0)	10 (11.0)	8 (8.8)	57 (62.6)	7 (7.7)
	C	14 (15.4)	8 (8.8)	26 (28.6)	41 (45.1)	14 (15.4)

A: Composition A, B: Composition B, C: Composition C

The higher-frequency grammatical items found in the data of Composition B (narrative essay, “My School Life of Last Year”) are shown in Table 16 below. They are “past tense *be*-copula”, “past tense verb”, “past progressive”, and “prepositional phrase as an adjective”. All these items were judged as being task-dependent by Cochran’s Q analysis.

Table 16

High-frequency items in Composition B (in comparison with A and C)

		be-copula (past)	verb (past)	past progressive	prep. phrase (adj.)
1st year (*N*=108)	A	35 (32.4)	39 (36.1)	4 (3.7)	13 (12.0)
	B	79 (73.1)	90 (83.3)	14 (13.0)	29 (26.9)
	C	3 (2.8)	6 (5.6)	0 (0.0)	11 (10.2)
2nd year (*N*=99)	A	43 (43.3)	52 (52.5)	8 (8.1)	30 (30.3)
	B	68 (68.7)	75 (75.8)	11 (11.1)	42 (42.4)
	C	10 (10.1)	16 (16.2)	4 (4.0)	20 (20.2)
3rd year (*N*=91)	A	47 (51.6)	58 (63.7)	4 (4.4)	16 (17.6)
	B	70 (76.9)	82 (90.1)	14 (15.4)	44 (48.4)
	C	12 (13.2)	24 (26.4)	1 (1.1)	21 (23.1)

A: Composition A, B: Composition B, C: Composition C

The items listed in Table 17 are the higher-frequency grammatical items used in Composition C (argumentative essay, “Should High School Students Have a Part-time Job?”). They are “S+V+that-clause”, “may/can/ might (possibility)”, “must (obligation)”, “should (advice)”, “have to”, “adverbial if-clause”, and “gerund as a subject”. According to Cochran’s Q analysis, all the items here were regarded as being task-dependent.

Table 17

High-frequency items in Composition C (in comparison with A and B)

		S+V+that	may/can/ might (possibility)	must (obligation)	should (advice)	have to-	if (adverbial)	gerund (as subject)
1st year (*N*=108)	A	53 (49.1)	19 (17.6)	5 (4.6)	5 (4.6)	4 (3.7)	18 (16.7)	0 (0.0)
	B	33 (30.6)	31 (28.7)	6 (5.6)	4 (3.7)	14 (13.0)	11 (10.2)	6 (5.6)
	C	63 (58.3)	49 (45.4)	23 (21.3)	46 (42.6)	36 (33.3)	46 (42.6)	30 (27.8)
2nd year (*N*=99)	A	52 (52.5)	36 (36.4)	5 (5.1)	3 (3.0)	5 (5.1)	20 (20.2)	5 (5.1)
	B	52 (52.5)	41 (41.4)	16 (16.2)	13 (13.1)	10 (10.1)	9 (9.1)	12 (12.1)
	C	66 (66.7)	66 (66.7)	20 (20.2)	56 (56.6)	33 (33.3)	51 (51.5)	22 (22.2)
3rd year (*N*=91)	A	58 (63.7)	23 (25.3)	6 (6.6)	7 (7.7)	16 (17.6)	24 (26.4)	5 (5.5)
	B	57 (62.6)	34 (37.4)	16 (17.6)	13 (14.3)	31 (34.1)	23 (25.3)	5 (5.5)
	C	75 (82.4)	50 (54.9)	22 (24.2)	48 (52.7)	45 (49.5)	48 (52.7)	24 (26.4)

A: Composition A, B: Composition B, C: Composition C

Chapter 5. Discussion

5.1 Development of Grammatical Items Emerging in the Different Writing Tasks

I compared the "acquired" grammatical items of all the three academic years that the Rasch model analyses extracted. Table 18 shows those items and the academic years in lower secondary school when the items were taught as the target grammar in government-authorized textbooks (the column "LSS" shows the academic years of lower secondary school).

Although their item difficulties were different, 14 items were extracted as being "acquired" in the three different tasks in common. They were "S+V", "S+V+O", "S+V+that-clause", "*be*-copula (present)", "verb (base form)", "*be*-copula (past)", "verb (past)", "S+*be*+C (adjective)", "personal pronoun (as object)", "can (ability)", "will", "infinitive (object)", "when (conjunction)", and "prepositional phrase (adjectival)". Among these, "S+V", "S+V+O", "*be*-copula (present)", "verb (base form)", "*be*-copula (past)", "verb (past)", and "S+be+C (adjective)" would be versatile items. Even in Composition B, in which the students wrote about their past experience, many of them used the present tense of "be-copula" and "verb" to mention their present situation or opinion in the last sentence after writing about their past episodes (e.g. *Now I am a third-year student. / I want to enjoy my school life this year, too*.). As for "S+V+that-clause", almost all the samples were "*I think that …*". This structure is considered to be acquired as a formulaic expression. Also, since the auxiliary verb "can" and "will" are learned in the early time of lower secondary school, they would be so familiar to the students. The same thing could be said of the conjunction "when". Many of the students learned "when-clause" as their first encounter with a subordinate structure. The item "infinitive (object)"

emerged as "*want to …*" in many examples, so this would have become a formulaic chunk in the students' language system. A qualifier "prepositional phrase (adjectival)" was also widely used in all the three tasks, but many of the samples were "*of …*". Sometimes it was used awkwardly, for example "*money of my mother*", instead of "*my mother's money*".

Table 18

Trajectories of the "acquired" grammatical items through the three years

1st-year items	LSS	2nd-year items	LSS	3rd-year items	LSS
S+V	*1	S+V	*1	S+V	*1
S+V+O	*1	S+V+O	*1	S+V+O	*1
S+V+O+O	2	S+V+that-clause	2	S+V+that-clause	2
S+V+that-clause	2	may/can/might (possibility)	2	may/can/might (possibility)	2
S+V+O+C (adj.)	3	*be*-copula (present)	1	must (obligation)	2
be-copula (present)	1	verb (base form)	1	should (advice)	2
verb (base form)	1	*be*-copula (past)	2	it … to-	3
be-copula (past)	2	verb (past)	1	*be*-copula (present)	1
verb (past)	1	S+*be*+C (adj.)	1	verb (base form)	1
S+*be*+C (adj.)	1	S+*look/feel/become*+C (adj.)	2	*be*-copula (past)	2
verb (3rd person singular)	1	personal pronoun (as obj.)	1	verb (past)	1
personal pronoun (as obj.)	1	can (ability)	1	S+*be*+C (adj.)	1
present progressive	1	will	2	personal pronoun (as obj.)	1
can (ability)	1	infinitive (as obj.)	2	present progressive	1
will	2	when (conj.)	2	there is/are	2
infinitive (as obj.)	2	because (conj.)	2	can (ability)	1
when (conj.)	2	prepositional phrase (adj.)	*?	will	2
prepositional phrase (adj.)	*?			have to-	2
				infinitive (as obj.)	2
				infinitive (adjectival)	2

	it is … for- to-	3
	if (conj.)	2
	when (conj.)	2
	because (conj.)	2
	gerund (as obj.)	2
	gerund (as sub.)	2
	comparative	2
	prepositional phrase (adj.)	*?
	relative pronoun (exc. whose)	3

LSS: academic year of lower secondary school
*1: not listed as target grammar, but learned in the first year of lower secondary school
*?: not listed as target grammer and learned in various times

A total of 18 items were extracted from the compositions written in the first year of upper secondary school, and all of them, except for "S+V+O+C (adjective)", were learned in the first and second years of lower secondary school. Also, all the CEFR-labeled features among them were of the A2 level. As for the item "S+V+O+C (adjective)", which was a B2 feature and taught in the third year of lower secondary school at that time, all of the examples were "(someone) make[s] me happy", probably because the topic was "The Key Person in My Life" and the expression had been acquired as a lexical chunk used in that context. Some grammatical items, including this "S+V+O+C (adjective)", might be acquired not as syntactic structures but as lexical chunks first, and used for a while as formulaic expressions —the same expressions as learned for the first time. It might take a while for learners to be able to apply particular grammatical items to various contexts that require conjugation or word choice. In addition to "S+V+O+C (adjective)", "S+V+O+O", "verb (third-person singular)", and "present progressive" were distinctively used in this composition. In order to describe the relationship and interaction between persons, the students would have recalled this structure that they had learned in

the second year of lower secondary school (e.g. *He taught us English*.). Many of them also needed to use "verb (third-person singular)" to introduce the present situation or habit of their key person (e.g. *She goes to another high school now*.). As for "present progressive", many students used the item for describing what their key person is doing recently (e.g. *He is playing handball at his high school*.). In the follow up study too, these four items "S+V+O+C (adjective)", "S+V+O+O", "verb (third-person singular)" and "present progressive" were proven to be task-dependent items by Cochran's Q test.

While 18 items were extracted as being "acquired", some items were judged as being far beyond the students' ability. In addition, there were quite a few items that no students used in this composition. Table 19 displays the items that showed the highest item difficulty in the Rasch model analysis. If they are the CEFR-labeled items, that is, the criterial features in the English Profile Programme, their CEFR levels are shown in the brackets. As for non CEFR-labeled items, the academic years of lower secondary school when they are learned are shown in the brackets (e.g. "LSS.2": the second year of lower secondary school). Also, Table 20 shows the items that were not used by any students in the first-year writing.

Table 19

Grammatical items showing the highest item difficulty

(Composition A: in the first year)

indirect WH-question (B1)	NP+-*ing* (B1)	as – as (LSS.2)
seem/be supposed+to- (B1)	NP+-*ed* (A2)	can (permission) (LSS.2)
it … that (B1)	should (advice) (A2)	infinitive (complement) (LSS.2)

Table 20

Unemployed grammatical items (Composition A: in the first year)

S+V+O+to- (B1)	adverbial *-ing* (post) (B1)	should (probability) (B1)
S+V+O+*-ing* (B1)	want/like/expect+O+to- (B1)	tag question (LSS.3)
S+V+to-+that (B1)	S is easy to- (B1)	can (request) (LSS.2)
whose (relative pronoun) (B1)	double embeddings: of [-'s] (B1)	could/would you -? (LSS.2)
pseudocleft I (what+S+V) (B1)	may (permission) (B1)	will you -? (LSS.2)
would rather/had better (B1)	must (necessity) (B1)	shall I -? (LSS.2)

If at least one student happens to use a particular item, the item difficulty is calculated. Among the items in Table 19, "indirect WH-question", "NP+-*ing*", and "NP+-*ed*", which are all learned in the third academic year of lower secondary school, are often considered to be difficult items to learn. Teachers would spend a lot of time instructing them, and students would practice with them intensively in drill books, for example. However, it seems that these items are so difficult to acquire that more time and exposure would be needed for them to emerge in language output. Meanwhile, none of the students used 18 items shown in Table 20. Indeed, 13 out of 16 B1-labeled items (more than 80 percent) are included here. Some items, such as "tag question", "could/would you -?", "will you -?", and "shall I -?", seem to have a convincing reason that they were not used. These items tend to be used in spoken production more often than written production. However, it is hard to judge whether the students did not need to use the items in this task, or whether they were not able to use those items.

There were 17 items extracted from the second-year compositions. The topic was "My School Life of Last Year." Although the average total word counts remarkably increased from the first year (94.7 to 143.4), the number of grammatical items used reduced by one, and over 80 percent (14 out of 17 items) were the same items as in the

first year. Here again, no B1-level features were extracted. Both of the compositions were narrative writing, and the students wrote the compositions based on their personal experience. In terms of the text type, the influence of the task difference between Compositions A and B should not have been so strong, but the development of the grammar use might not have kept up with the development of the writing fluency. It is assumed that in their first and second years of upper secondary school, the students had acquired only the basic grammatical items learned early in lower secondary school and managed to use them repeatedly whenever they had to express themselves in English. In this composition, "may/can/might (possibility)", "S+*look/feel/become*+C (adjective)", "because (conjunction)" were distinctively extracted. As for "may/can/might (possibility)", many samples of "dubious" use of "can" were counted as being correct. Many students used "can" in a sentence, such as *"I could make a lot of friends."* (instead of *"I made a lot of friends."*). I interpreted the function of this "can" as "possibility" of being blessed with a good chance, not as personal "ability". The item "S+*look/feel/become*+C (adjective)" was frequently used in the same way as *"I felt happy (sad)"*. The conjunction "because" was used incorrectly in most of the attempts in their first-year composition (Murakoshi, 2012a). After a year, many students might have become able to use the item successfully in their writings. The follow-up study found that "may/can/might (possibility)" and "S+*look/feel/become*+C (adjective)" were task-dependent, but that "because (conjunction)" was task-independent. As mentioned above, in the case of the main subjects, the success rate of using "because" was extremely low in their first academic year, so the item was not extracted as being "acquired" at that time. Then, in this composition, written in their second academic year, the item "because" was distinctively extracted as an "acquired" item. In the case of the students in the follow-up study, the gap of its use between the three compositions was relatively small in all the

students in three academic years (Table 11), which made "because" judged as "task-independent".

The items that showed the highest item difficulty are displayed in Table 21. Also, the items that no students used in Composition B are shown in Table 22.

Table 21

Grammatical items showing the highest item difficulty

(Composition B: in the second year)

want/like/expect+O+to- (B1)	can (request) (LSS.2)
NP+-*ed* (A2)	can (permission) (LSS.2)

Table 22

Unemployed grammatical items (Composition B: in the second year)

something+to- (A2)	S+V+to-+that (B1)	must (necessity) (B1)
NP+-*ing* (B1)	whose (relative pronoun) (B1)	should (probability) (B1)
double embeddings: of-[of-] (A2)	would rather/had better (B1)	tag question (LSS.3)
S+V+O+to- (B1)	seem/be supposed +to- (B1)	could/would you -? (LSS.2)
S+V+O+-*ing* (B1)	S is easy to- (B1)	will you -? (LSS.2)
it … that (B1)	double embeddings: of [-'s] (B1)	shall I -? (LSS.2)

As well as "NP+-*ed*", "want/like/expect+O+to-" is usually taught in the third academic year of lower secondary school, and they are often considered to be difficult items for students to learn. Thus, it seems to be natural that not many students did not have a command of those items yet. The other items "can (request)" and "can (permission)" would have been hard to emerge in Composition B, in which the students wrote about their past experiences or past events. In the meantime, 18 items shown in Table 22 were not used by any students in their second-year writing. 12 out of 16 B1-labeled items (more

than 80 percent) are still included here. Again, the items “tag question”, “could/would you -?”, “will you -?”, and “shall I -?” left the question whether these items were hard to emerge in written production or whether the students were not able to manipulate these items yet at that point of time.

In the third year, while the average total word counts decreased from the second year (143.4 to 112.7), the Rasch-extracted grammatical items increased from 17 to 29. 16 out of 29 items were the same ones as in the second year, and the new 13 items included three items that were learned in the third academic year of lower secondary school. The CEFR-labeled features here were of the A2 level, except for “it … to-” (B2). The composition was an argumentative essay in which students had to answer the question, “Should High School Students Have a Part-time Job?”, so there would be some commonly used items, such as “should (advice)” or “comparative”. They needed to use “should (advice)” to express their assertion (answer to the topical question). Also, they needed to use “comparative” structures to compare the advantages and disadvantages of working part-time. The other items extracted distinctively in this composition were “may/can/might (possibility)”, “must (obligation)”, “it … to-”, “present progressive”, “there is/are”, “have to-”, “infinitive (adjectival)”, “it is … for- to-”, “if (conjunction: adverbial)”, “because (conjunction)”, “gerund (as object)”, “gerund (as subject)”, and “relative pronoun (except *whose*)”. Among these, the items learned most recently, in the third year of lower secondary school, were “it … to-”, “it is … for- to-”, and “relative pronoun (except *whose*)”. While some items seem to have been influenced by the task requirement, it may be worthy of attention that more varieties were found in the use of modal verbs and verbals (infinitives and gerunds). When looking back to the investigation of task dependence in the follow-up study, “may/can/might (possibility)”, “must (obligation)”, “should (advice)”, “it … to-”, “present progressive”, “have to-”, “infinitive (adjectival)”,

"if (conjunction: adverbial)", and "gerund (as subject)" were regarded as being task-dependent, and "there is/are", "infinitive (adjectival)", "it is … for- to-", "because (conjunction)", "gerund (as object)", and "relative pronoun (except *whose*)" were judged as being task-independent. Among these, "relative pronoun" was extracted as an "acquired" item for the first time in the third-year composition written by the main subjects, though it was found to be a task-independent item in the follow-up study. However, when looking at its frequencies examined in the follow-up study, the progress of its use was seen between the academic years, especially from the first to the second academic year (Table 11). Thus, "relative pronoun" would be widely used in different contexts, but it would take a certain amount of time for learners to be able to manipulate the item.

Table 23 shows the highest difficulty items, which were barely used in Composition C, and Table 24 displays the items that never emerged in the students' writing.

Table 23

Grammatical items showing the highest item difficulty

(Composition C: in the third year)

S+V+O+-*ing* (B1)	can (request) (LSS.2)	could/would you -? (LSS.2)
something+to- (A2)	tag question (LSS.3)	

Table 24

Unemployed grammatical items (Composition C: in the third year)

NP+-*ed* (A2)	would rather/had better (B1)	should (probability) (B1)
NP+-*ing* (B1)	seem/be supposed +to- (B1)	S+V+O+C (noun) (LSS.3)
double embeddings: of-[of-] (A2)	S is easy to- (B1)	will you -? (LSS.2)
S+V+to-+that (B1)	double embeddings: of [-'s] (B1)	shall I -? (LSS.2)
whose (relative pronoun) (B1)	must (necessity) (B1)	

The item "S+V+O+-*ing*" is usually learned in the first or second academic year of upper secondary school, so it must have been hard for the students to use, or it might not have been needed in Composition C. The items "can (request)", "tag question", and "could/would you -?", which seem to be used in speaking tasks, were used by only one student each. They used those items in quotes of someone's speech (e.g. *She said, "You have a lot of free time, don't you?"*). The unemployed grammatical items decreased a little, to 15 items. They include nine B1-labeled items, fewer than 12 of them in the second academic year. There would be only a few students becoming able to use the B1-labeled items in the third academic year. The item "S+V+O+C (noun)", which is usually learned in the third academic year of lower secondary school, doesn't seem to have been relevant to the task requirement of Composition C. It is often used for defining persons or things, such as in *"We call him Peter"*, or *"They named the dog Roy"*. The argumentative topic "Should High School Students Have a Part-time Job?" would not have brought out the use of the item.

The extraction of grammatical items by the Rasch model analyses in this study was lenient because the item was extracted even if only one person used it successfully, and because the analyses were carried out on the premise that the success rate was 50 percent. Also, unlike many of the studies concerning the Processability Theory in which multiple times of use or multiple occasions are prerequisite, only one-time use was regarded as an "emergence" in this study. Nevertheless, the numbers of the items extracted as being "acquired" were 18, 17, and 28, out of 71 items in the checklist, which comprised the A2- and B1-labeled items and the others learned in lower secondary school. There must have been some influence of the task requirements in choosing which items to use, but the huge gap between the input of the grammar knowledge and its proficient use of it is undeniable.

Also, it was revealed that there was a time lag between the learning of the grammatical items and the proficient use of them. Almost all the items extracted from the compositions written in the first and second years of upper secondary school were the ones learned in the first and second years of lower secondary school. It was not until in the third year that the students gained a command of a few items learned in the third year of lower secondary school.

5.2 High-Frequency Grammatical Items in Each Writing Task and Development of "Task-Independent" Grammatical Items

Since the main subjects wrote each one of the different essays in different academic years, it was not certain whether the emergence of a particular grammatical item was due to the learner's development, or due to the task requirement. The data from the students of another upper secondary school, who wrote all the three essays at the same time helped make it clearer which particular grammatical items were used in which writing tasks. First, I scrutinized the frequency data of use of listed grammatical items from the students of another upper secondary school, and analyzed the relationship between the topic of each task and the highly-frequent items used in it. Then, I evaluated again the trajectories of the "acquired" grammatical items of the main subjects while considering the task-dependent items which were identified through the statistical process.

