Exploring English Grammar Knowledge and Written Production of Japanese EFL learners

Miki, TOKUNAGA

Exploring English Grammar Knowledge and Written Production of Japanese EFL learners

Miki, TOKUNAGA   (Fukuoka University)

Abstract

To explore which English grammar structures seem to be understood and can be used accurately by Japanese university students, a grammar rule test, a grammaticality judgment test, and a written translation test were administered to 345 university students. The results suggested that participants had enough procedural knowledge on some easy structures such as the regular past, third person -*s* and comparative sentences to benefit from more practice using those structures, while some difficult structures, such as embedded questions, tag questions and relative clauses, may benefit from some instruction for more accurate knowledge. Progressive sentences and plural -*s* showed interesting results, indicating that knowledge of these structures was limited, and some explicit instruction followed by examples and practice may be beneficial.

1. Introduction

In an increasingly global society, command of English is more important than ever. However, the Japanese English education system has often been criticized for failing to produce more Japanese with functional English abilities despite the Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT) shifting its English education guidelines from grammar-focused to communication-focused instruction since the mid 1980's. In 2003, MEXT adopted a five-year Action Plan to cultivate "Japanese with English Abilities" (MEXT 2002). The plan was to ensure that, by 2008, junior high school graduate acquire the ability to hold simple conversations in English and to pass Grade 3 of the EIKEN test (Test in Practical English Proficiency), while senior high school graduates can hold normal conversations on everyday topics and pass Grade 2 or pre-2 of the EIKEN test (MEXT 2002). However, these goals were not met, resulting in the revised five proposals published in 2011 (MEXT 2011), which also have not proven its effect. MEXT's 2015 report shows that only 37% of junior high school seniors and 34% of high school seniors hold the aforementioned grades on the EIKEN (MEXT 2015).

In addition, Japanese students have little opportunity to use English outside classes, and despite the abundance of English content on the internet, only a fraction of learners seem to take advantage of it, possibly because of the contents' difficultly or simply from lack of interest. In a survey of 3,154 high school students (Benesse, 2014), only 1.3 % of students answered that they used materials on the internet to improve their English, and 21.6% answered that they did not study English at all outside classes, including homework. In the same survey, 37.3% answered they did not like the English language itself, 29.6% answered they had no interest in other countries or cultures, and 46.4% answered they expected to use little or no English in real life when they become adults. Thus, despite the government's efforts, the reality of English education at many Japanese universities involves

teaching beginner to low-intermediate level learners in an environment where exposure to English outside classrooms is very low, and the learners' sense of necessity for the language in real life is not very high.

Even in ESL settings, where learners are surrounded in English outside the classroom and opportunities to practice is abundant, acquisition of second language in adults usually requires some explicit learning as what can be acquired from communicative contexts is limited (N. Ellis, 2011). In skill acquisition theory, learners must first acquire knowledge about the language (declarative knowledge), learn how to use that knowledge (procedural knowledge), then, through considerable practice, the knowledge may become automatized (N. Ellis, 2011; DeKeyser, 2015). While proceduralization does not take a lot of time, it is an important stage as automatization does not occur if learners lack the rule usage to be automatized to start with (DeKeyser, 2007). In Japanese EFL education, it seems even more critical for students to gain declarative knowledge through instruction and proceduralize it in classrooms so they can take full advantage of the rare real-life practice opportunities.

Japanese EFL learners are often said to have good knowledge of English grammar but cannot communicate in English because of lack of practice. However, since junior and senior high school English classes' emphasis has shifted from a grammar and reading focus to a communication focus, Japanese students' English levels have reportedly been declining (Ono, 2005; Saita, 2003; Shite, 2007). At the same time, the number of 18-year-old Japanese, which peaked at 2.5million in 1966, has been steadily decreasing since the 1990's to 1.2 million in 2017 (MEXT, 2017). As a result, abilities in all subjects are declining for students entering university (Ishii, Shiina, Maeda, & Yanai, 2007) as universities compete for students to survive financially, accepting students whose scholastic abilities are less than ideal. Thus, it can no longer be assumed that Japanese students know English grammar and just need practice.

To facilitate the process of learning and practicing for Japanese EFL learners, it is important to investigate what learners already know as declarative knowledge and how much of that knowledge has been proceduralized or even automatized. This research aimed to investigate Japanese university students' knowledge of English grammar structures and how well they can reflect their knowledge in their productions through a battery of tests. The objective of this research was to find out which basic grammar structures seem to be in need of more instruction (for declarative knowledge) and/or more practice (for procedural knowledge) in Japanese university English classes to best facilitate language acquisition. There are two research questions:

1. Which basic English grammar structures seem to be understood by Japanese university students?
2. Which basic English grammar structures can Japanese university students use in written tasks?

2. Methods

2.1 Participants

The battery of tests was given to 345 first or second-year students at a private Japanese university taking required English classes, and only the data from participants who signed the consent form were used in analysis. All participants' L1 was Japanese. In the pre-test survey, only 11 students acknowledged having taken TOEIC, of which only five students had scores above 500. On the question asking about their use of English outside English classes, the majority ($n = 290$) answered "none" or "very little," while only eight answered "often".

2.2 Test Batteries

Participants' knowledge was measured by a grammar rule test and an untimed grammaticality judgment test (GJT), while a Japanese to English sentence translation test was used to measure production. To minimize the effect of participants' vocabulary knowledge, Japanese translation accompanied each sentence of the GJT, and some necessary words were given for the production task. After pilot tests, 16 grammar structures to be measured were selected, and test sentences were written to cover those structures. The battery of tests was administered in three separate class periods.

2.2.1 Grammar Rule Test

The grammar rule test used in this study consisted of 24 multiple choice grammar questions on 19 grammar structures: third person -s, adverb placement, base form of a verb after an auxiliary verb, definite article, comparative –er, embedded question, *go + ing*, indefinite article, negative, perfect tense, plural -s, possessive -s, regular past, relative clause, *since/for*, *so/because*, tag question, *than* in comparative sentence, and verb complement. *Go + ing* refers to errors of adding to before gerunds. Japanese students often add *to* after *go* regardless of what follows *go*, creating such erroneous sentence as "*Did you go to shopping yesterday?" An example of the *so/because* structure item is "*I'm going home so I'm tired." This structure was added as confusion of usage between the two were common in students' written homework in the past at the same university. The questions were in Japanese, with English example sentence to show the particular grammar rules. There were two types of questions; one asking for parts of a sentence, the other asking for correct translation of a whole sentence. This depended on the nature of the target grammar structure. The test was conducted in class, with no time limits, and all participants were given the same test. Examples of each type of items from the grammar rule test used in this study are shown below:

Example 1：「Ken はリンゴが好きだ。」を英語にすると、主語が Ken なので動詞（この場合 like）の語尾に(A. ed B. ing C. s)がつく。
Example 2: 「お腹がすいたのでパンを買った。」を英語にすると、(A. I was hungry because I bought some bread. B. I bought some bread because I was hungry. C. I bought some bread, so I was hungry.)

2.2.2 Grammaticality Judgment Test

GJTs have been used in many previous studies of second language acquisition, and while timed GJTs are thought to measure constructs related to implicit knowledge, untimed GJTs are presumed to measure constructs of explicit or declarative knowledge (e.g., Ellis, 2009; Green & Hecht, 1992; Roeher, 2008; Sakai, 2008; Shimada, 2010; Tokunaga, 2013). Even though clear distinctions of constructs measured by timed and untimed GJTs are still being disputed, many previous studies agree that the tests measure different factors (e.g., Ellis, 2005; Godfroid, et al., 2015; Zhang, 2015). However, Shimada (2010) found that the difference in mean scores was larger between grammatical and ungrammatical items than between timed and untimed items among Japanese EFL learners. Gutiérrez (2013) presented a similar conclusion that, even though both time pressure and grammaticality of items significantly affected participants' performance, grammaticality of items had a stronger effect than time pressure. The pilot test of this study with 219 Japanese university students (Tokunaga, 2016) did not find a statistically significant difference in performance between timed and untimed GJTs. Since the GJT was used to measure the participants' explicit or declarative knowledge in this study, only an untimed GJT was used.