5.2.1 High-Frequency Grammatical Items in Composition A

The topic of Composition A was "The Key Person in My Life", and there were several items that seemed to have strong relevance to the topic. First, compared to the other two tasks, third-person singular verbs, "S+V+O+O", and "S+V+O+C (adj.)" structures appeared most frequently in this task. This is probably because the description

of the key persons in their life required more use of the verb forms of the third-person singular, and because the "S+V+O+O" and "S+V+O+C (adj.)" structures were needed more to write about their relationship with those persons (e.g. *She gave me some good advice*. / *They make me happy when I am sad*.). In addition, the present progressive form was also used more frequently than the other tasks since students used it to refer to habitual actions of/with the person (e.g. *We are going shopping together*.). The objective personal pronouns, such as *me*, *you*, *him*, *her*, *us*, and *them*, were also used much more frequently in this narrative task, probably because the students needed to express thematic relations of the people (e.g. *I respect him very much*. / *She always helps me*.). They were almost always used in the "S+V+O+O" and "S+V+O+C(adj.)" structures, as well as in the canonical "S+V+O". On the contrary, these objective personal pronouns were the least frequent in Composition C, the argumentative essay. This might be because, compared to the other two topics, the argumentative topic elicited more general statements, not stories about particular individuals.

5.2.2 High-Frequency Grammatical Items in Composition B

The topic of Composition B was "My School Life of Last Year". Thus, it was natural that past tense verbs and past tense *be*-copulas were most frequently used in this task in order for the students to express the past events and experiences (e.g. *I really enjoyed the school festival*. / *I was a member of Class 1-3*.). The past progressive form was also used quite frequently in order to express the habitual actions in the past, even though the percentage was not as high, compared to those of past tense verbs and past tense *be*-copulas (e.g. *I was always chatting with my friends after school*.). Interestingly, the adjectival use of a prepositional phrase (following a noun phrase) was also the most frequent in this task. Scrutinized closely, the essays written by its users turned out to

contain prepositional phrases of possessive meaning, such as *"The program of our class (in the school festival) was an amusement park"*, and of attributive meaning, such as *"The teachers in our school were all good"*. Those kinds of expressions might have been needed to elaborate on events, things or people concerning their past experience.

5.2.3 High-Frequency Grammatical Items in Composition C

The students wrote an argumentative essay to express their affirmative or negative opinions about the topic, "Should High School Students Have a Part-time Job". One of the most frequently used items in this task was the "S+V+that-clause" structure. It seems natural because they needed to express their opinions in their arguments using this structure, as in *"I think that high school students need money to buy clothes"*. Some modal auxiliary verbs, such as "may/can/might (possibility)", "must (obligation)", "should (advice)", and "have to-" were also more frequently used in arguments, probably because the students chose to use those items to make value judgments about the proposition given (e.g. *Some students may have to support their family*.). Also, the use of "if-adverbial subordinate clause" was more frequent in this task than the other two. There seems to be a good deal of reason in that because arguments very often require to give conditional or hypothetical cases as affirmative or negative reasons to support them (e.g. *High school students should work if they have enough free time*.). In addition, the "subjective use of gerund" was much more frequent than the other two tasks. The demand for making an assertion about a certain action probably elicited the use of this structure, as in *"Making money is important to help our parents"*.

5.2.4 Development of "Task-Independent" Grammatical Items Extracted from the Longitudinal Data of Writing

As shown together in Table 18, the Rasch analyses extracted the "acquired" grammatical items from the compositions written by the main subjects (209 upper secondary school students). These compositions were written over the course of three years, with the topic of the writing task being different each year. On the other hand, another group of upper secondary school student (108 first-year, 99 second-year, 91 third-year) wrote the three compositions at the same time in the same year, and from their data, some task-dependent grammatical items came to light (Table 13). When eliminating those task-dependent items from the transitional "acquired" items in Table 18, the genuine task-independent items were left in the list (Table 25 below). Naturally, the numbers greatly declined, but the developmental tendency seemed to be similar to that of the items in Table 18: a rapid increase was found in the third year. Also, a progressive accumulation was seen from the first year to the second year to the third year. In the first year, the acquired grammatical items were "S+V+O", "S+*be*+adjective", and "will". In the second year, students were also able to use "because" in addition to the previous three items. In the third year, they were able to use five more items ("there is/are", "gerund as object", "comparative", "it is … for- to-", "relative pronoun").

This pattern of progression indicates that there would be certain delay before the language input becomes the learners' language repertoire to be used in their output, and that the manipulatable repertoire would grow exponentially. First, the learners would learn the grammatical items and practice with them explicitly in class. Then, while receiving accumulative exposure of the grammar explicitly and implicitly, the learners would manage to acquire a command of grammar after necessary intervals, just like a run-up before jump. The more recent the grammatical items are learned as input, the

longer they would have to wait until they are used in output. Therefore, at least in the case of the main subjects in this study, the B1-labled grammatical items, none of which emerged as being "acquired", might become manipulatable in the latter period of the third year of upper secondary school, or after the graduation.

Table 25

Trajectories of task-independent "acquired" grammatical items over the three years

1st-year items	LSS	2nd-year items	LSS	3rd-year items	LSS
S+V+O	*1	S+V+O	*1	S+V+O	*1
S+*be*+C (adj.)	1	S+*be*+C (adj.)	1	S+*be*+C (adj.)	1
will	2	will	2	will	2
		because (conj.)	2	because (conj.)	2
				there is/are	2
				gerund (as obj.)	2
				comparative	2
				it is … for- to-	3
				relative pronoun (exc. whose)	3

*1: not listed as target grammar, but learned in the first year of lower secondary school

5.3 Developmental Sequences of Particular Grammatical Items

5.3.1 Development of Relative Pronoun Clauses

In Chapter 2, I mentioned some of the previous research on the developmental sequence of particular grammatical items. One of them was on relative pronoun clauses. According to the research, referred to in Doughty (1991), parts of speech that relative pronouns represent in relative clauses show the developmental sequence of acquisition: a clause in which a relative pronoun serves as a subject is the easiest to acquire, and the most difficult is a clause in which a relative pronoun serves as an object of "*than*" (Table

2). When examining the development of use of relative pronouns in the longitudinal composition data of this study in this framework, there were several characteristics observed (Table 26). In addition to the increase of total number of examples through the three years, a wider variety of relative clauses were observed in the second and third years, compared to the first year, just at the beginning of the upper secondary school. In the first academic year, there were only eight examples, and all of them were a relative clause in which a relative pronoun serves as a direct object, in such a sentence as "*I have many memories which I can't forget*". In the second and the third years, however, the use of subjectival relative pronoun gradually closed the gap. Moreover, there was only one example in which a relative pronoun was used as an object of preposition: *We always talk with someone with whom we have talked before*. There were no examples of a relative pronoun as an indirect object, a possessive relative pronoun (*whose*), or a relative pronoun as an object of comparison through the three years. According to the research mentioned above, a subjectival relative pronoun is said to be the easiest for learners to access. However, the data of present study seems to show that it is a direct-objectival relative pronoun that is the easiest to access. The subjectival use seems to be following it in the three-year acquisition path. Regarding this point, I asked several lower secondary school teachers of English which relative pronoun they think is the easiest for their students to learn. A little to my surprise, all the teachers said that it is a subjectival use of relative pronoun, which seems to support the findings of the previous research mentioned in Doughty (1991). However, on further inquiry, I learned that it is fill-in-the-blank type of questions or combining-clause tasks that their students do well in with subjectival relative pronouns, and that the teachers rarely give the opportunities, such as free compositions where the students use relative pronouns at their will. I suppose that especially in free compositions, it is easier for the students to use an objectival relative clause or zero

relative clause, because the basic S+V structure remains even in the modified clause.

Table 26

Distribution of different types of relative clauses used in the three compositions

Part of Speech	Composition A (1st year)		Composition B (2nd year)		Composition C (3rd year)	
Subject	0	0.0%	10	43.5%	22	46.8%
Direct object	8	100.0%	13	56.5%	24	51.1%
Indirect object	0	0.0%	0	0.0%	0	0.0%
Object of preposition	0	0.0%	0	0.0%	1	2.1%
Possessive	0	0.0%	0	0.0%	0	0.0%
Object of comparison	0	0.0%	0	0.0%	0	0.0%
Example total	8		23		47	

The examples of relative clauses found in the three longitudinal compositions show not only quantitative development, but also qualitative development, when looking at the variation of structural connections between main clauses and relative clauses. As shown in Table 27, in Composition A, written in the first academic year, only two types of combination were observed: the O-O and S-O types. Among the only eight examples of relative structures, six of them (75.0%) were the O-O type. In Composition B in the second year, the variations of relative structures increased to seven, with the increase of total number of examples. In addition to the types listed in Yule (1998), three other types were found in Composition B. The first one, “exS-S” is the combination of an existential head noun, or a logical subject, and a subjectival relative clause in the *there* structure (e.g. *There are many students who join the club activities*). Another one, “exS-O” is the combination of an existential head noun, or a logical subject, and an objectival relative clause in the *there* structure (e.g. *There are a lot of things which we should do instead of*

working). The other one, "C-O" is the combination of a complement and an objectival relative clause, in such a sentence as "*This is one step which we will take in society in the future*". In Composition C, written in the third academic year, the variations of combination remained seven types, including the additional three, while the total number of examples was doubled to 47 from 23 in the second academic year. If the frequencies of relative clauses reflected their development, this result would not correspond with the difficulty order of relative clauses shown in Yule (1998) and Celce-Murcia and Larsen-Freeman (1999). In the productive tasks of this study, at least, it seems that the students first began to be able to use the O-O and S-O combination, and then extended the variations of relative structures. However, throughout the three years, the relative clauses in final position account for about 70 to 75 percent of the total examples, which seems to correspond with the claim in Yule (1998) that the relative clauses in final position are more successfully learned to use than those in medial position.

Table 27

Variations of main-relative clause combination in the three compositions

Position	Combination	Composition A (1st year)		Composition B (2nd year)		Composition C (3rd year)	
Final	O-S	0	0.0%	6	26.1%	9	19.1%
	O-O	6	75.0%	6	26.1%	12	25.5%
	exS-S	0	0.0%	2	8.7%	5	10.6%
	exS-O	0	0.0%	1	4.3%	7	14.9%
	C-O	0	0.0%	1	4.3%	2	4.3%
Medial	S-S	0	0.0%	2	8.7%	8	17.0%
	S-O	2	25.0%	5	21.7%	4	8.5%
	Example total	8		23		47	

5.3.2 Development of Future Tenses

As mentioned in Chapter 2, concerning the use of the future expressions, Bardovi-Harlig (2002) found that learners begin to use "will" quite early, and that "going to" is first used in formulaic expressions and then comes to be manipulated at will. When looking at the use of "will" and "going to" in the three compositions in this study, there seem to be some points in common with the findings of previous study. If we consider that the more frequent a particular grammatical item is used, the easier it is for learners to use, "will" seems to have been much easier for the students in this study to express the future events or intentions than "[be] going to" all in the three compositions. As shown in Table 28 below, the percentage of use of "will" was over 90 percent in each academic year, even though the total examples of these two items increased every year. On the other hand, there was one example of "[be] going to" found in each year which seems to have been formulaically used in the beginning of each writing task, such as "*I am going to write about ...*". The other one in Composition A, two in Composition B, and four in Composition C are used in the sentences constructed on the writers' volition. Overall, even in the sentences where the students would like to express future certainties or future intentions, they might have chosen "will", because it was just simple, or because they might not have known well how to use "[be] going to" in non-formulaic sentences. The true reasons for their choice could not be figured out without interviewing them. However, such a high percentage of "will" throughout the three years would hint that many students might have substituted "will" for "[be] going to" in many cases, because they did not fully understand the difference between these two future expressions.

Table 28

Frequencies of Future Expressions

	Composition A (1st year)		Composition B (2nd year)		Composition C (3rd year)	
will	47	95.9%	61	95.3%	74	93.7%
be going to	2	4.1%	3	4.7%	5	6.3%
Example total	49		64		79	

5.4 From Formulaic to Extensive Use of the Particular Grammatical Items

While checking the grammatical items used in the main subjects' compositions, I noticed that in several structures, the students tend to use the same words (verbs or adjectives), which makes those structures look like formulaic expressions for the users. If there can be a sign of transition from formulaic use to extensive or systematic use of a structure observed through the three years, the development of processability of the structure might be admitted. I looked over all the examples of particular items used in the students' compositions with eyes again, and sorted those examples according to the constituent words of items. More than one example used by one student were counted, so the numbers of examples counted can be larger than the numbers of users.

5.4.1 Infinitive (object)

The first case is infinitive as an object of a verb. Table 29 displays the variation and frequencies of verbs preceding objectival infinitives that were found in the students' compositions in each of three years. Each entry of verb includes different conjugations, such as 3rd-person singular, past tense, or present perfect. Throughout the three years, "*want*" was used by far most frequently. The structure "*want*+to-infinitive" is usually

learned in the second year of lower secondary school, when it is introduced as the very first example of infinitive (Table 3). The students seem to have acquired "*want to*" as a chunk and become used to manipulating it. From the first year to the second, the number of examples increased greatly (62 to 99), but there was no increase exhibited in the variety of verbs (7 to 6). However, from the second year to the third, while the number of examples slightly decreased (99 to 94), there was a perceptible increase in the variety of preceding verbs (6 to 10). What is also noteworthy is that the most frequent "*want*+to-infinitive" (61.7%) yielded some percentage to the "*need*+to-infinitive" structure (14.9%). That might be because in the argumentative writing task of the third year, the students thought about the topic more objectively and used "*need*" naturally in some occasions, instead of "*want*". In the narrative essays, on the other hand, they used "*want*+to-infinitive" more in order to expressed direct or personal desire.

In the view of Processability Theory, learners are judged to have acquired a systematic use of particular grammar when they can manipulate it using different words or morphemes. Although the formulaic use of structure still remains strong-rooted and there might be some influence of the task difference, there seems to be a sign that the students have been becoming more proficient at using infinitive as a verbal object, with more frequency and with more variety of verbs.

Table 29

Frequencies of objectival infinitive with different verbs

	Composition A (1st year)		Composition B (2nd year)		Composition C (3rd year)	
want	54	87.1%	86	86.9%	58	61.7%
like	2	3.2%	3	3.0%	2	2.1%
begin	1	1.6%			2	2.1%

start	1	1.6%			4	4.3%
need	2	3.2%	3	3.0%	14	14.9%
hope	1	1.6%	1	1.0%	1	1.1%
forget	1	1.6%				
decide			3	3.0%	4	4.3%
try			3	3.0%	6	6.4%
continue					1	1.1%
learn					2	2.1%
Example total	62		99		94	
Verb variation	7		6		10	

5.4.2 S+V+O+O

Another structure in which the students tended to cling to particular constituents is the ditransitive clause, or the double object construction (S+V+O+O). It is labeled as an A2 feature in Hawkins and Filipović (2012), and the students learned the structure for the first time in the second year of lower secondary school (Table 3). Most of the cases, the first and representative ditransitive verb that they first learn is "*give*". Table 30 shows the variation of ditransitive verbs and the total numbers of examples of each S+V+O+O structure, found in the compositions that the students wrote in three academic years. Multiple use of the same ditransitive verb in the same composition was counted as one. When looking at the total frequencies of ditransitive clauses that they used in their compositions of each year, the first encounter with "give" seems to have long affected their grammar manipulation. There does not seem to be progress in the variation of verbs through the three years. In every task, "*give*" and "*teach*" account for 75 to 90% altogether. Still, some students were able to use some other verbs in the S+V+O+O structure. On the other hand, there was a big difference in the total numbers of examples of the S+V+O+O

structure between the three writing tasks. The topic of Composition A was "The Key Person in My Life", and the students may have had more opportunities to use the item when writing about the relationship with the key person in their life, in such a sentence as "*He gave me good advice*". In Composition B, the students wrote about their school life of the previous year, and some of them may have often used the item in their personal anecdotes about what the people around them did to them, or what they themselves did to the other people, in such a sentence as "*The teacher taught me how to play the guitar*". However, in Composition C, the students wrote about their opinions about an argumentative topic, and there may have been less room for referring to the relationship with a particular person or their personal anecdotes, compared to Compositions A and B. As a consequence, in Composition C, the total number of S+V+O+O examples decreased, and there were only a few variations in the use of ditransitive verbs constituting the S+V+O+O structure. On the whole, although the imprint of the early-learned formulae appeared to be rather persistent, the students seem to have been able to vary the ditransitive verbs in the S+V+O+O structures to suit the task requirements up to a point.

When examining formulaic sequences inside the structure formed by the verb "*give*", "*give me*" and "*gave me*" account for the majority of the examples in Compositions A and B. There were no "*gives me*" examples found in all the three compositions. Although it might be arguable whether "*gave me*" can be judged as a formulaic chunk since "*gave*" is a conjugated form of "*give*", the number of students who used only "*give me*" or only "*gave me*" sequence as the S+V+O+O constituent amount to 72.0% of all the "S+[*give/gave*]+O+O" examples found in Composition A (18 out of 25 examples), and 72.2% in Composition B (13 out of 18 examples). Thus, "*give/gave me*" might have become a prefabricated pattern that can be easily drawn and used in the students' memory. In Composition C, on the other hand, there were only two examples

of "*give me*" out of 13 examples, and there were no "*gave me*" examples. Instead, there were eight examples of "*give us*" and one "*give you*" found in Composition C. That might be because in the argumentative essay, the students wrote their opinions, often referring to people or students in general as *we (us)* or *you*.

Table 30

Frequencies of S+V+O+O with different ditransitive verbs

	Composition A (1st year)		Composition B (2nd year)		Composition C (3rd year)	
give	25	40.3%	18	56.3%	13	61.9%
teach	22	35.5%	6	18.8%	6	28.6%
send	6	9.7%	1	3.1%		
make	4	6.5%	1	3.1%		
tell	3	4.8%	2	6.3%	2	9.5%
show	1	1.6%	2	6.3%		
buy	1	1.6%				
ask			1	3.1%		
lend			1	3.1%		
Example total	62		32		21	
Verb variation	7		8		3	

5.4.3 S+V+that

The next item that takes on a formulaic complexion is S+V+that-clause. This structure is labeled as an A2 level feature of the CEFR (Hawkins & Filipović, 2012), and the students learned this structure in the second year of lower secondary school, as shown in Table 3 in Chapter 3. Table 31 shows the frequencies of S+V+that-clause with different predicate verbs. Since this structure was often used with the same verb more than one

time in the same composition, the multiple examples with the same verb were counted as one. Still, through the three years, "think" was used in over 80% of all the S+V+that examples as a predicate verb. As the total number of examples increased from 98 to 183, there appeared more and more variety of verbs that constitute the structure year by year (5 to 13 verbs). It seems that this grammatical item became more and more processable enough for the students to use it at their will in production tasks.

The formulaic sequence within this item that was distinctively frequent is "*I think* [*that*]" or "*I thought* [*that*]". Some might think that "*I thought* [*that*]" is less chunk-like compared to "*I think* [*that*]" because the verb is conjugated, but as in the case of "S+V+O+O", in this study, the present tense and past tense of the verb were not discriminated when counting them as chunk constituents. However, if there was any word between *I* and *think/thought* (e.g. *I also think*, *I don't think*), it was not counted as the formulaic expression. The redundant use of the "*I think/thought*" structure, in such a sentence as "*I think* [*that*] *I want to help my parents*", was counted as an "*I think*" chunk. However, if it was added after the end of sentence (e.g. *We should have a part-time job, I think*.), it was not counted as the formulaic sequence. This additional use of "*I think*" in the sentence final could be an evidence of its formulaicity, but it does not form the S+V+that structure.

When counting the number of students who used only "*I think* [*that*]" or "*I thought* [*that*]" sequence as the S+V+that constituent, its percentage to the total number of S+[*think/thought*]+that users went 86.0% (the first year), 86.7% (the second year), and 77.8% (the third year). It seems to be natural that in many sentences of personal or opinion essays, the agent of the verb "*think*" tends to be "*I*". Still, it is interesting that in the third year compositions, more students showed different use of the S+[*think/thought*]+that structure (e.g. different subjects, negative, etc.) than in the first and second years.

Table 31

Frequencies of S+V+that structure with different verbs

	Composition A (1st year)		Composition B (2nd year)		Composition C (3rd year)	
think	86	87.8%	98	85.2%	153	83.6%
hope	9	9.2%	4	3.5%	2	1.1%
wish	1	1.0%			1	0.5%
decide	1	1.0%	1	0.9%		
hear	1	1.0%				
feel			3	2.6%	1	0.5%
know			3	2.6%	10	5.5%
believe			2	1.7%		
expect			1	0.9%		
consider			1	0.9%		
mean			1	0.9%	3	1.6%
learn			1	0.9%	6	3.3%
agree					2	1.1%
notice					1	0.5%
understand					1	0.5%
realize					1	0.5%
tell					1	0.5%
say					1	0.5%
Example total	98		115		183	
Verb variation	5		10		13	

5.4.4 S+V+O+C (adjective)

One more item that the students seem to have used somehow formulaically was the

S+V+O+C (=adjective) structure. This item is labeled as a B2 feature in Hawkins and Filipović (2012), so it is judged as a rather difficult item for L2 learners to use in their written tasks. However, the students had already learned the structure in the third year of lower secondary school, as shown in Table 3 in Chapter 3. They can use different verbs in this structure, and the slot for a complement can also be filled with a variety of adjectives. Table 32 shows the frequencies of combinations of verbs and adjectives that the students chose when they used the S+V+O+O structure in their three compositions. Almost all the adjectives used were emotive, such as "*happy*", "*sad*", and so on. It stands out that the frequency of the structure extremely decreased in the third year. It might be because Composition C, written in the third year, was an argumentative essay in which the students wrote their opinions in an objective manner, so that would have held back the use of emotive adjectives. On the other hand, from the first year to second year, in the same narrative essays, there appeared to be some progress in the total number of examples and the adjective variations.

However, some students seem to have stored the item as a particular formulaic sequence. As shown in Table 28, through the three years, the most frequent combination was "*make/made*" and "*happy*", which makes the S+[*make/made*]+O+[*happy*] structure. Among those S+[*make/made*]+O+[*happy*] examples, "*make me happy*", "*makes me happy*", and "*made me happy*" comprised high percentages: 40.9% in the first year (9 out of 22 examples), 63.6% in the second year (14 out of 22 examples), 80% in the third year (4 out of 5 examples). As mentioned above, the absolute number of S+V+O+C examples was small in the third-year compositions, so the highest ratio of "*make/makes/made me happy*" sequence (80%) cannot be compared with those in the other two academic years. In the second year, when the students wrote about their school life of the previous year, many of them showed a strong tendency toward writing about pleasant memories. That

would have promoted the use of "*make/makes/made me happy*" sequence.

Table 32

Frequencies of combinations of verbs and adjectives in S+V+O+C

		Composition A (1st year)		Composition B (2nd year)		Composition C (3rd year)	
make	happy	22	75.9%	22	56.4%	5	71.4%
	cheerful	1	3.4%				
	angry	1	3.4%	1	2.6%		
	powerful	1	3.4%				
	fine	1	3.4%				
	sad			4	10.3%		
	bright			3	7.7%		
	dark			1	2.6%		
	unhappy			1	2.6%		
	strong			1	2.6%		
	glad			1	2.6%		
	true			1	2.6%		
	surprised			1	2.6%		
	funny			1	2.6%		
	good			1	2.6%		
	interesting					1	14.3%
	comfortable					1	14.3%
keep	happy	1	3.4%				
	cheerful			1	2.6%		
think	important	2	6.9%				
	Example total	29		39		7	
combination variation		7		13		3	

5.5 Development of Sentence Complexity and Its Relationship with Fluency and Overall Writing Proficiency

5.5.1 The MLUs and Word Counts

The development of the MLU was confirmed as it had already been found in Murakoshi (2015), though even the average MLU of their third year (=8.8) stayed at the A2 level. In Murakoshi (2015), I examined the use of subordinate structures as the MLU enhancement factor. Taking another look at the "acquired" items in Table 18, I recognized the accumulative progression of use of subordinate structures anew (Table 33). This progress would contribute to the constant development of the MLU.

Table 33

Progress of "acquired" subordinate structures over the three years

1st-year items	2nd-year items	3rd-year items
S+V+that	S+V+that	S+V+that
when (conjunction)	when (conjunction)	when (conjunction)
	because (conjunction)	because (conjunction)
		if (conjunction: adverbial)
		relative pronoun (exc. whose)

Seemingly, the development of the MLU was steady. However, the standard deviation (*SD*) of the third year was the largest and the minimum value was the smallest. That indicates the difference between individuals also widened, while the average MLU progressed steadily.

As for the total word counts, they did not develop in a linear manner: there was a

decrease from the second to the third year. One of the reasons may be that in Composition B in the second year (writing about their school life), it was probably easier for the students to write about their past experiences, enumerating facts and their own feelings. On the other hand, in Composition C in the third year, the argumentative logic might have limited their free description. They had to write their assertion and the reason(s), but not every one of them were able to deepen their arguments by giving examples or attempting refutations.

When looking at the relationship between the MLUs and the total word counts of the three academic years, the strength of correlation showed gradual development. The correlation coefficient went from .05 (NS) in the first year to .30 (weak) in the second year to .48 (moderate) in the third year. This could mean that the more time students studied English, the more balanced their writing skills became in terms of complexity and fluency. In the process of acquiring writing proficiency, much of students' attention would be paid to communicating their thoughts in the early stage. Students would write what occurs to them in shorter sentences using their limited language repertoire. Later, they would become more and more conscious of language forms in their writing. As their linguistic knowledge increases, they would consciously try to use more complex structures, such as subordination or modification. Then, their writing would be well-balanced —communicative in content and sophisticated in language.

Still, even in the third year, the strength of the correlation was just moderate, so there were some outliers to be aware of. As an example, one sample of Composition C had a larger MLU than the average by 2.5 points (words), but its word count was less than one-third of the average. The following is that sample:

I think that we should work. I think that we will learn various things through work experience. I think that I will enjoy working, because I think that many people will become my friends.

(34 words, MLU=11.3)

Although some complex structures can be found in this composition and the accuracy was relatively high, the same structure was used repeatedly. The writer might have had some explicit grammatical knowledge to write longer sentences, but might not have had enough discourse competence or motivation to write a longer passage or communicate her/his ideas.

At the opposite extreme is the following sample. The total number of words is 1.7 times as large as the average, but the MLU is smaller than the average by 1.7 points (words).

I agree with this title because it is important for us to work. I am working in restaurant now. It is very hard work. However, it is very interesting for me. I learned many important things. However, we must study now. We became 3 grade this year. Study is more important than to work. If we passed the test, we should begin work. However, money is important too. So, this title is very difficult for me. Study is important. Money is important. Both things are very important for me. I don’t decide only one. I want money. I can’t buy clothes, food, and other things if I don’t work. My mother said, “Why do you want to work?” I said, “Because I want money.” She said, “I’ll give you money.” I was very happy then. However, I don’t like to receive money from my mother because it is my mother’s money. My mother is working now. I want to use only

my money. So, I can understand important of money. If I don't work, I can't understand important of money. So, working is very important things.

(193 words, MLU=7.1)

Some short sentences with three to five words are found here and there. Another characteristic is that the same discourse markers are used repeatedly to develop his/her argument, such as "However" and "So". Also, some redundancy of statement was seen (e.g. *Study is more important ... However, money is important too ... Study is important. Money is important ...*). This is probably because this student just wrote his/her opinions and experiences as they occurred to him/her. It is assumed that not enough time was spent planning for the writing in this case.

5.5.2 The MLUs and the GTEC Writing Scores

The strength of correlations between the students' MLU values and the scores of GTEC writing test might be able to reveal another aspect of the development of their grammatical knowledge for constructing more coherent and logical sentences. As mentioned in Chapter 3, the GTEC writing test that the students took was assessed in terms of opinion, reason, vocabulary, grammar, and organization. The MLU, or the average of sentence length, might be concerned with expressing constructive opinions or giving persuasive reasons in essay writing, because if the writer wants to develop logical argument, he or she would need to use complex sentences with subordinate clauses which would help develop the MLU values. As shown in Chapter 3, regarding Composition A, written in the first academic year, there was no correlation found between the MLU and the score of GTEC writing test (r = - .04, p = .61). In Composition B, written in the next academic year, there was still no correlations between the MLU and the GTEC writing

score (r = .05, p = .48). Finally, in Composition C, written in the third academic year, the MLU had weak correlation with the score of GTEC writing test (r = .28*, p = .00). As I already mentioned in Chapter 3, the MLUs and the GTEC scores of each academic year were elicited at almost the same time of the year, but from different pieces of writing. However, if both data show certain aspects of the students' proficiency at the same time of each academic year, the transition of their correlations through the three years would imply something about the students' linguistic development. The correlation coefficient transitioned from - .04 to .05, and finally to .28* in the third year, with statistical significance. This might indicate that the students' efforts to write longer and more complex sentences moderately but gradually led to success in describing their thoughts more logically and persuasively.

5.6 Pedagogical Implications

One of the reasons that the input of the grammar was not effectively reflected in the output may be that the grammatical items were not taught in the context in which they are often used, but were taught explicitly as independent linguistic forms. The students might have been good at answering fill-in-the-blank grammar questions, but could not have retrieved the items when they had to expressing themselves in language production. Another reason may be that the students had not had many opportunities to use their grammar knowledge, where they could analyze the task requirement, choose the grammatical items and vocabulary words needed, and use them in appropriate forms in production tasks. Nassaji and Fotos (2011) argue that implicit or explicit instruction should be supported by the provision of discourse-level input to expose learners to repeated use of target forms in natural input, and that learner discourse-level output

producing target forms are also essential to promote learners noticing and ultimate acquisition of the target structures (p.51). They suggest that L2 grammar instruction, either explicit or implicit, should take place in extended contexts rather than in isolation, and should include opportunities for learners to receive meaningful input and to produce meaningful output containing the target form (p.52). Therefore, teachers of English need to understand and teach the context in which each grammatical item is often used, using natural and practical examples that are as relevant to the students' life as possible. Moreover, it should be important to deal with the same grammatical items over and over again, sometimes with certain intervals, because only one-time instruction cannot be expected to guarantee the acquisition of those items. Also, many writing opportunities should be given in class in order for the students to try to use their linguistic knowledge at will and acquire it more steadily through trial and error.

Then, it should be taken into consideration that particular types of tasks may induce learners to use particular grammatical items. In the follow-up study, nearly 40 percent of the checked grammatical items (28 out of 71 items) were proven to be "task-dependent". Larsen-Freeman (2003) points out that it is important to design activities so as to elicit particular grammar structures. In class at secondary schools, topics of writing tasks tend to be chosen based on themes of reading passages in textbooks, or students' experiences, such as summer vacation or school events. As well as topics, teachers should pay more attention to text types or genres of writing tasks. A wide variety of tasks should be given to learners so that they can try to use as various grammatical items as possible.

Teachers should not only focus on classroom activities, but also the course syllabi as well. The traditional grammatical syllabus still lies beneath the procedures of editing English textbooks, explicitly or implicitly. However, it might not necessarily match the actual developmental sequence of Japanese EFL learners' linguistic knowledge. For

example, in the English Grammar Profile, part of the English Profile Programme, the passive structure is listed as A2 to C1 items, depending on its usage. The numbers of its entries are, three as A2, nine as B1, 14 as B2, three as C1, and two as C2. In Core Inventory for General English, it is listed as a B2 item to teach. However, in Japan, it is usually taught in the second year of lower secondary school, when many students are not expected to have reached A2 level yet, to say nothing of B2 level. That may be why its incorrect use is often seen in compositions written by lower secondary school students (e.g. *Tennis is played by Tom). Those students might not be ready yet to use the structure in a more natural context. Like the English Profile Programme, collecting learners' output data as much as possible and analyzing it would help design more practical syllabi for teaching English at secondary education. In that respect, the CEFR-J Grammar Profile will play a much more important role in constructing more effective syllabi and editing textbooks or classroom materials, reinforced by more output data in the future. In the CEFR-J, the lower levels are branched into the subdivisions, taking the situations of English education in Japan into consideration (e.g. A2.1, A2.2). The CEFR allows a common set of levels and/or descriptors to be cut into practical local levels at different points in accordance with local needs. As the CLC has been used for extracting the criterial features that discriminate the CEFR levels, more output data from different levels of Japanese EFL learners could contribute to extracting criterial features that discriminate finer branches of the CEFR-J. That would help teachers design more practical syllabi, teaching plans, language activities, and test items. That would also help students set more concrete goals of learning English. Furthermore, that would be so much helpful to textbook publishers in deciding what grammatical items to pick and how to arrange them in their products.

The detection of grammatical items used in learners' language production has a

potential for predicting learners' writing proficiency. Many English teachers in lower and upper secondary schools have trouble assessing their students' compositions. One of the reasons may be that they think assessment of writing tends to be subjective and unreliable. That thought would sometimes make teachers hesitate to introduce writing activities or tests in their teaching. Also, the thought would make it difficult for teachers at upper secondary schools to prepare a writing task in entrance examinations to their schools. Teachers would prefer the only one correct answer to any question because it is easier to mark. Then, that would give negative backwash effect to teaching writing in lower secondary schools. If writing is not part of entrance examination to upper secondary school, students would lose interest in learning how to write, and teachers would focus on other teaching points than writing. Another reason may be that teachers think assessing writing is time-consuming. That would also make them reluctant to give their students many opportunities of writing. They might think that time-consuming writing assessment would deprive them of their limited time and energy unreasonably. If an objective check of use of particular grammatical items could reliably predict learners' writing skills, it would be greatly helpful to English teachers. In order to realize that new measuring method, it is essential to collect a large amount of composition data from learners in various proficiency levels, and investigate the correlation between the emergence of particular grammatical items and the holistic human ratings. Also, it would be necessary to more clarify the relationship between particular grammatical items and discourse that those items tend to create. That clarification of the relationship between grammar and discourse would make more simple application possible: the detection of particular grammatical items could be used as part of analytic rating of writing. Using rubrics for assessing writing has been more and more common in lower and upper secondary school. Many teachers would like to assess accuracy of their students' compositions as one aspect

of writing, but they would sometimes be in trouble deciding what to assess or how to assess it. More limited and focused check of grammatical items would help teachers carry out more effective and more economical assessment of writing.

5.7 Limitations and Outlook for the Future Research

The present study came to a tentative conclusion, but there were some limitations to highlight. However, these limitations can provide hints for future research.

First, the view of the "acquisition" of the grammatical items might have been too absolute. In this study, in order to conduct the Rasch model analyses, the checklist of grammatical items was used, regarding their use or disuse as a correct or incorrect answer of a test. Only the correct use of an item was regarded as being "acquired", so any interlanguage or attempt at using the language was dealt with as being "incorrect". That assumption would have led to a kind of rough view of the acquisition of the grammar knowledge. Future research could try to analyze the learners' attempt to use their grammar knowledge. In relation to that, learners' errors could also be worthy of investigation. It may be interesting to track the transition of a particular error through a certain period of time. Some might be corrected relatively in a short while, but others could be rather persistent, such as the isolated subordinate "because", which was investigated in Murakoshi (2012a).

Second, the detection of grammatical items may not have been entirely reliable. The compositions written by the main subjects of this study were not given as tools for detecting the "acquired" grammatical items, but just as writing tasks which were part of the pedagogical syllabus. Therefore, in the composition tasks, there were not multiple obligatory contexts that promoted the learners to use particular grammatical items. Even

if the student used the item just once, the emergence was regarded as being “acquired”, though the item could have been produced only by chance. It would be helpful to teachers and researchers if more diagnostic and effective writing tasks were invented that could assess learners’ grammar knowledge more accurately and fulfill pedagogical objectives as well.

Moreover, there is room for consideration regarding the checklist items. Basically, they comprise the CEFR-labeled grammatical items extracted from learner corpora and the target grammatical items in EFL textbooks. Even though some grammatical items were not used in the compositions by any of the students, it is not certain whether that was because the students were not proficient enough to use them, or because those items are rarely used even by native speakers of English after all. Thus, visiting a corpus of native speakers’ written production would be helpful for more careful selection of the checklist items.

Finally, the results and conclusion may not be applicable to the general population of Japanese upper secondary school students. The main subjects of this study belonged to the same upper secondary school, so their proficiency in English spanned within a narrow range. In order to gain a more general view of students’ grammatical acquisition, more data should be elicited from learners of different proficiency levels.

Chapter 6. Conclusion

This study identified the grammatical items that the Japanese upper secondary school students have acquired and can use at will in written production tasks, through the Rasch model analyses. Whether the task-dependent items were taken into consideration or not, it was not until in the third year of upper secondary school that the items learned in the third year of lower secondary school emerged as "acquired" grammatical items. The writing task was given early in the academic year, so at least, up until the early time of their second year, the students managed to write their English compositions using the grammatical items that they had learned in the first and second years of lower secondary school. The MEXT's goal of English education for upper secondary school students is acquiring the CEFR B1 level proficiency by the time of graduation. Supposing that grammar knowledge is considered to be part of this "proficiency," the fact that no B1 grammatical items were judged as "acquired" in the students' compositions, may be a testament to the issues of teaching practice in English class in Japan. As for the sentence-level proficiency, the mean length of utterance (MLU) was proven to be still at the upper A2 level of the CEFR. However, these two indicators of proficiency —the manipulatable grammatical items and the MLU showed the students' certain progress throughout their three academic years of upper secondary school, even though the changes were rather gradual.

I hope that the findings of this study could help improve the approaches of teaching grammatical items in English class of secondary schools. The fact that the grammatical knowledge as learning input is not easy to emerge in learners' output would promote the paradigm of classroom practice to be shifted to context-based and communication-oriented. Also, the application of the grammatical item detection to assessing writing

could motivate teachers to introduce more writing activities in their class, which would give students more opportunities to learn how to construct cohesive and coherent discourse, using more sophisticated and complex sentences.

Lastly, the longitudinal composition data used in this study is so precious and informative that more investigations should be undertaken in further research. It could be utilized as part of learner corpus of Japanese EFL learners. There would be a lot of aspects left to analyze: errors or non-native use of grammar, development of collocations, or vocabulary use. Also, basically in this study, the main subjects were treated as a group of EFL learners, so there is scope for focusing more on the development of individuals. Moreover, the students' answer to a questionnaire could be compared with the results of this study. The questionnaire was given at about the same time as each writing test. It asked about their attitudes toward studying English, study hours, skills they wanted to improve, and so on. Similarly, the composition data from another upper secondary school, which was mainly used to identify task-dependent grammatical items in the follow-up study, deserves another research too. The students in each academic year worked on three compositions at the same time, so they produced more language than the main subjects did. Thus, with their composition data put together, more reliable observation of learners' output would be possible. Their data could also be used as a learner corpus, which would enable us to examine different aspects of their language production more precisely. I feel it my duty to make the most use of the data and the findings for the improvement of the secondary education of English.

References

Andersen, R. W. (1978). An implicational model for second language research. *Language Learning,* 28 (2), 221-282.

AntConc [computer software]

Available at https://www.laurenceanthony.net/software/antconc/

Apple, M. T. (2013). Using Rasch analysis to create and evaluate a measurement instrument for foreign language classroom speaking anxiety. *JALT Journal*, 35 (1), 5-28.