Three different forms of GJTs were designed, with each form consisting of 30 items from a 49-item pool covering the 16 grammar structures listed on Table 1. The 16 grammar structures were chosen based on Ellis (2005) and Shimada (2010), plus two added structures: *go + ing*, and *so/because*, the reason for testing both of which are explained in the grammar rule test section above. The dative alteration, ergative verbs, and unreal conditionals were excluded as they seemed too difficult to be considered as basic grammar structures. In fact, in a pilot test (Tokunaga 2013), the first two proved to be too difficult, and unreal conditional sentences tended to be judged as incorrect only because of their unmatched tenses in the two clauses, with very few participants being able to correct the error. The 49-item pool included only 12 grammatical items, because Ellis (2009) had found that grammatical and ungrammatical items in untimed GJT correlated with different knowledges, and used only the ungrammatical items to measure explicit knowledge. Also, in our pilot test with equal numbers of grammatical and ungrammatical items, participants ($N = 179$) tended to choose "correct" more often than "incorrect", with lower ability group ($n = 85$) choosing more "correct" than higher ability group ($n = 94$), indicating that they might be choosing "correct" whenever an obvious error could not be found, not because the sentences seemed correct. Each of the three forms of the tests consisted of 30 items with 24 incorrect items, of which nine were common items between the forms, and six correct items. The data from the three forms were merged using the nine common items and other overlapping items between two of the forms. This was done to keep each test relatively short, so the participants could stay focused throughout the tests, while having more than one sentence for each target grammar point.

Table 1. Grammatical Structures Included in the GJT

Ellis (2005)	Current Study
Adverb placement	Adverb placement
Comparatives	Comparatives
Dative alternation	
Embedded questions	Embedded questions
Ergative verbs	
Indefinite article	Indefinite article
Modal verbs	Auxiliary + base form
Plural *–s*	Plural *–s*
Possessive *–s*	Possessive *–s*
Questions tags	Tag questions
Regular past	Regular past
Relative clauses	Relative clauses
Since and for	Since/for
Third person *–s*	Third person *–s*
Unreal conditionals	
Verb complement	Verb complement
Yes/no questions	Yes/no questions
	So/because
	Go + ing

Each sentence in the GJT was accompanied by its Japanese translation in an attempt to minimize the effect of participants' vocabulary knowledge and reading ability, and to focus on their understanding of target grammar structures. Ellis (2004) stated that GJTs potentially involve three processing operations: understanding the meaning of a sentence (semantic), deciding whether something is formally incorrect in the sentence (noticing), and considering what is incorrect and why (reflecting). However, the first step involves more than the test takers' grammatical knowledge, and for EFL learners of beginner to intermediate levels of English, the first step could prevent them from moving to step two, where they are to judge the grammaticality of the sentence. An example item from the GJT used in this study is shown below:

<div align="center">* Aya like English.　　Aya は英語が好きだ。</div>

The test was conducted on paper, and participants were first asked to judge whether each sentence was grammatically correct, then to correct the error for those sentences they had marked incorrect. The test was started around 30 minutes before the end of the class period, and the participants were required to stay for 20 minutes. After the initial 20 minutes, participants who had completed all the items were allowed to leave the classroom. In all groups, only a few students had to stay beyond the first 20 minutes, and no one needed more than 30 minutes. The results were analyzed for judgment, and for correction if the sentence was ungrammatical.