Baker, R. (1997). Classical test theory and Item Response Theory in test analysis, *Language Testing Update*, Special Report No.2, Centre for Research in Language Education, Lancaster University.

Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly,* 26 (2), 390-395.

Bardovi-Harlig, K. (2002). A new starting point?: Investigating formulaic use and input in future expression. *Studies in Second Language Acquisition, 24*, 189-198.

Baten, K., Buyl, A., Lochtman, K., and Herreweghe, M. V. (2015). Processability theory: Assessment of the current state and future perspectives. In K. Baten et al. (Eds.), *Theoretical and Methodological Developments in Processability Theory*. Amsterdam: Benjamins, 1-18.

Beers, S. F. and Nagy, W. E. (2009). Syntactic complexity as a predictor of ado;escent writing quality: Which measures? which genre? *Reading and Writing*, 22 (2), 185-200.

Bond, T. G. and Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Routledge: New York.

Bresnan, J. (2016). *Lexical-functional syntax (2nd ed.)*. Malden, MA: Wiley Blackwell.

Briscoe, E., Carroll, J. and Watson, R. (2006). The second release of the RASP system. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, Sydney, Australia. Stroudsburg, PA: Association for Computational Linguistics, pp.77-80.

Retrieved from: http://www.aclweb.org/anthology/P06-4020 [2019.11.30]

Brown, S. and Larson-Hall, J. (2012). *Second language acquisition myth: Applying second language research to classroom teaching*. Ann Arbor, MI: The University of Michigan Press.

Celce-Murcia, M and Larsen-Freeman, D. (1999). *The grammar book*. Boston, MA: Heinle.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

DeKeyser, R. M. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction*, pp.97-113. New York, NY: Routledge.

DeKeyser, R. M. (2009). Cognitive-psychological processes in second language learning. In M. Long & C. Doughty (Eds.), *The handbook of Language teaching*, pp.119-138. Oxford: Blackwell.

Doughty, C. (1991). Second language instruction does make a difference; Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition, 13* (4), 431-469.

Dyson, B. (2009). Processability theory and the role of morphology in English as a second language development: A longitudinal study. *Second Language Research, 25* (3), 355-376.

Dyson, B. (2010). Learner language analytic methods and pedagogical implications. *Australian Review of Applied Linguistics, 33* (3). MONASH UNIVERSITY ePRESS.

Ellis, N. (2002). Frequency effects in language processing: Review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188.

Ellis, N. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism*, pp. 35-47. Washington, D. C.: Georgetown University Press.

Ellis, R. (2005a). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*, 141-172.

Ellis, R. (2005b). Instructed language learning and task-based teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*, pp.713-728. Mahwah, NJ: Lawrence Erlbaum.

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly, 40*, 83-107.

Ellis, R. (2008). *The studies of second language acquisition (2nd ed.)*. Oxford University Press.

English Profile. (2015). Retrieved from: http://www.englishprofile.org/ [2019.11.30]

Ferris, D. R. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly,* 28 (2), 414-420.

Gass, S. and Selinker, L. (Eds.). (2008). *Second language acquisition: An introductory course (3rd ed.)*. New York, NY: Routledge.

Glahn, E., Håkansson, G., Hammarberg, B., Holmen, A., Hvenekilde, A., and Lund, K. (2001). Processability in Scandinavian second language acquisition. *Studies in Second Language Acquisition,* 23, 389–416.

Hammarberg, B. (1996). Examining the Processability hierarchy: The case of adjective agreement in L2 Swedish. In E. Kellerman, B. Weltens and T. Bongaerts (Eds.), *Eurosla 6: A Selection of Papers. Toegepaste Taalwetenschap in artikelen,* 55/2, 75-88.

Harrison, J. (2015a). What is English Profile? In J. Harrison and F. Barker (Eds.), *English Profile in Practice*. Cambridge: Cambridge University Press.

Harrison, J. (2015b). The English Grammar Profile. In J. Harrison and F. Barker (Eds.). *English Profile in Practice*. Cambridge: Cambridge University Press.

Hawkins, J. A. and Buttery, P. (2009). Using learner language from corpora to profile levels of proficiency: insights from the English Profile Programme. *Language Testing Matters: Investigating the wider social and educational impact of assessment*, pp.158-175. Cambridge: Cambridge University Press.

Hawkins, J. A. and Buttery, P. (2010). Criterial Features in Learner Corpora: Theory and Illustrations. *English Profile Journal,* 1(1), 1-23. DOI: 10.1017 /S2041536210000103

Hawkins, J. A., and Filipović, L. (2012). *English Profile Studies I: Criterial Features in L2 English*. Cambridge: Cambridge University Press.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (Research Report No. 3). Urbana, IL: National Council of Teachers of English.

Hunt, K. W. (1970). Syntactic maturity in school children and adults. *Monographs of the Society for Research in Children Development, 35* (1).

IBM SPSS Statistics (Version 22) [computer software]. IBM Corporation.

Ishii, Y. and Tono, Y. (2016). CEFR-J Grammar Profile no tame no Bunpo koumoku hindo-chosa. [A frequency analysis of grammar categories for the CEFR-J Grammar Profile.] Paper presented at the Symposium on the CEFR-J RLD Project: Developing Grammar, Text and Error Profiles Using Textbooks and Learner Corpora. March 5, 2016. Tokyo University of Foreign Studies.

Klein, W. (1995). The acquisition of English. In R. Dietrich, W. Klein and C. Noyau (Eds.) *The acquisition of temporality in a second language*. Amsterdam: John Benjamins.

Koike, I. (Ed.) (2008). *Leading research to coordinate primary, secondary, and tertiary level English education based on second language research*. Grant-in-Aid for Scientific Research (A): No. 16202010. An unpublished final report and collected papers. Meikai University.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston, MA: Heinle.

Lightbown, P. and Spada, N. (2013). *How languages are learned (4th ed.)*. Oxford: Oxford University Press.

Lim, H. (1983). The development of syntax in the writing of university ESL students. *TESOL Quarterly,* 17, 313-314.

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly,* 45 (1), 36-62.

Mansouri, F. (2005). Agreement marking in Arabic as a second language: typological features and their processing implications. In M. Pienemann. (Ed.), *Cross-linguistic Aspects of Processability Theory*. Amsterdam: John Benjamins.

Meisel, J. M., Clahsen, H., and Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition,* 3, 109-135.

Milanovic, M. (2009). Cambridge ESOL and the CEFR. *Research Notes*, 37, 2-5. Cambridge ESOL. Retrieved from: http://www.cambridgeenglish.org/images/23156-research-notes-37.pdf [2019.11.30]

Murakoshi, R. (2012a). The development of the use of 'because-clauses' by Japanese high school students. *The Research Possibilities of Corpus-based Linguistics IV*. Research report for Corpus-based Linguistics and Language Education 8, pp. 369-376. Tokyo University of Foreign Studies.

Murakoshi, R. (2012b). Nihonjin Koukousei Eigo-gakushusha no Eisakubun ni Miru Bunpou-tokusei [Grammatical features in Japanese high school students' English compositions]. *ARCLE Review,* 6, pp. 90-99. Tokyo: Action Research Center for Language Education.

Murakoshi, R. (2013). Nihonjin Koukousei Eigo-gakushusha no Eisakubun ni Miru Bunpou-tokusei no Hattatsu [The development of grammatical features in English compositions of Japanese high school students]. *ARCLE Review,* 7, 24-33. Tokyo: Action Research Center for Language Education.

Murakoshi, R. (2015). *Nihonjin-koukousei eigo-gakushusha no eisakubun ni miru tougo-teki fukuzatsusa no hattatsu.* [The development of syntactic complexity in Japanese high school students' English compositions]. *ARCLE Review,* 9, 17-26. Tokyo: Action Research Center for Language Education.

Nassaji, H. and Fotos, S. (2011). *Teaching grammar in second language classroom*. New York: Routledge.

Negishi, M. (2012a). Two approaches for describing the CEFR levels of English: Core Inventory and English Profile Programme. In N. Tomimori (Ed.), *The General Study of Education Policies and Evaluation Systems of Foreign Languages in Higher Education in EU and Japan.* Project Report for Grant-in-Aid for Scientific Research (B): No. 21320101. An unpublished final report and collected papers. Tokyo University of Foreign Studies, pp.23-30.

Negishi, M. (2012b). *CEFR kijun-tokusei ni motozuku checklist houshiki ni yoru eisakubun no saiten kanousei* [Exploring the possibility of assessing English writing based on the checklist of CEFR criterial features]. *ARCLE Review,* 6, 80-89. Tokyo: Action Research Center for Language Education.

Negishi, M., Takada, T. and Tono, Y. (2012) A progress report on the development of the CEFR-J. *Studies in Language Testing* 36: 137-165. Cambridge: Cambridge University Press

Newton, J. and Kennedy, G. (1996). Effects of communication tasks on the grammatical relations marked by second language learners. *System*, 24 (3), 309-322.

North, B., Ortega, A., and Sheehan, S. (2010). A Core Inventory for General English. Retrieved from: https://englishagenda.britishcouncil.org/sites/default/files/attachments/pub-british-council-eaquals-core-inventoryv2.pdf.

British Council/EAQUALS (European Association for Quality Language Services) [2019.11.30]

Ohtomo, K. (1996). *Koumoku-outou Riron Nyumon －Gengo-tesuto Deta no Atarashii Bunseki-hou* [*The Introduction of Item Response Theory: A New Method of Analyzing Language Test Data*]. Tokyo: Taishukan Publishing.

Ortega, L. (2007). Meaningful L2 practice in foreign language classroom: A cognitive-interactionist SLA perspective. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology,* pp.180-207. Cambridge: Cambridge University Press.

Pallotti, G. (2007). An operational definition of the emergence criterion. *Applied Linguistics,* 28 (3), 361-382.

Pienemann, M. (1998). *Language processing and second language development: Processability Theory*. Amsterdam: John Benjamins.

Pienemann, M. (2003). Language processing capacity. In C. J. Doughty and M. H. Long, *The Handbook of Second Language Acquisition*, pp.674-714. Malden, MA: Blackwell.

Pienemann, M. (Ed.). (2005). *Cross-linguistic aspects of Processability Theory*. Amsterdam: John Benjamins.

Pienemann, M. (2007). Processability Theory. In B. VanPatten & J. Williams, *Theories in Second Language Acquisition: An Introduction*, pp. 137-154. Mahwah, NJ: Lawrence Erlbaum.

Pienemann, M. (2008). A brief introduction to Processability Theory. In J. -U. Keßler (Ed.), *Processability Approaches to Second Language Development and Second Language Learning,* pp. 9-29. Newcastle: Cambridge Scholars Publishing.

Pienemann, M., Johnston, M., and Brindley, G. (1988). Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition, 10* (2), 217-243.

Pienemann, M. and Keßler, J. -U. (2012). Processability Theory. In S. Gass & A. Mackey, *The Routledge Handbook of Second Language Acquisition,* pp. 228-246. New York: Routledge.

Polio, C. (2001). Research methodology in second language writing research: The case of text-based studies. In T. Silva & P. K. Matsuda (Eds.), *On second language writing* (pp. 91-116). Mahwah, NJ: Lawrence Erlbaum.

Rahkonen, M. and Håkansson, G. (2008). Production of written L2-Swedish—Processability or input frequencies? In J. –U. Keßler (Ed.), *Processability Approaches to Second Language Development and Second Language Learning*, pp. 9-29. Newcastle: Cambridge Scholars Publishing.

RASCAL (Version 3.50) [computer software]. St. Paul, MN: Assessment Systems Corporation.

Rovai, A. P., Jason, D. B., and Michael, K. P. (2013). *Social science research design and statistics: A practitioner's guide to research methods and IBM SPSS analysis (2nd ed.)*. Watertree Press. Chesapeake: VA.

Salamoura, A. and Saville, N. (2009). Criterial Features of English across the CEFR levels: evidence from the English Profile Programme. *Research Notes*, 37, 34-40. Cambridge ESOL. Retrieved from: http://www.cambridgeenglish.org/images/23156-research-notes-37.pdf [2019.11.30]

Salamoura, A. and Saville, N. (2010). Exemplifying the CEFR: criterial features of written learner English from the English Profile Programme. In I. Bartning, M. Martin & I. Vedder (Eds.), *Communicative Proficiency and Linguistic Development: Intersections between SLA and language testing research*, pp. 101-132. Eurosla.

Saville, N. (2010). The English Profile Programme－the first three years. *English Profile Journal,* 1(1), 1-14. DOI: 10.1017/S2041536210000061

Schumann, J. (1980). The acquisition of English relative clauses by second language learners. In R. C. Scarcella & S. D. Krashen (Eds.), *Research in second language acquisition*. Rowley, MA: Newbury House.

Shizuka, T. (2007). *Kiso-kara Fukaku Rikai-suru Rasch Modelling －Koumoku －outou Riron towa Nite Hinaru Sokutei no Paradigm* [*Understanding the Rasch Modeling from the Basics: the Paradigm of Measurement Similar to but Different from the Item Response Theory*]. Osaka: Kansai University Press.

Takahashi, M. (2002). *Koumoku-hannou Riron Nyumon －Atarashii Zettai-hyouka* [*The Introduction of Item Response Theory: A New Absolute Assessment Standard*]. Tokyo: Idea Press.

The Ministry of Education, Culture, Sports, Science and Technology (MEXT). (1989). The Couse of Study for Junior High School: Foreign Language (Japanese). Retrieved from: http://www.nier.go.jp/guideline/h01j/chap2-9.htm [2019.11.30]

The Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2008a). The Couse of Study for Junior High School (Japanese). Retrieved from: http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/chu/__icsFiles/afieldfile/2010/12/16/121504.pdf [2019.11.30]

The Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2008b). The Teaching Guide for the Couse of Study for Junior High School: Foreign Language (Japanese). Retrieved from: http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2011/01/05/1234912_010_1.pdf [2019.11.30]

The Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2009a). The Couse of Study for High School (Japanese). Retrieved from: http://www.mext.go.jp/component/a_menu/education/micro_detail/_icsFiles/afieldfile/2011/03/30/1304427_002.pdf [2019.11.30]

The Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2009b). The Teaching Guide for the Couse of Study for High School: Foreign Language (Japanese). Retrieved from: http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afieldfile/2010/01/29/1282000_9.pdf [2019.11.30]

The Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2015). *Heisei 26 nendo Eigo-kyouiku Kaizen no tameno Eigo-ryoku Chousa Jigyou-houkoku*. Retrieved from: http://www.mext.go.jp/a_menu/kokusai/gaikokugo/1358258.htm [2019.11.30]

The Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2016). *Heisei 27 nendo Eigo-kyouiku Kaizen no tameno Eigo-ryoku Chousa Jigyou-houkoku*. Retrieved from: http://www.mext.go.jp/a_menu/kokusai/gaikokugo/1377767.htm [2019.11.30]

The Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2018). *Heisei 29 nendo Eigo-kyouiku Kaizen no tameno Eigo-ryoku Chousa Jigyou-houkoku*. Retrieved from http://www.mext.go.jp/a_menu/kokusai/gaikokugo/1403470.htm [2019.11.30]

Tono, Y. (Ed.). (2012). *Research into development and validation of English language proficiency guidelines for Japanese learners of English at primary, secondary, and tertiary education.* Project report for Grant-in-Aid for Scientific Research (A): No. 20242011. An unpublished final report and collected papers. Tokyo University of Foreign Studies.

Tono, Y. (Ed.). (2013). *The CEFR-J Handbook: A Resource Book for Using CAN-DO Descriptors for English Language Teaching*. Tokyo: Taishukan Publishing.

Tono, Y. (Ed.). (2016). *The CEFR-J RLD project: Developing grammar, text and error profiles using textbooks & learner corpora*. Project report for Grant-in-Aid for Scientific Research (A): No. 20242017. An unpublished final report and collected papers. Tokyo University of Foreign Studies.

Tono, Y. (2017). The CEFR-J and its Impact on English Language Teaching in Japan. *JACET Selected Papers*, 4, 31-52.

Toyoda, H. (2002). *Koumoku-hannou Riron (Nyumon-hen)* [*The Item Response Theory: An Introduction*]. Tokyo: Asakura Publishing.

Wolfe-Quintero, K., Inagaki, S., and Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity* (Technical Report #17). Honolulu: University of Hawai'i Second Language Teaching and Curriculum Center.

Wong, J. (1991). Learnability of relative clauses: A Hong Kong case. *Working papers of the department of English 3* (1): 108-117. City Polytechnic of Hong Kong.

Yau, M. and Belanger, J. (1985). Syntactic development in the writing of ESL students. *English Quarterly, 18*, 107-118.

Yu, C. H. (2013). A simple guide to the Item Response Theory (IRT) and Rasch modeling. Retrieved from: http://www.creative-wisdom.com/computer/sas/IRT.pdf [2019.11.30]

Yule, G. (1998). *Explaining English grammar*. Oxford: Oxford University Press.

Zhang, Y. (2004). Processing constraints, categorical analysis, and the second language acquisition of the Chinese adjective suffix *–de* (ADJ). *Language Learning*, 54 (3), 437-468.

Zhang, Y. (2005). Processing and formal instruction in the L2 acquisition of five Chinese grammatical morphemes. In M. Pienemann (Ed.). *Cross-Linguistic Aspects of Processability Theory*. Amsterdam: John Benjamins.

Appendix A: Numbers of Emergence of the Grammatical Items
(including 16 eliminated list items*)

	S+V	S+V+O	S+V+O+O	S+V+that	direct WH−question	something +to−	NP+ *−ed*	NP+ *−ing*	double embedding:of −[of−]	may/can/ might (possibility)	must (obligation)	should (advice)
Composition A (1st year)	125	193	56	92	12	2	1	1	2	23	12	1
Composition B (2nd year)	166	206	29	105	18	0	1	0	0	62	21	11
Composition C (3rd year)	190	198	20	159	5	1	0	0	0	135	59	135

	S+V+O + to−	S+V+O + *−ing*	it ... that	S+V+ to− +that	whose (relative pronoun)	pseudocleft (WH−+S+V)	indirect WH−question	WH−+ to−	would rather/ had better	adverbial clause with −ing: (following main clause)	seem/ be supposed + to−	want, like, expect+O + to−
Composition A (1st year)	0	0	1	0	0	0	1	2	0	0	1	0
Composition B (2nd year)	0	0	0	0	0	4	4	3	0	4	0	1
Composition C (3rd year)	6	1	5	0	0	10	7	14	0	4	0	5

						*		*	*		*	*
	S is easy to−	of [−'s] −	may (permission)	must (necessity)	should (probability)	adverbial clause with −ing: (preceding main clause)	it ... to−	pseudocleft (WH−+V)	S+V+O+that	S+V+O+C (adj.)	V/Adj + to− (appear to, be sure to, etc.)	imagine/ prefer*O+ to−
Composition A (1st year)	0	0	0	0	0	0	7	1	2	26	0	0
Composition B (2nd year)	0	0	2	0	0	0	11	0	0	31	0	0
Composition C (3rd year)	0	0	6	0	0	0	52	0	3	6	1	0

	*	*	*	*	*	*	*	*	*	*	*
	be known/ obliged/ thought + to−	S is difficult/ good/hard+ to−	[...of−]'s	chance to−	believe/ find/ suppose/ take+O+ to−	be assumed/ discovered/ felt/ found/ proved to−	[...'s−]'s	might (permission)	declare/ presume/ remember+O + to−	be presumed to−	S is tough to−
Composition A (1st year)	0	0	0	0	0	0	0	0	0	0	0
Composition B (2nd year)	0	0	0	0	0	0	0	0	0	0	0
Composition C (3rd year)	0	0	0	0	0	0	0	0	0	0	0

	be−copula (present)	verb (base form)	*be*−copula (past)	verb (past)	S+*be*+C(adj.)	S+*look / feel / become* +C(adj.)	verb (3rd person singular)	personal pronoun (obj.)	present progressive	past progressive	S+V+O+C (noun)	tag question
Composition A (1st year)	206	201	70	102	185	21	54	179	36	5	3	0
Composition B (2nd year)	200	207	155	189	199	41	31	166	34	21	8	0
Composition C (3rd year)	189	209	38	76	161	18	24	112	29	4	0	1

	There is/are	can (ability)	can (request)	can (permission)	be going to−	will	have to−	could/would you −?	will you −?	shall I −?	infinitive (purpose)	infinitive (obj.)
Composition A (1st year)	8	44	0	1	2	47	10	0	0	0	3	60
Composition B (2nd year)	36	68	1	1	3	61	11	0	0	0	14	93
Composition C (3rd year)	28	87	1	6	5	74	28	1	0	0	18	77

	infinitive (comp.)	infinitive (adj.)	infinitive (reason of feeling)	It is ... for– to–	if (conj.: adverbial)	when (conj.)	because (conj.)	gerund (obj.)	gerund (subj.)	comparative	superlative	as – as
Composition A (1st year)	1	6	8	9	23	92	7	22	7	9	11	1
Composition B (2nd year)	3	9	26	8	17	113	51	23	8	32	20	3
Composition C (3rd year)	4	34	8	32	103	74	55	43	42	49	20	3

	passive	present perfect	relative pronoun (except whose)	prepositional phrase (adj.)
Composition A (1st year)	24	11	5	35
Composition B (2nd year)	13	27	16	97
Composition C (3rd year)	22	15	29	67

Appendix B: Output of the Rasch Model Analyses

1. The First-Year

```
                    RASCAL for Windows95 (tm) Version 3.50           Page  1
          Copyright (c) 1982 - 1996 by Assessment Systems Corporation

                     Rasch Model Item Calibration Program

 The number of items was  71

 The key was:
 11111111111111111111111111111111111111111111111111111111111111111111111

 The numbers of alternatives were:
 22222222222222222222222222222222222222222222222222222222222222222222222

 The inclusion specifications were:
 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY


 Items lost to editing: 18
 Total remaining items: 53

 Examinees lost to editing: 0
 Total remaining examinees: 209


 Scale Centered on:  Difficulty

 Model: Logistic [ D = 1.0 ]

 Correction for Bias in Final Estimates: YES

 Scale Adjustment Information:          Multiplicative Constant =     9.1000
                                              Additive Constant =   100.0000


 On loop  1 the average difficulty parameter change was     1.0042
 On loop  2 the average difficulty parameter change was     0.1516
 On loop  3 the average difficulty parameter change was     0.0288
 On loop  4 the average difficulty parameter change was     0.0061
 On loop  5 the average difficulty parameter change was     0.0013
 On loop  6 the average difficulty parameter change was     0.0003


 RASCAL converged after 6 Loops
```

Rasch Model Item Calibration Program

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
1	-3.158	0.156	10.250	15	71
2	-5.599	0.266	13.366	15	49
3	-1.497	0.169	22.563	15	86
4	-2.390	0.154	28.259	15	78
5	0.433	0.303	13.366	15	104
6	2.280	0.699	6.926	15	121
7	2.969	0.977	42.131	15	127
8	2.969	0.977	31.226	15	127
9	2.280	0.699	6.806	15	121
10	-0.305	0.230	16.930	15	97
11	0.433	0.303	11.854	15	104
12	2.969	0.977	18.728	15	127
13	--Deleted--				
14	--Deleted--				
15	2.969	0.977	7.807	15	127
16	--Deleted--				
17	--Deleted--				
18	--Deleted--				
19	2.969	0.977	9.030	15	127
20	2.280	0.699	12.826	15	121
21	--Deleted--				
22	--Deleted--				
23	2.969	0.977	8.027	15	127
24	--Deleted--				
25	--Deleted--				
26	--Deleted--				
27	--Deleted--				
28	--Deleted--				
29	--Deleted--				
30	1.006	0.386	6.476	15	109
31	-0.453	0.219	11.255	15	96
32	-7.440	0.557	3.260	15	32
33	-6.393	0.357	5.328	15	42
34	-1.865	0.160	14.570	15	83
35	-2.621	0.153	4.243	15	76
36	-5.099	0.226	17.093	15	54
37	-0.197	0.239	17.231	15	98
38	-1.441	0.171	20.890	15	87
39	-4.807	0.208	9.036	15	56
40	-0.866	0.194	27.866	15	92
41	1.354	0.452	18.600	15	112
42	1.873	0.575	6.490	15	117
43	--Deleted--				
44	0.866	0.363	15.056	15	108

Rasch Model Item Calibration Program

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
45	-1.140	0.182	12.928	15	90
46	--Deleted--				
47	2.969	0.977	17.307	15	127
48	2.280	0.699	40.611	15	121
49	-1.234	0.178	11.361	15	89
50	0.630	0.329	11.275	15	106
51	--Deleted--				
52	--Deleted--				
53	--Deleted--				
54	1.873	0.575	30.417	15	117
55	-1.606	0.166	17.073	15	85
56	2.969	0.977	7.807	15	127
57	1.166	0.415	14.065	15	111
58	0.866	0.363	16.343	15	108
59	0.742	0.345	6.201	15	107
60	-0.305	0.230	10.038	15	97
61	-2.390	0.154	21.191	15	78
62	1.006	0.386	8.244	15	109
63	-0.252	0.234	18.353	15	98
64	1.006	0.386	20.223	15	109
65	0.742	0.345	18.039	15	107
66	0.527	0.315	18.722	15	105
67	2.969	0.977	12.866	15	127
68	-0.356	0.226	10.573	15	97
69	0.527	0.315	20.208	15	105
70	1.354	0.452	16.373	15	112
71	-0.828	0.196	11.031	15	92

Rasch Model Item Calibration Program

Raw Score Conversion Table

Number Correct	(Theta) Ability	Std. Error	Freq-uency	Cum Freq	Percentile	Scaled Score
0	*****	*****	0	0	1	***
1	-7.71	1.223	0	0	1	30
2	-6.60	0.968	1	1	1	40
3	-5.78	0.863	0	1	1	47
4	-5.09	0.801	8	9	4	54
5	-4.48	0.747	10	19	9	59
6	-3.96	0.690	16	35	17	64
7	-3.52	0.636	24	59	28	68
8	-3.15	0.589	25	84	40	71
9	-2.83	0.550	19	103	49	74
10	-2.54	0.520	18	121	58	77
11	-2.28	0.495	22	143	68	79
12	-2.05	0.474	17	160	77	81
13	-1.83	0.457	14	174	83	83
14	-1.62	0.442	12	186	89	85
15	-1.43	0.430	10	196	94	87
16	-1.25	0.420	7	203	97	89
17	-1.08	0.411	3	206	99	90
18	-0.91	0.403	1	207	99	92
19	-0.75	0.396	1	208	99	93
20	-0.59	0.390	0	208	99	95
21	-0.44	0.385	1	209	99	96
22	-0.29	0.380	0	209	99	97
23	-0.15	0.376	0	209	99	99
24	-0.00	0.373	0	209	99	100
25	0.14	0.371	0	209	99	101
26	0.27	0.368	0	209	99	102
27	0.41	0.367	0	209	99	104
28	0.55	0.365	0	209	99	105
29	0.68	0.365	0	209	99	106
30	0.82	0.365	0	209	99	107
31	0.95	0.365	0	209	99	109
32	1.09	0.366	0	209	99	110
33	1.22	0.367	0	209	99	111
34	1.36	0.368	0	209	99	112
35	1.50	0.371	0	209	99	114
36	1.64	0.373	0	209	99	115
37	1.78	0.377	0	209	99	116
38	1.92	0.381	0	209	99	117
39	2.07	0.385	0	209	99	119
40	2.22	0.391	0	209	99	120
41	2.38	0.398	0	209	99	122
42	2.54	0.406	0	209	99	123
43	2.71	0.416	0	209	99	125

Rasch Model Item Calibration Program

Raw Score Conversion Table

Number Correct	(Theta) Ability	Std. Error	Freq-uency	Cum Freq	Percentile	Scaled Score
44	2.88	0.427	0	209	99	126
45	3.07	0.442	0	209	99	128
46	3.27	0.461	0	209	99	130
47	3.49	0.486	0	209	99	132
48	3.74	0.519	0	209	99	134
49	4.03	0.566	0	209	99	137
50	4.38	0.637	0	209	99	140
51	4.84	0.761	0	209	99	144
52	5.59	1.049	0	209	99	151
53	*****	*****	0	209	99	***

Rasch Model Item Calibration Program

Sorted in Item Difficulty Order

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
18	--Deleted--				
53	--Deleted--				
46	--Deleted--				
21	--Deleted--				
22	--Deleted--				
14	--Deleted--				
24	--Deleted--				
43	--Deleted--				
26	--Deleted--				
27	--Deleted--				
13	--Deleted--				
29	--Deleted--				
17	--Deleted--				
52	--Deleted--				
25	--Deleted--				
16	--Deleted--				
51	--Deleted--				
28	--Deleted--				
32	-7.440	0.557	3.260	15	32
33	-6.393	0.357	5.328	15	42
2	-5.599	0.266	13.366	15	49
36	-5.099	0.226	17.093	15	54
39	-4.807	0.208	9.036	15	56
1	-3.158	0.156	10.250	15	71
35	-2.621	0.153	4.243	15	76
4	-2.390	0.154	28.259	15	78
61	-2.390	0.154	21.191	15	78
34	-1.865	0.160	14.570	15	83
55	-1.606	0.166	17.073	15	85
3	-1.497	0.169	22.563	15	86
38	-1.441	0.171	20.890	15	87
49	-1.234	0.178	11.361	15	89
45	-1.140	0.182	12.928	15	90
40	-0.866	0.194	27.866	15	92
71	-0.828	0.196	11.031	15	92
31	-0.453	0.219	11.255	15	96
68	-0.356	0.226	10.573	15	97
60	-0.305	0.230	10.038	15	97
10	-0.305	0.230	16.930	15	97
63	-0.252	0.234	18.353	15	98
37	-0.197	0.239	17.231	15	98
5	0.433	0.303	13.366	15	104
11	0.433	0.303	11.854	15	104
66	0.527	0.315	18.722	15	105

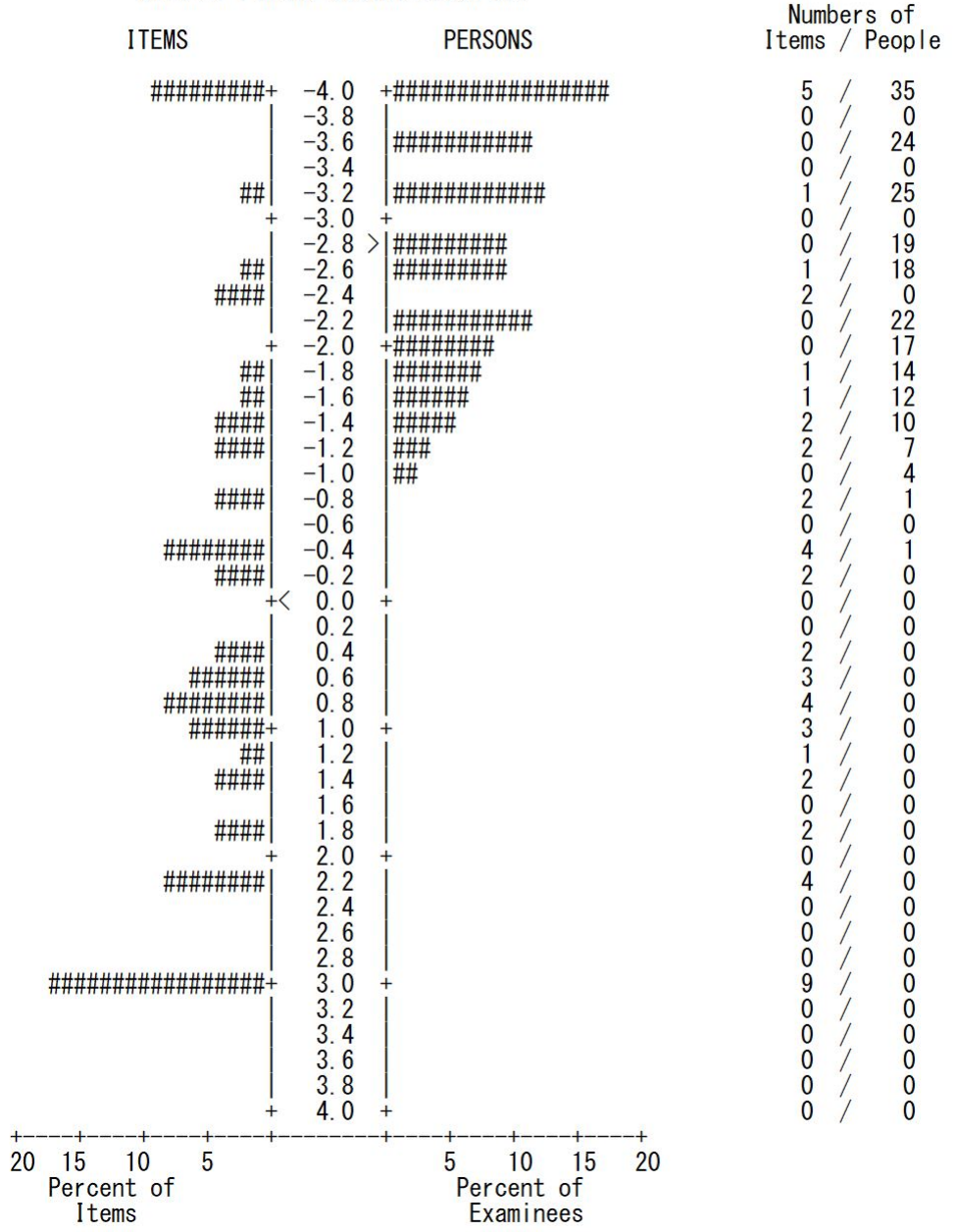
Rasch Model Item Calibration Program

Sorted in Item Difficulty Order

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
69	0.527	0.315	20.208	15	105
50	0.630	0.329	11.275	15	106
59	0.742	0.345	6.201	15	107
65	0.742	0.345	18.039	15	107
58	0.866	0.363	16.343	15	108
44	0.866	0.363	15.056	15	108
64	1.006	0.386	20.223	15	109
30	1.006	0.386	6.476	15	109
62	1.006	0.386	8.244	15	109
57	1.166	0.415	14.065	15	111
41	1.354	0.452	18.600	15	112
70	1.354	0.452	16.373	15	112
42	1.873	0.575	6.490	15	117
54	1.873	0.575	30.417	15	117
20	2.280	0.699	12.826	15	121
6	2.280	0.699	6.926	15	121
48	2.280	0.699	40.611	15	121
9	2.280	0.699	6.806	15	121
56	2.969	0.977	7.807	15	127
47	2.969	0.977	17.307	15	127
7	2.969	0.977	42.131	15	127
15	2.969	0.977	7.807	15	127
12	2.969	0.977	18.728	15	127
8	2.969	0.977	31.226	15	127
67	2.969	0.977	12.866	15	127
23	2.969	0.977	8.027	15	127
19	2.969	0.977	9.030	15	127

Rasch Model Item Calibration Program

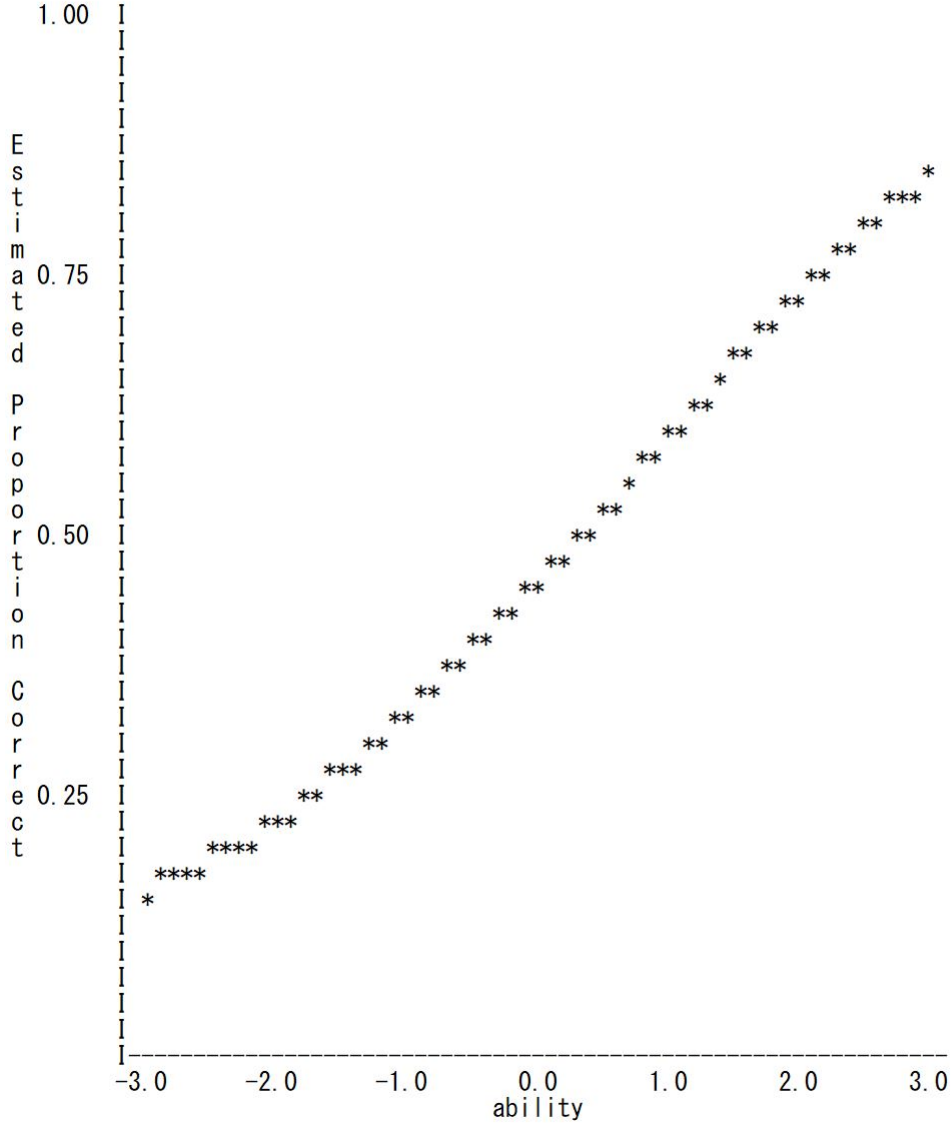
ITEM BY PERSON DISTRIBUTION MAP



Summary Information:	Average Difficulty	S.D. Difficulty	Average ability	S.D. ability
(Theta Metric)	-0.00	2.55	-2.76	1.04
(Scaled Score Metric)	100.0	23.2	74.9	9.5