2.2.3 Japanese to English Sentence Translation Test

The written translation test used 25 sentences from the GJT covering the 16 grammar structures. Three sets of tests with 15 sentences each were designed, and each participating group was given a different set of sentences from the GJT, except for the common items. Five common items and additional matching items between two of the three forms were used to merge the results. The nature of the translation test increased the variety of possible errors, and 53 items including 22 grammar structures were checked and recorded. The 22 grammar structures were the 16 structures in GJT plus *go home*, negative sentence, possessive pronoun, preposition, progressive, and *there is*. As Japanese learners of English tend to add *to* in *go + ing* phrases mentioned earlier, they also tend to write *Go to home*, so this was checked. Starting a sentence with *there is* was difficult in a pilot test, so this structure was also checked.

The test was conducted on paper, and the participants were asked to translate the 15 sentences into English with no time limit. Each Japanese sentence was accompanied by essential content words in English to prevent the participants from not producing any sentence. The content words were given in alphabetical order and in base form, and the participants were instructed to change the forms when necessary. An example item from the translation test, including third person -*s* and plural -*s* is shown below:

<div align="center">Lisa は犬を２匹飼っている。　dog, have</div>

The results were analyzed for the 53 items and 25 sentences. For example, if a participant wrote the third person -*s* correctly but missed the plural -*s*, the score was 1 for third person -*s*, 0 for plurals -*s*, and 0 for the sentence.

3. Results

Rasch analysis using the Winsteps® software package (Linacre, 2017a) was conducted on the data from all the tests. While adjacent raw scores do not represent equal intervals of knowledge, Rasch analysis produces measures of item difficulty and person ability on a common equal interval scale measured in log-odds units, or *logits*. The mean item difficulty is set to 0 logits by convention, and the distance between each interval on the scale represents an equal difference in ability or difficulty. Item difficulties in logits are used in measurement of items, but percentages correct are also shown on each table on this paper for those who are not familiar with Rasch measurement. However, as described in the test batteries section, not all items have been taken by the same number of participants and percentages are not adjusted for those differences, whereas the calculation of logit measures automatically adjusts for the difficulty of different test forms and the ability of the students who took each item.

Table 2 shows the summary statistics of the tests. The total number of participants varied for the different tests as some were absent when the tests were conducted in class. As mentioned above, mean item difficulty is conventionally set to 0 logits. In Rasch analysis, person and item reliabilities do not report on the quality of the data. Rather, they show the reproducibility of the results. Thus, a high reliability coefficient means that persons (or items) estimated to have high measures actually do have

higher measures than persons (or items) estimated with low measures (Linacre, 2017b). Person separation and item separation show how many statistically significant levels the persons and the items can be separated into by the test.

Table 2. Summary Statistics for the Rule Test, the GJT, and the Written Translation Test (N = 345)

Tests	Items / Persons	n	M (logit)	SE	SD	Separation	Rasch Reliability
Rule Test	Items	24	0.00	0.25	1.51	5.22	.96
	Persons	325	2.84	0.86	1.03	1.05	.52
GJT	Items	49	0.00	0.22	1.16	4.79	.96
(Judge Only)	Persons	326	1.21	0.51	1.04	1.83	.77
GJT	Items	37	0.00	0.20	1.33	5.98	.97
(Correction)	Persons	326	0.36	0.56	1.40	2.31	.84
Translation	Items	53	0.00	0.27	1.56	4.90	.96
(Items)	Persons	333	1.70	.56	1.13	1.81	.77
Translation	Items	25	0.00	0.23	1.72	6.82	.98
(Sentences)	Persons	333	0.37	0.72	1.33	1.69	.74

Item reliabilities are higher than person reliabilities in all tests as there were more participants to measure the difficulty of the items than items to measure abilities of participants. Low person separation (< 2.0, corresponding to reliability of < .80) means that the tests are not sensitive enough to distinguish between high and low performers, while high item separation (> 3.0, corresponding to reliability of > .90) indicates that there are enough participants to confirm the item difficulty hierarchy (Linacre, 2017b). As the main purpose of this research project is to measure the difficulties of different grammar structures and not to rank the participants' L2 levels, the low person reliability and separation do not substantively affect the measurement of item difficulty.