Rasch Model Item Calibration Program

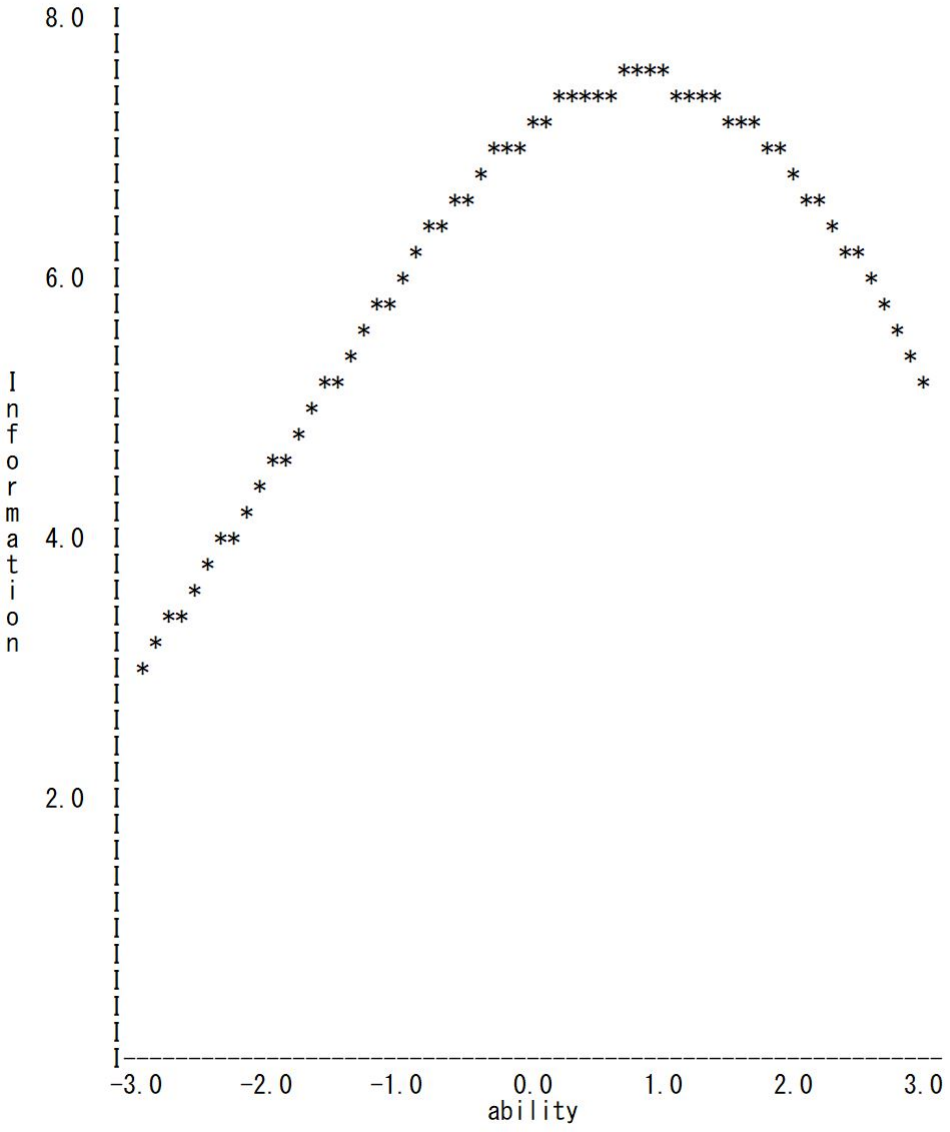
Test Characteristic Curve



Rasch Model Item Calibration Program

Test characteristics:	Estimated reliability	Expected Information	Average Information
	0.870	6.798	6.049

Test Information Curve



Elapsed Time: 0 second

2. The Second-Year

```
                    RASCAL for Windows95 (tm) Version 3.50           Page  1
          Copyright (c) 1982 - 1996 by Assessment Systems Corporation

                     Rasch Model Item Calibration Program


The number of items was  71

The key was:
11111111111111111111111111111111111111111111111111111111111111111111111

The numbers of alternatives were:
22222222222222222222222222222222222222222222222222222222222222222222222

The inclusion specifications were:
YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY


Items lost to editing: 18
Total remaining items: 53

Examinees lost to editing: 0
Total remaining examinees: 209


Scale Centered on:  Difficulty

Model: Logistic [ D = 1.0 ]

Correction for Bias in Final Estimates: YES

Scale Adjustment Information:          Multiplicative Constant =     9.1000
                                             Additive Constant =   100.0000


On loop  1 the average difficulty parameter change was     0.6678
On loop  2 the average difficulty parameter change was     0.1208
On loop  3 the average difficulty parameter change was     0.0183
On loop  4 the average difficulty parameter change was     0.0030
On loop  5 the average difficulty parameter change was     0.0005


RASCAL converged after 5 Loops
```

Rasch Model Item Calibration Program

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
1	-3.447	0.180	9.145	15	69
2	-6.485	0.555	5.033	15	41
3	0.088	0.207	17.106	15	101
4	-1.926	0.149	17.734	15	82
5	0.646	0.251	16.660	15	106
6	--Deleted--				
7	3.619	0.970	7.152	15	133
8	--Deleted--				
9	--Deleted--				
10	-0.949	0.161	13.401	15	91
11	0.470	0.236	24.037	15	104
12	1.186	0.312	12.134	15	111
13	--Deleted--				
14	--Deleted--				
15	--Deleted--				
16	--Deleted--				
17	--Deleted--				
18	2.235	0.498	22.672	15	120
19	2.235	0.498	13.191	15	120
20	2.525	0.571	35.157	15	123
21	--Deleted--				
22	2.235	0.498	7.358	15	120
23	--Deleted--				
24	3.619	0.970	14.261	15	133
25	--Deleted--				
26	--Deleted--				
27	2.931	0.694	20.168	15	127
28	--Deleted--				
29	--Deleted--				
30	1.186	0.312	12.975	15	111
31	0.006	0.202	13.879	15	100
32	-5.336	0.336	5.619	15	51
33	-6.895	0.672	4.427	15	37
34	-3.113	0.168	12.135	15	72
35	-4.432	0.239	11.059	15	60
36	-5.221	0.321	11.992	15	52
37	-0.354	0.183	17.426	15	97
38	0.006	0.202	10.978	15	100
39	-3.447	0.180	20.566	15	69
40	-0.110	0.195	19.832	15	99
41	0.470	0.236	22.154	15	104
42	1.522	0.360	11.591	15	114
43	--Deleted--				
44	-0.183	0.191	23.857	15	98

Rasch Model Item Calibration Program

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
45	-1.097	0.157	13.222	15	90
46	3.619	0.970	7.152	15	133
47	3.619	0.970	9.808	15	133
48	2.525	0.571	21.742	15	123
49	-0.923	0.162	14.186	15	92
50	1.186	0.312	13.272	15	111
51	--Deleted--				
52	--Deleted--				
53	--Deleted--				
54	0.925	0.280	7.501	15	108
55	-1.665	0.149	11.250	15	85
56	2.525	0.571	12.084	15	123
57	1.399	0.341	18.261	15	113
58	0.220	0.216	19.005	15	102
59	1.522	0.360	11.395	15	114
60	0.710	0.258	17.487	15	106
61	-2.100	0.149	16.446	15	81
62	-0.656	0.170	12.061	15	94
63	0.365	0.227	41.032	15	103
64	1.522	0.360	12.194	15	114
65	-0.034	0.200	7.825	15	100
66	0.527	0.241	27.589	15	105
67	2.525	0.571	11.614	15	123
68	1.006	0.289	13.913	15	109
69	0.175	0.213	9.711	15	102
70	0.778	0.264	15.359	15	107
71	-1.752	0.149	9.916	15	84

Rasch Model Item Calibration Program

Raw Score Conversion Table

Number Correct	(Theta) Ability	Std. Error	Freq- uency	Cum Freq	Percentile	Scaled Score
0	*****	*****	0	0	1	***
1	-7.52	1.188	0	0	1	32
2	-6.48	0.935	0	0	1	41
3	-5.72	0.832	0	0	1	48
4	-5.09	0.769	0	0	1	54
5	-4.53	0.720	4	4	2	59
6	-4.04	0.676	2	6	3	63
7	-3.61	0.636	6	12	6	67
8	-3.23	0.599	11	23	11	71
9	-2.88	0.567	15	38	18	74
10	-2.58	0.537	15	53	25	77
11	-2.30	0.512	27	80	38	79
12	-2.05	0.489	17	97	46	81
13	-1.82	0.469	23	120	57	83
14	-1.60	0.452	27	147	70	85
15	-1.40	0.437	19	166	79	87
16	-1.22	0.423	14	180	86	89
17	-1.04	0.412	10	190	91	91
18	-0.87	0.402	6	196	94	92
19	-0.71	0.393	4	200	96	94
20	-0.56	0.386	2	202	97	95
21	-0.41	0.379	0	202	97	96
22	-0.27	0.374	5	207	99	98
23	-0.13	0.369	2	209	99	99
24	0.01	0.365	0	209	99	100
25	0.14	0.362	0	209	99	101
26	0.27	0.360	0	209	99	102
27	0.40	0.358	0	209	99	104
28	0.53	0.357	0	209	99	105
29	0.66	0.357	0	209	99	106
30	0.79	0.357	0	209	99	107
31	0.92	0.358	0	209	99	108
32	1.05	0.359	0	209	99	110
33	1.18	0.361	0	209	99	111
34	1.31	0.364	0	209	99	112
35	1.45	0.367	0	209	99	113
36	1.58	0.371	0	209	99	114
37	1.72	0.375	0	209	99	116
38	1.87	0.380	0	209	99	117
39	2.01	0.387	0	209	99	118
40	2.17	0.394	0	209	99	120
41	2.33	0.402	0	209	99	121
42	2.49	0.412	0	209	99	123
43	2.67	0.423	0	209	99	124

Rasch Model Item Calibration Program

Date: 02 2 2020 Time: 20,36

Raw Score Conversion Table

Number Correct	(Theta) Ability	Std. Error	Freq-uency	Cum Freq	Percentile	Scaled Score
44	2.85	0.437	0	209	99	126
45	3.05	0.453	0	209	99	128
46	3.26	0.474	0	209	99	130
47	3.50	0.499	0	209	99	132
48	3.76	0.533	0	209	99	134
49	4.06	0.580	0	209	99	137
50	4.43	0.651	0	209	99	140
51	4.91	0.774	0	209	99	145
52	5.69	1.061	0	209	99	152
53	*****	*****	0	209	99	***

Rasch Model Item Calibration Program

Sorted in Item Difficulty Order

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
53	--Deleted--				
52	--Deleted--				
28	--Deleted--				
21	--Deleted--				
13	--Deleted--				
6	--Deleted--				
15	--Deleted--				
8	--Deleted--				
9	--Deleted--				
26	--Deleted--				
43	--Deleted--				
29	--Deleted--				
17	--Deleted--				
14	--Deleted--				
23	--Deleted--				
16	--Deleted--				
25	--Deleted--				
51	--Deleted--				
33	-6.895	0.672	4.427	15	37
2	-6.485	0.555	5.033	15	41
32	-5.336	0.336	5.619	15	51
36	-5.221	0.321	11.992	15	52
35	-4.432	0.239	11.059	15	60
1	-3.447	0.180	9.145	15	69
39	-3.447	0.180	20.566	15	69
34	-3.113	0.168	12.135	15	72
61	-2.100	0.149	16.446	15	81
4	-1.926	0.149	17.734	15	82
71	-1.752	0.149	9.916	15	84
55	-1.665	0.149	11.250	15	85
45	-1.097	0.157	13.222	15	90
10	-0.949	0.161	13.401	15	91
49	-0.923	0.162	14.186	15	92
62	-0.656	0.170	12.061	15	94
37	-0.354	0.183	17.426	15	97
44	-0.183	0.191	23.857	15	98
40	-0.110	0.195	19.832	15	99
65	-0.034	0.200	7.825	15	100
38	0.006	0.202	10.978	15	100
31	0.006	0.202	13.879	15	100
3	0.088	0.207	17.106	15	101
69	0.175	0.213	9.711	15	102
58	0.220	0.216	19.005	15	102
63	0.365	0.227	41.032	15	103

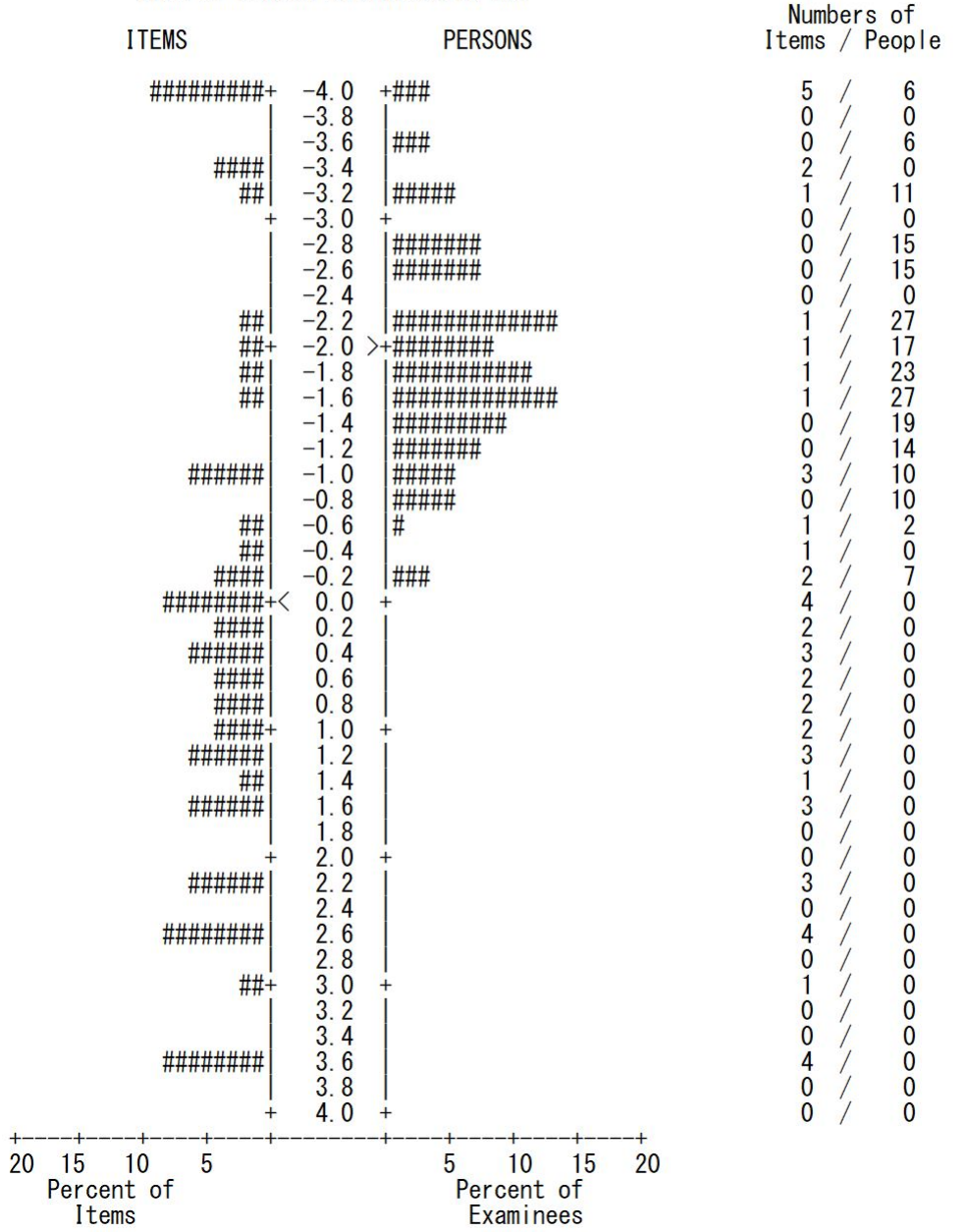
Rasch Model Item Calibration Program

Sorted in Item Difficulty Order

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
41	0.470	0.236	22.154	15	104
11	0.470	0.236	24.037	15	104
66	0.527	0.241	27.589	15	105
5	0.646	0.251	16.660	15	106
60	0.710	0.258	17.487	15	106
70	0.778	0.264	15.359	15	107
54	0.925	0.280	7.501	15	108
68	1.006	0.289	13.913	15	109
50	1.186	0.312	13.272	15	111
12	1.186	0.312	12.134	15	111
30	1.186	0.312	12.975	15	111
57	1.399	0.341	18.261	15	113
64	1.522	0.360	12.194	15	114
42	1.522	0.360	11.591	15	114
59	1.522	0.360	11.395	15	114
19	2.235	0.498	13.191	15	120
22	2.235	0.498	7.358	15	120
18	2.235	0.498	22.672	15	120
67	2.525	0.571	11.614	15	123
56	2.525	0.571	12.084	15	123
20	2.525	0.571	35.157	15	123
48	2.525	0.571	21.742	15	123
27	2.931	0.694	20.168	15	127
47	3.619	0.970	9.808	15	133
24	3.619	0.970	14.261	15	133
7	3.619	0.970	7.152	15	133
46	3.619	0.970	7.152	15	133

Rasch Model Item Calibration Program

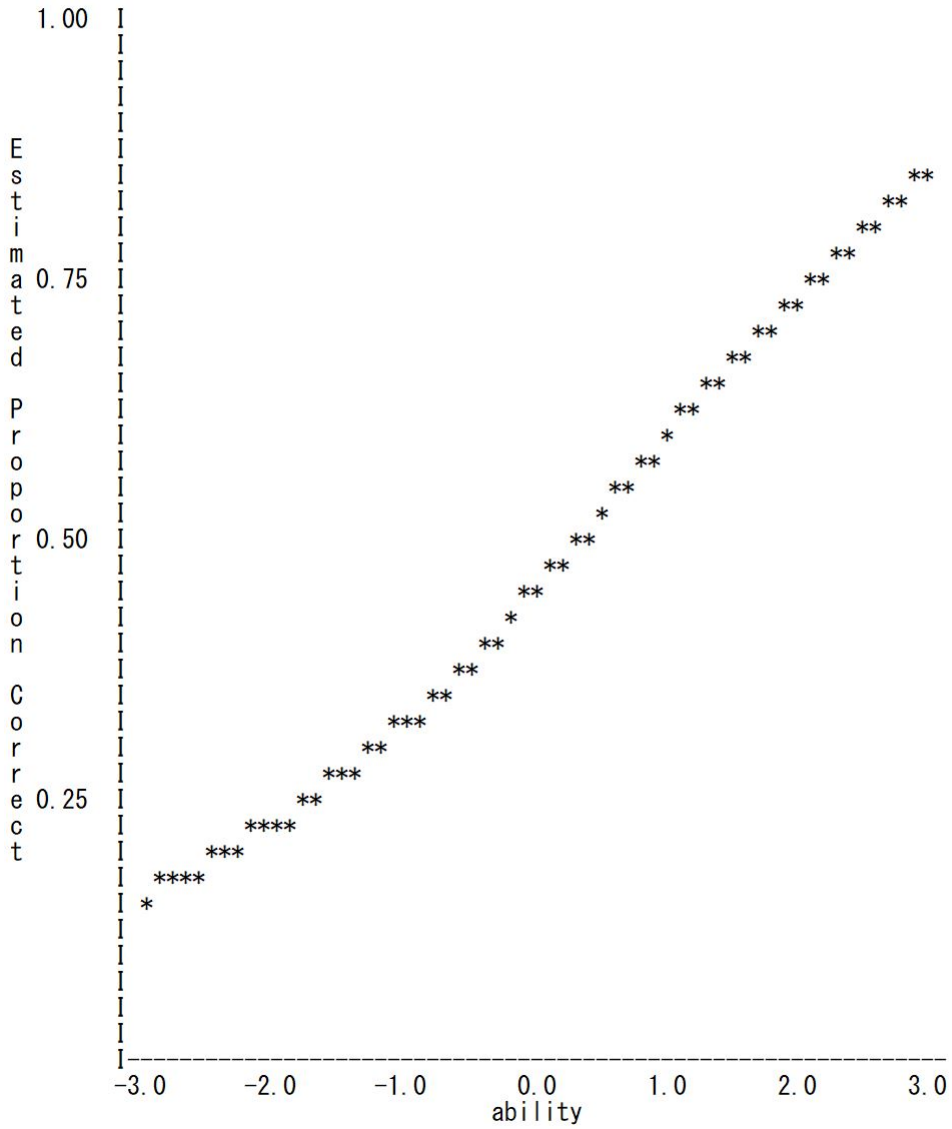
ITEM BY PERSON DISTRIBUTION MAP



Summary Information:	Average Difficulty	S.D. Difficulty	Average ability	S.D. ability
(Theta Metric)	-0.00	2.54	-1.97	0.86
(Scaled Score Metric)	100.0	23.1	82.1	7.8

Rasch Model Item Calibration Program

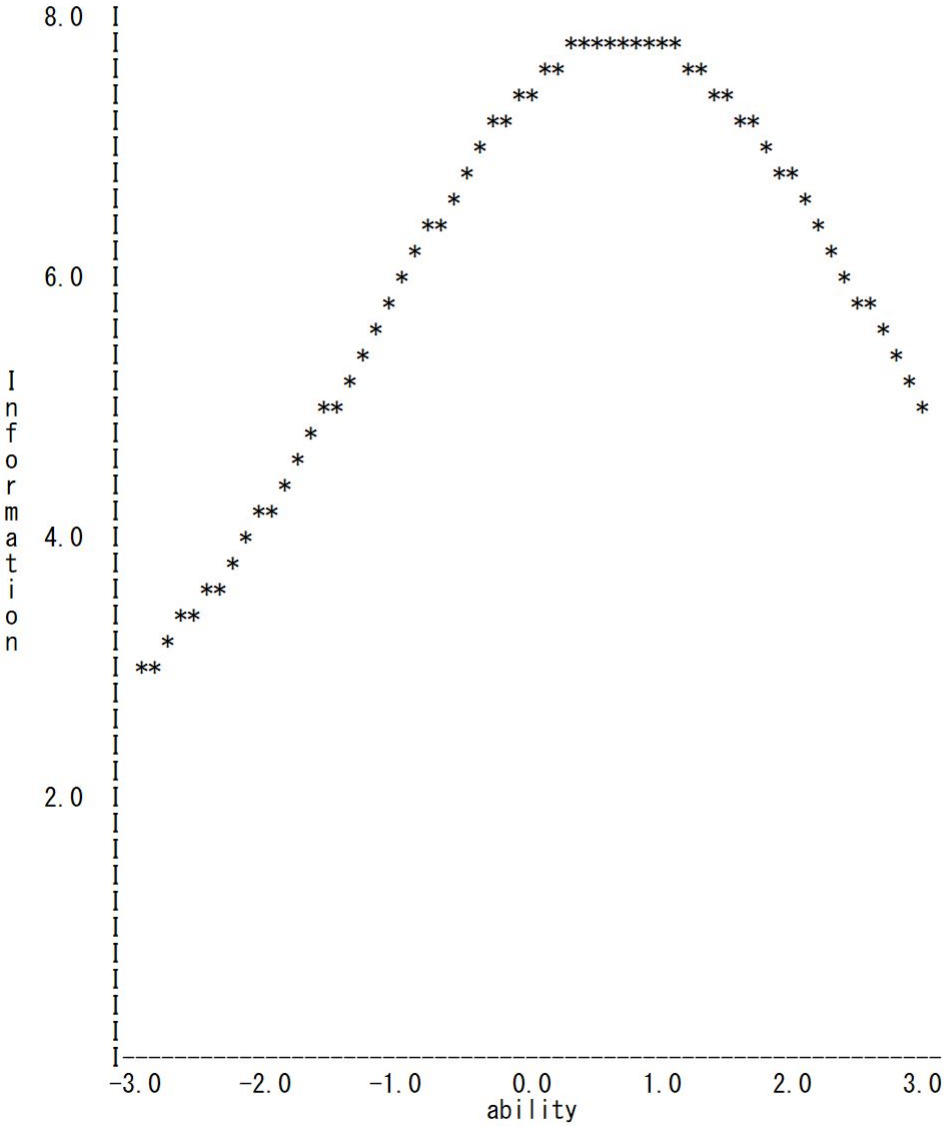
Test Characteristic Curve



Rasch Model Item Calibration Program

Test characteristics:	Estimated reliability	Expected Information	Average Information
	0.871	6.944	6.033

Test Information Curve



Elapsed Time: 0 second

3. The Third-Year

```
                    RASCAL for Windows95 (tm) Version 3.50           Page  1
          Copyright (c) 1982 - 1996 by Assessment Systems Corporation

                     Rasch Model Item Calibration Program

The number of items was  71

The key was:
11111111111111111111111111111111111111111111111111111111111111111111111

The numbers of alternatives were:
22222222222222222222222222222222222222222222222222222222222222222222222

The inclusion specifications were:
YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY


Items lost to editing: 15
Total remaining items: 56

Examinees lost to editing: 0
Total remaining examinees: 209


Scale Centered on:  Difficulty

Model: Logistic [ D = 1.0 ]

Correction for Bias in Final Estimates: YES

Scale Adjustment Information:          Multiplicative Constant =     9.1000
                                             Additive Constant =   100.0000


On loop  1 the average difficulty parameter change was     0.7287
On loop  2 the average difficulty parameter change was     0.1057
On loop  3 the average difficulty parameter change was     0.0150
On loop  4 the average difficulty parameter change was     0.0023
On loop  5 the average difficulty parameter change was     0.0003


RASCAL converged after 5 Loops
```

Rasch Model Item Calibration Program

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
1	-4.803	0.249	15.819	19	56
2	-5.454	0.315	6.754	19	50
3	0.350	0.242	18.760	19	103
4	-3.497	0.174	37.856	19	68
5	1.844	0.450	10.808	19	117
6	3.460	0.974	39.434	19	131
7	--Deleted--				
8	--Deleted--				
9	--Deleted--				
10	-2.848	0.157	30.237	19	74
11	-1.074	0.165	52.089	19	90
12	-2.848	0.157	15.267	19	74
13	1.656	0.413	24.196	19	115
14	3.460	0.974	19.277	19	131
15	1.844	0.450	11.762	19	117
16	--Deleted--				
17	--Deleted--				
18	1.119	0.327	22.363	19	110
19	1.495	0.385	24.566	19	114
20	0.753	0.282	18.699	19	107
21	--Deleted--				
22	2.072	0.500	17.274	19	119
23	--Deleted--				
24	1.844	0.450	26.694	19	117
25	--Deleted--				
26	--Deleted--				
27	1.656	0.413	46.674	19	115
28	--Deleted--				
29	--Deleted--				
30	-0.882	0.171	16.044	19	92
31	1.656	0.413	23.442	19	115
32	-4.740	0.244	13.928	19	57
33	--Deleted--				
34	-0.443	0.189	22.551	19	96
35	-1.498	0.156	17.222	19	86
36	-3.558	0.176	13.875	19	68
37	0.471	0.253	15.318	19	104
38	0.135	0.225	16.280	19	101
39	-2.313	0.151	12.057	19	79
40	-0.095	0.209	32.979	19	99
41	2.072	0.500	23.010	19	119
42	--Deleted--				
43	3.460	0.974	11.108	19	131
44	-0.052	0.212	26.473	19	100

Rasch Model Item Calibration Program

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
45	-1.753	0.152	20.197	19	84
46	3.460	0.974	41.897	19	131
47	1.656	0.413	9.332	19	115
48	1.844	0.450	14.230	19	117
49	-1.451	0.156	20.130	19	87
50	-0.052	0.212	13.946	19	100
51	3.460	0.974	25.754	19	131
52	--Deleted--				
53	--Deleted--				
54	0.471	0.253	24.692	19	104
55	-1.522	0.155	19.228	19	86
56	2.072	0.500	6.588	19	119
57	-0.297	0.197	21.220	19	97
58	1.356	0.362	11.375	19	112
59	-0.219	0.201	20.327	19	98
60	-2.112	0.151	27.633	19	81
61	-1.451	0.156	29.092	19	87
62	-0.966	0.168	24.307	19	91
63	-0.610	0.181	13.203	19	94
64	-0.578	0.183	13.895	19	95
65	-0.795	0.174	19.041	19	93
66	0.350	0.242	43.225	19	103
67	2.363	0.573	20.833	19	122
68	0.239	0.233	10.100	19	102
69	0.677	0.274	26.783	19	106
70	-0.095	0.209	30.405	19	99
71	-1.280	0.160	13.717	19	88

Rasch Model Item Calibration Program

Raw Score Conversion Table

Number Correct	(Theta) Ability	Std. Error	Freq-uency	Cum Freq	Percentile	Scaled Score
0	*****	*****	0	0	1	***
1	-6.34	1.127	1	1	1	42
2	-5.43	0.854	1	2	1	51
3	-4.82	0.736	2	4	2	56
4	-4.34	0.663	3	7	3	61
5	-3.94	0.610	5	12	6	64
6	-3.59	0.569	7	19	9	67
7	-3.29	0.535	16	35	17	70
8	-3.02	0.507	10	45	22	73
9	-2.77	0.484	18	63	30	75
10	-2.55	0.464	14	77	37	77
11	-2.34	0.446	15	92	44	79
12	-2.15	0.432	15	107	51	80
13	-1.96	0.419	19	126	60	82
14	-1.79	0.408	20	146	70	84
15	-1.63	0.398	11	157	75	85
16	-1.47	0.390	16	173	83	87
17	-1.32	0.383	6	179	86	88
18	-1.18	0.377	7	186	89	89
19	-1.04	0.371	4	190	91	91
20	-0.90	0.367	3	193	92	92
21	-0.76	0.363	10	203	97	93
22	-0.63	0.359	2	205	98	94
23	-0.50	0.356	1	206	99	95
24	-0.38	0.354	0	206	99	97
25	-0.25	0.352	1	207	99	98
26	-0.13	0.351	0	207	99	99
27	-0.00	0.350	2	209	99	100
28	0.12	0.349	0	209	99	101
29	0.24	0.349	0	209	99	102
30	0.37	0.349	0	209	99	103
31	0.49	0.349	0	209	99	104
32	0.61	0.350	0	209	99	106
33	0.74	0.351	0	209	99	107
34	0.86	0.353	0	209	99	108
35	0.99	0.355	0	209	99	109
36	1.12	0.357	0	209	99	110
37	1.25	0.360	0	209	99	111
38	1.38	0.363	0	209	99	113
39	1.51	0.367	0	209	99	114
40	1.65	0.371	0	209	99	115
41	1.79	0.376	0	209	99	116
42	1.94	0.383	0	209	99	118
43	2.09	0.390	0	209	99	119

Rasch Model Item Calibration Program

Raw Score Conversion Table

Number Correct	(Theta) Ability	Std. Error	Freq-uency	Cum Freq	Percentile	Scaled Score
44	2.24	0.398	0	209	99	120
45	2.40	0.408	0	209	99	122
46	2.57	0.420	0	209	99	123
47	2.76	0.434	0	209	99	125
48	2.95	0.451	0	209	99	127
49	3.16	0.471	0	209	99	129
50	3.39	0.498	0	209	99	131
51	3.66	0.532	0	209	99	133
52	3.96	0.579	0	209	99	136
53	4.33	0.650	0	209	99	139
54	4.81	0.773	0	209	99	144
55	5.58	1.059	0	209	99	151
56	*****	*****	0	209	99	***

Rasch Model Item Calibration Program

Sorted in Item Difficulty Order

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
53	--Deleted--				
52	--Deleted--				
28	--Deleted--				
21	--Deleted--				
9	--Deleted--				
23	--Deleted--				
7	--Deleted--				
8	--Deleted--				
17	--Deleted--				
26	--Deleted--				
25	--Deleted--				
29	--Deleted--				
33	--Deleted--				
42	--Deleted--				
16	--Deleted--				
2	-5.454	0.315	6.754	19	50
1	-4.803	0.249	15.819	19	56
32	-4.740	0.244	13.928	19	57
36	-3.558	0.176	13.875	19	68
4	-3.497	0.174	37.856	19	68
10	-2.848	0.157	30.237	19	74
12	-2.848	0.157	15.267	19	74
39	-2.313	0.151	12.057	19	79
60	-2.112	0.151	27.633	19	81
45	-1.753	0.152	20.197	19	84
55	-1.522	0.155	19.228	19	86
35	-1.498	0.156	17.222	19	86
49	-1.451	0.156	20.130	19	87
61	-1.451	0.156	29.092	19	87
71	-1.280	0.160	13.717	19	88
11	-1.074	0.165	52.089	19	90
62	-0.966	0.168	24.307	19	91
30	-0.882	0.171	16.044	19	92
65	-0.795	0.174	19.041	19	93
63	-0.610	0.181	13.203	19	94
64	-0.578	0.183	13.895	19	95
34	-0.443	0.189	22.551	19	96
57	-0.297	0.197	21.220	19	97
59	-0.219	0.201	20.327	19	98
40	-0.095	0.209	32.979	19	99
70	-0.095	0.209	30.405	19	99
44	-0.052	0.212	26.473	19	100
50	-0.052	0.212	13.946	19	100
38	0.135	0.225	16.280	19	101

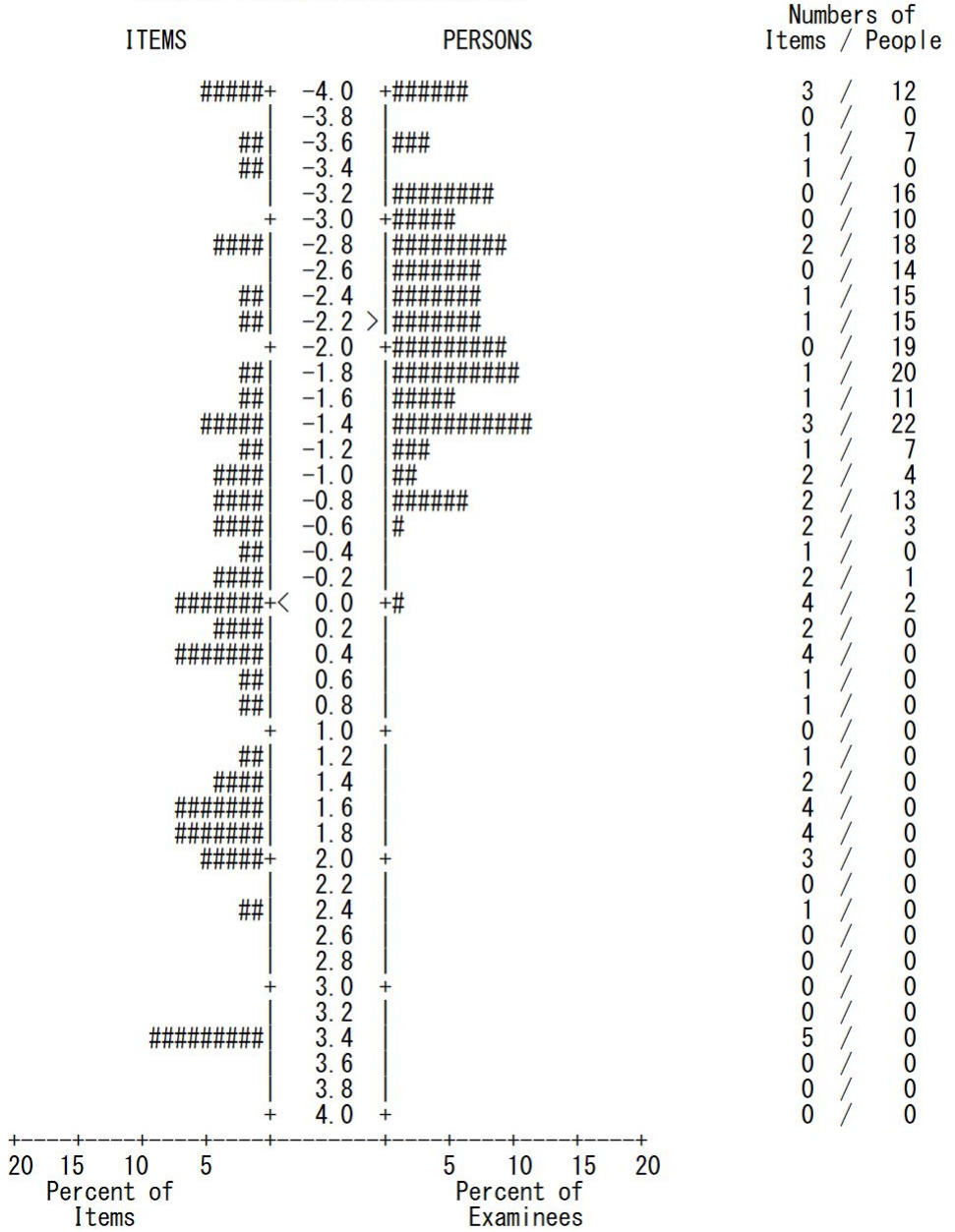
Rasch Model Item Calibration Program

Sorted in Item Difficulty Order

Item	Difficulty	Std. Error	Chi Sq.	df	Scaled Diff
68	0.239	0.233	10.100	19	102
3	0.350	0.242	18.760	19	103
66	0.350	0.242	43.225	19	103
54	0.471	0.253	24.692	19	104
37	0.471	0.253	15.318	19	104
69	0.677	0.274	26.783	19	106
20	0.753	0.282	18.699	19	107
18	1.119	0.327	22.363	19	110
58	1.356	0.362	11.375	19	112
19	1.495	0.385	24.566	19	114
47	1.656	0.413	9.332	19	115
27	1.656	0.413	46.674	19	115
13	1.656	0.413	24.196	19	115
31	1.656	0.413	23.442	19	115
15	1.844	0.450	11.762	19	117
24	1.844	0.450	26.694	19	117
5	1.844	0.450	10.808	19	117
48	1.844	0.450	14.230	19	117
41	2.072	0.500	23.010	19	119
56	2.072	0.500	6.588	19	119
22	2.072	0.500	17.274	19	119
67	2.363	0.573	20.833	19	122
43	3.460	0.974	11.108	19	131
6	3.460	0.974	39.434	19	131
51	3.460	0.974	25.754	19	131
14	3.460	0.974	19.277	19	131
46	3.460	0.974	41.897	19	131

Rasch Model Item Calibration Program

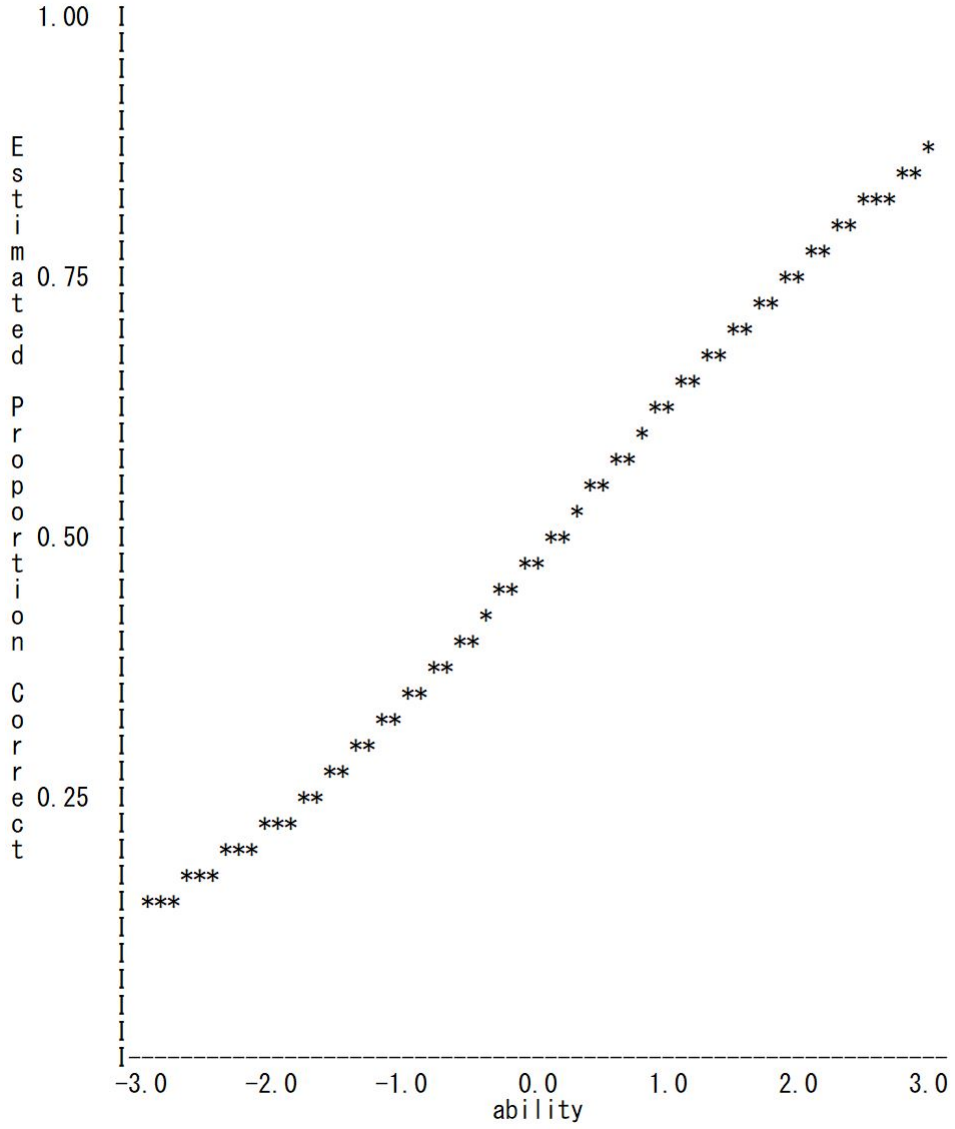
ITEM BY PERSON DISTRIBUTION MAP



Summary Information:	Average Difficulty	S.D. Difficulty	Average ability	S.D. ability
(Theta Metric)	-0.00	2.15	-2.21	0.97
(Scaled Score Metric)	100.0	19.6	79.9	8.9