3.1 Grammar Rule Test

Figure 1 shows the Winsteps variable map of the rule test with persons ranked by ability on the left and items by difficulty on the right. Each "X" represents one item, while "." and "#" represent one and two participants respectively. A position higher on the scale represents greater ability for a person or greater difficulty for an item. When person ability precisely matches item difficulty, the person has a 50% expectation of success on the item. The M on each side of the axis shows the mean of person ability and item difficulty. On Figure 1, it is clear that participants' mean ability was higher than the mean item difficulty of items on the rule test, indicating that this rule test was easy for most of the participants.

However, there were clear ranges in difficulties of items. Table 3 shows the difficulty of the 24 items from the rule test. The most difficult item was the plural -*s* in the sentence, "Does Ken like apples?"(3.12 logits). This item had a much higher difficulty than the other plural -*s* in "Ken ate three apples" (0.27 logits) where the number made plurality of the noun very clear. The second most difficult item was *go – ing* in "I went skiing" (2.89 logits), with 141 participants answering, "I went to skiing" as correct.

```
INPUT: 345 Person  24 Item  REPORTED: 325 Person  24 Item  2 CATS WINSTEPS 3.93.2
------------------------------------------------------------------------------

Higher ability                              |                        More difficult
   <logit> ----------------- Person -+- Item ----------------- <logit>
      5              .################## +                            5
                                        T|
      4       ######################### +                            4
                                        S|
              .########################  |
      3       .########################+T XX                         3
              .######################## M|
              .######################## |
      2          .################### + X                            2
                      ##########      |S X                           2
                       ########      S|
      1              ########## + X                                  1
                       ####     | XXX
                        ###   T| XXXXX
      0                  #   +M XXX                                  0
                         #    |
                              | X
     -1              + X                                            -1
                          | X
                          |S XX
     -2              +                                             -2
                          | XX
                          |
     -3              +T                                            -3
                          | X
     -4              +                                             -4
   ----------------- Person -+- Item  -----------------
Lower ability                                         Less difficult
EACH "#" IN THE Person COLUMN IS 2 Person: EACH "." IS 1
```
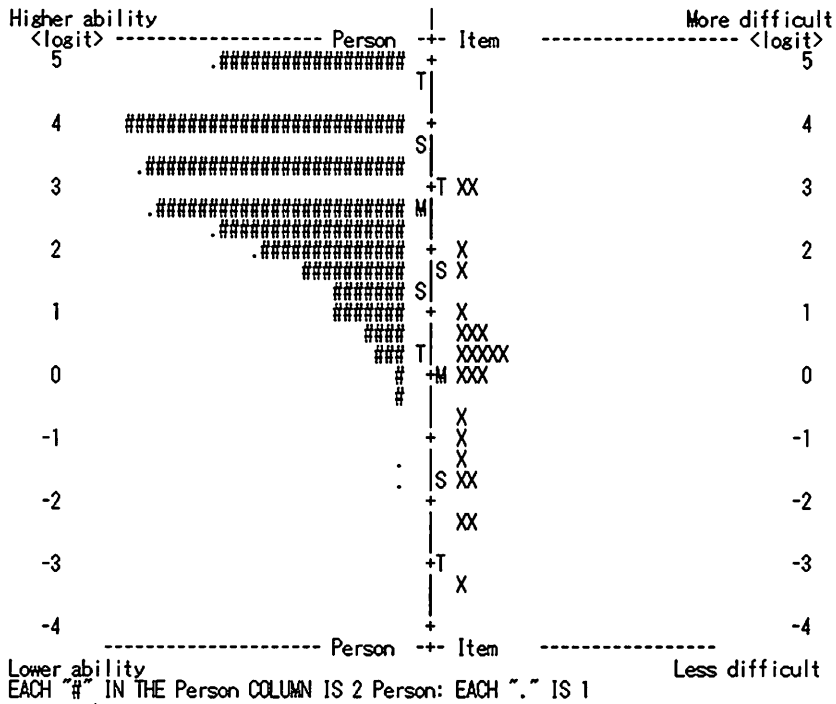
Figure 1.   Variable map of the rule test showing abilities of 325 participants on the left and difficulties of 24 items on the right on an equal interval scale. Each "#" represents two persons and each and "." equals one person, and each "X" is one item.