Rasch Model Item Calibration Program

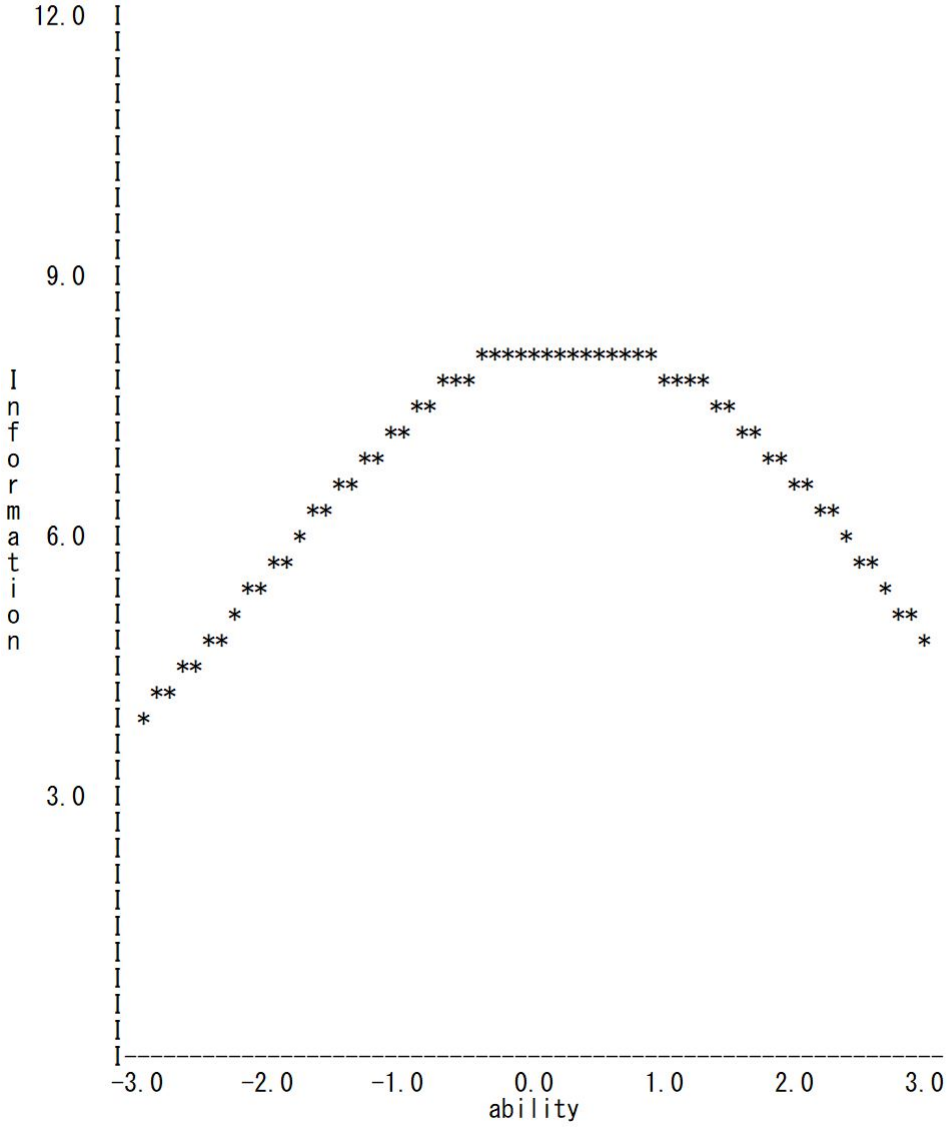
Test Characteristic Curve



Rasch Model Item Calibration Program

Test characteristics:	Estimated reliability	Expected Information	Average Information
	0.884	7.671	6.704

Test Information Curve



Elapsed Time: 0 second

Appendix C: Word Counts, MLUs and GTEC Writing Scores

ID	Word Count (1st year)	MLU (1st year)	GTEC/W (1st year)	Word Count (2nd year)	MLU (2nd year)	GTEC/W (2nd year)	Word Count (3rd year)	MLU (3rd year)	GTEC/W (3rd year)
21001	162	7.4	98	115	8.8	107	136	9.1	106
21002	43	6.1	77	99	5.8	100	52	8.7	114
21005	120	6.3	93	299	8.5	122	266	9.2	133
21006	75	6.8	72	99	6.6	100	74	8.2	97
21007	54	7.7	93	182	7.3	102	59	7.4	106
21008	79	8.8	88	147	9.2	106	129	7.6	112
21009	124	6.9	98	172	7.2	103	171	8.1	116
21010	159	6.1	98	122	5.8	110	160	8.0	126
21011	45	5.0	77	78	5.2	98	63	7.9	90
21012	118	5.9	95	160	5.7	109	50	7.1	109
21013	208	8.7	90	178	8.1	111	129	9.9	111
21014	70	5.4	90	129	6.8	102	126	6.6	112
21015	95	5.9	90	157	6.8	107	122	7.6	108
21016	96	6.9	88	150	5.6	114	119	8.5	120
21017	133	6.7	93	154	6.4	122	193	8.8	129
21018	111	5.8	83	170	9.4	106	124	8.9	108
21019	138	6.9	83	154	8.6	96	92	6.6	105
21020	117	6.2	98	119	7.0	107	170	8.9	130
21021	61	7.6	88	174	6.2	106	101	9.2	117
21022	65	9.3	88	145	8.5	101	66	7.3	95
21023	20	10.0	64	71	7.1	89	34	11.3	106
21024	25	6.3	69	54	5.4	79	59	6.6	89
21025	132	7.3	114	283	7.9	126	193	10.7	131
21026	41	5.9	83	191	11.2	93	67	8.4	107
21027	137	7.2	93	132	10.2	107	254	13.4	131
21029	22	5.5	72	95	13.6	95	132	13.2	109
21030	44	6.3	72	152	5.8	105	157	7.5	115
21031	77	7.0	88	112	5.9	96	131	6.9	94
21032	110	7.9	80	173	9.1	96	114	14.3	112
21033	257	6.6	82	213	8.5	111	186	13.3	116
21034	112	5.6	82	138	7.3	101	110	8.5	114
21035	70	5.8	82	175	8.0	92	159	8.0	110
21036	109	7.3	98	122	8.7	122	114	11.4	121
21037	100	5.9	88	182	6.7	106	112	8.6	110
21038	100	5.3	90	124	7.3	100	125	6.9	96
21039	90	6.9	77	238	9.2	110	159	9.4	122
21040	45	7.5	83	107	8.9	106	73	7.3	98
21041	72	6.5	77	131	6.9	95	117	9.0	111
21042	88	5.9	77	105	7.5	112	95	7.3	107
21043	115	8.2	93	206	8.6	122	183	15.3	124
21044	94	5.9	88	156	6.0	102	60	7.5	110
21045	124	7.8	88	151	7.9	101	129	9.9	108
21046	102	6.0	98	146	6.1	107	53	5.9	109
21047	189	6.3	98	145	8.1	125	146	8.6	124
21048	98	5.8	104	113	7.5	118	83	7.5	123
21049	39	7.8	104	163	6.8	113	126	9.7	122
21050	55	7.9	77	117	9.8	90	53	8.8	87
21051	122	5.5	88	90	5.6	91	48	5.3	106
21053	114	8.8	104	183	7.0	108	149	11.5	113
21054	63	6.3	93	138	8.6	107	112	10.2	113
21055	110	5.8	112	115	6.4	104	97	9.7	105
21056	92	6.6	93	133	7.8	107	115	10.5	125
21057	67	6.1	85	83	6.9	102	99	9.0	102

ID	Word Count (1st year)	MLU (1st year)	GTEC/W (1st year)	Word Count (2nd year)	MLU (2nd year)	GTEC/W (2nd year)	Word Count (3rd year)	MLU (3rd year)	GTEC/W (3rd year)
21059	65	6.5	72	180	5.3	114	33	4.1	115
21060	71	5.9	80	189	8.2	106	66	9.4	114
21061	104	6.9	93	95	6.3	110	56	8.0	109
21062	122	6.8	88	176	8.4	102	122	12.2	111
21063	164	6.6	88	156	10.4	106	130	8.7	116
21064	115	6.1	104	211	5.1	123	56	6.2	122
21066	127	4.9	101	210	6.8	105	76	5.1	115
21067	180	5.8	120	262	6.4	130	119	7.0	130
21068	151	5.8	88	123	6.8	99	63	6.3	100
21069	57	4.8	93	166	7.2	107	117	9.0	115
21070	90	5.6	88	149	6.0	112	47	6.7	122
21072	37	7.4	77	91	8.3	105	62	10.3	105
21073	77	6.4	98	176	6.5	107	102	6.8	119
21074	45	5.6	88	93	5.8	111	86	9.6	122
21075	142	7.9	99	191	8.0	107	91	13.0	118
21076	179	6.6	72	81	11.6	105	105	8.8	108
21077	73	7.3	88	119	7.0	102	106	8.2	108
21078	128	7.1	77	122	8.1	101	111	8.5	112
21079	75	9.4	66	118	6.9	110	55	9.2	122
21080	88	8.0	82	144	7.2	106	135	9.6	111
21081	30	7.5	80	96	7.4	88	47	6.7	106
21082	110	6.1	98	201	7.2	117	74	9.3	119
21083	118	6.9	93	216	8.3	107	80	8.0	115
21085	38	9.5	75	66	6.6	84	54	9.0	107
21086	101	9.2	98	29	7.3	92	67	8.4	104
21087	72	7.2	72	113	7.5	98	137	9.8	110
21088	49	7.0	88	84	7.6	91	60	7.5	107
21089	125	8.3	88	243	10.6	106	169	7.7	121
21090	78	7.8	66	73	5.6	84	23	5.8	94
21091	73	6.6	93	145	6.0	91	50	6.3	80
21092	76	9.5	88	107	6.3	97	85	9.4	106
21093	73	7.3	77	121	7.1	101	60	7.5	107
21094	52	5.8	75	54	6.8	100	72	8.0	94
21095	71	7.1	93	151	6.6	109	151	7.9	116
21097	72	5.1	88	139	6.3	106	81	7.4	111
21098	100	6.3	93	109	7.8	97	95	7.9	111
21099	110	7.3	82	133	8.9	106	141	8.3	121
21100	144	8.0	120	114	7.6	112	112	8.0	122
21101	69	6.3	88	118	7.4	106	149	7.8	108
21102	103	5.7	77	160	7.0	105	113	7.5	108
21103	73	6.1	96	157	6.8	100	125	6.6	115
21104	66	6.0	101	105	6.6	102	66	6.6	105
21105	67	8.4	98	166	6.6	124	51	8.5	107
21106	100	7.7	117	213	10.1	94	95	7.9	122
21107	128	6.4	98	160	7.6	79	122	7.6	102
21109	115	7.7	86	206	6.4	102	76	8.4	105
21110	37	7.4	72	125	7.8	105	113	8.7	117
21111	62	7.8	88	132	7.8	99	158	9.9	114
21112	92	6.6	115	170	6.3	109	165	10.3	121
21114	86	6.1	91	127	6.7	81	76	6.9	100
21116	185	6.9	93	168	9.3	110	107	13.4	121
21117	34	4.9	88	77	7.0	96	144	9.6	97
21118	56	6.2	85	125	8.9	97	183	12.2	107

ID	Word Count (1st year)	MLU (1st year)	GTEC/W (1st year)	Word Count (2nd year)	MLU (2nd year)	GTEC/W (2nd year)	Word Count (3rd year)	MLU (3rd year)	GTEC/W (3rd year)
21119	115	5.2	104	157	10.5	113	145	14.5	133
21120	125	7.8	101	138	8.6	108	174	10.9	116
21121	169	7.7	75	156	7.4	96	56	7.0	104
21122	27	5.4	80	65	5.0	106	79	6.6	108
21123	181	6.7	83	119	6.6	121	109	8.4	117
21124	72	7.2	88	135	7.9	101	31	7.8	32
21125	87	7.3	93	93	7.8	99	24	6.0	90
21126	31	6.2	75	127	10.6	90	92	7.7	103
21127	97	8.1	101	136	6.2	115	178	8.5	112
21128	64	7.1	85	95	9.5	91	83	8.3	107
21129	29	7.3	72	175	7.0	114	78	7.1	121
21132	81	5.8	90	187	7.5	107	178	9.4	106
21133	82	7.5	82	141	7.8	106	187	8.5	112
21134	56	7.0	82	130	8.7	85	78	8.7	102
21136	102	7.8	82	122	8.1	80	44	11.0	77
21137	242	6.2	117	233	7.1	127	102	10.2	124
21138	56	6.2	90	104	6.5	96	118	7.9	110
21140	147	6.4	101	187	6.7	110	193	7.1	120
21141	107	6.3	88	127	8.5	106	115	9.6	107
21142	84	7.6	88	81	7.4	91	95	9.5	102
21143	51	6.4	91	170	8.9	107	207	8.3	121
21144	99	7.6	80	107	7.6	106	185	8.0	112
21145	43	7.2	72	100	7.7	100	101	7.8	106
21146	142	7.1	85	216	9.0	109	126	10.5	112
21147	111	6.9	77	262	9.7	113	149	9.9	114
21148	215	5.0	109	224	9.7	116	283	17.7	122
21149	71	6.5	77	118	6.9	105	152	7.6	116
21150	111	8.5	88	133	8.3	86	70	11.7	110
21151	92	7.7	88	154	8.1	112	100	9.1	116
21152	69	4.9	72	89	5.9	95	90	5.6	105
21153	59	6.6	85	151	5.6	101	113	7.1	115
21154	152	7.6	88	139	9.9	101	90	12.9	108
21157	139	5.8	77	109	6.4	68	56	4.7	96
21159	125	8.3	88	97	8.8	101	155	7.8	106
21160	80	8.9	88	97	5.7	106	102	11.3	113
21161	54	6.0	67	88	5.5	87	99	7.6	110
21163	45	7.5	83	168	9.3	96	81	5.8	110
21164	101	7.2	83	128	8.5	106	55	11.0	95
21165	124	8.3	90	136	7.2	107	20	5.0	98
21166	56	6.2	13	144	6.9	86	67	5.2	92
21167	124	9.5	85	216	12.7	98	185	13.2	123
21168	56	8.0	88	59	7.4	91	70	7.0	93
21169	61	7.6	80	128	10.7	98	26	6.5	113
21170	113	7.5	88	202	10.1	106	121	8.6	118
21172	36	5.1	109	204	12.8	121	151	8.4	121
21173	75	6.8	89	180	8.2	100	135	7.5	90
21174	102	7.8	93	232	8.3	107	187	7.8	121
21175	195	8.1	101	179	7.2	108	136	10.5	114
21176	61	7.6	88	94	7.8	97	137	9.1	91
21177	75	9.4	72	112	7.0	88	138	17.3	105
21178	92	6.1	83	75	7.5	98	56	7.0	93
21179	80	7.3	82	101	7.2	104	107	7.6	86
21180	58	6.4	99	159	9.9	103	179	9.4	114

ID	Word Count (1st year)	MLU (1st year)	GTEC/W (1st year)	Word Count (2nd year)	MLU (2nd year)	GTEC/W (2nd year)	Word Count (3rd year)	MLU (3rd year)	GTEC/W (3rd year)
21181	39	6.5	82	124	8.3	95	105	6.2	105
21182	58	6.4	77	113	7.1	95	140	8.2	102
21184	196	6.8	96	148	7.0	125	184	10.8	120
21185	119	9.9	93	172	7.5	122	39	9.8	81
21187	172	7.5	90	270	10.8	116	177	9.3	131
21188	137	6.2	91	112	8.6	107	125	13.9	108
21190	164	10.3	98	187	8.9	112	195	11.5	125
21191	73	7.3	88	149	9.3	96	109	8.4	110
21192	82	6.8	88	136	7.6	101	130	8.1	106
21193	47	5.9	88	123	6.5	99	141	10.1	112
21194	173	8.2	82	161	7.0	106	125	8.9	111
21195	95	7.9	80	205	8.9	118	250	10.4	126
21196	158	6.9	93	132	8.8	102	54	7.7	101
21197	36	5.1	82	105	5.3	101	81	6.8	99
21198	68	5.7	83	108	5.4	102	76	6.3	101
21199	68	7.6	98	102	6.0	102	143	8.9	114
21200	37	4.6	77	138	7.7	101	70	8.8	115
21201	128	6.4	104	192	9.1	123	173	10.8	122
21202	101	6.7	90	162	8.1	99	124	9.5	112
21203	96	5.3	88	120	7.5	106	144	6.5	121
21204	130	5.7	96	175	7.6	114	105	10.5	114
21205	54	9.0	88	200	9.1	116	161	10.7	122
21206	93	6.6	83	140	6.7	93	85	7.7	102
21208	102	8.5	77	58	6.4	68	168	10.5	110
21209	97	8.1	80	103	9.4	106	133	11.1	116
21210	46	5.8	77	74	9.3	90	74	9.3	103
21211	155	6.2	82	155	6.7	109	127	7.5	104
21212	62	5.6	77	114	8.1	50	41	6.8	71
21213	127	11.5	75	166	7.2	105	58	9.7	110
21215	52	5.8	98	189	9.0	117	170	10.6	112
21217	104	7.4	88	125	8.9	105	59	6.6	108
21218	27	6.8	72	117	9.8	75	129	12.9	99
21219	32	5.3	88	39	4.9	100	29	5.8	100
21220	249	5.9	96	213	6.5	118	128	6.1	125
21221	64	10.7	88	157	7.9	97	45	6.4	107
21222	55	7.9	98	158	9.9	112	86	8.6	109
21223	94	8.5	88	155	9.1	96	174	12.4	116
21224	114	6.7	88	192	7.4	106	157	8.7	118
21225	71	5.9	88	131	6.0	102	195	12.2	106
21226	152	6.3	93	304	9.2	138	295	10.5	136
21227	40	6.7	72	123	7.2	105	120	8.0	95
21228	128	5.6	93	121	6.7	100	126	6.6	112
21230	50	12.5	90	118	6.2	107	48	6.9	110
21231	24	8.0	78	86	7.8	95	87	9.7	112
21232	121	8.1	98	142	8.9	100	172	9.1	108
21233	171	6.8	90	244	7.2	121	187	8.9	121
21234	134	7.9	93	188	9.4	117	158	10.5	119
21237	76	5.8	82	127	7.5	106	135	8.4	108
21238	67	5.6	83	99	5.5	96	59	7.4	110
21239	62	6.2	61	117	8.4	91	77	6.4	104
Average	94.7	6.9	87.4	143.4	7.7	103.4	112.7	8.8	109.9