Table 3. Item Difficulties of the Rule Test in Order of Difficulty (*N* = 325)

Grammar Structure	Item Difficulty (logit)	% Correct	Grammar Structure	Item Difficulty (logit)	% Correct
plural -*s* (unclear)	3.12	45.5	perfect	0.17	88.6
go + ing	2.89	49.5	adverb placement (before verb)	0.04	89.8
relative clause (object)	1.92	66.2	tag question	0.01	89.8
embedded question	1.83	67.7	verb complement	-0.11	90.8
adverb placement (after be verb or auxiliary verb)	0.99	79.9	negative (doesn't)	-0.52	93.4
indefinite article	0.75	82.7	comparative -*er*	-1.05	95.7
relative clause (subject)	0.56	84.9	possessive -*s*	-1.33	96.6
negative (isn't)	0.54	85.2	aux. + base form	-1.56	97.2
negative (base form after doesn't)	0.40	86.5	definite article	-1.62	97.5
plural -*s* (obvious)	0.27	87.7	comparative *than*	-2.21	98.5
since/for	0.27	87.7	third person -*s*	-2.45	98.8
so/because	0.27	87.7	regular past	-3.18	99.4

Other high difficulty grammar structures were relative clauses (1.92 and 0.56 logits), embedded question (1.83 logits), adverb placement (0.99 logits), indefinite article (0.75 logits) and negatives (0.54 and 0.40 logits). The easiest grammar structures were regular past (-3.18 logits), third person -s (-2.45 logits), comparative *than* (-2.21 logits), and definite article (-1.62 logits.) Both the indefinite and definite articles were in the same sentence 「その山の上に城がある。」 [There is ( ) castle on ( ) mountain]. The question in Japanese specified that the mountain was a particular mountain but the castle was an unidentified one, and the answer choices for both items were a, *the*, and "nothing is needed." The definite article *the* has a Japanese equivalent その in this sentence, which may have helped it to be easier than the indefinite article.

3.2 Grammaticality Judgment Test (Judge only)

Tables 4 and 5 list the 15 most difficult items and 10 easiest items from the 49 grammaticality judgment items. In this section of the test, participants were only asked to judge whether each sentence was grammatically correct, and each sentence had a Japanese translation. It is easily seen on Table 4 that 14 of the 15 most difficult items were ungrammatical sentences, and six of the 10 easiest items were correct sentences. This indicates that judging grammatical items as correct was easier than judging ungrammatical items as incorrect, or that participants tended to mark "correct" unless they noticed an obvious error, as found in the pilot test.

Table 4: The Most Difficult 15 Items in GJT Judgment ($N = 326$)

Grammar Structure	Sentence	Item Difficulty (logit)	% Correct (*n* varies)
embedded question	*Does she know who are you?	2.48	20.7
relative clause	*The man lives in that house is very tall.	2.48	20.7
indefinite article	*There is concert tonight.	2.24	29.8
indefinite article	*I bought a new shoes yesterday.	2.21	30.4
relative clause	*The woman works in that cafe is very kind.	1.73	41.0
embedded question	*Do you know what is this?	1.62	43.2
indefinite article	*He is riding red bicycle.	1.60	41.3
plural -s	*Do you like cat?	1.51	42.9
go + ing	*Did you go to shopping yesterday?	1.20	49.1
Plural -s	*How many student were late for class?	1.06	51.8
embedded question	Do you know what this is?	0.96	48.9
indefinite article	*She is reading new book.	0.95	54.7
aux. + base form	*Will she comes to the party?	0.89	55.2
tag question	*John likes beer, isn't he?	0.54	62.3
verb complement	*Jane enjoyed to study French.	0.50	58.7

The most difficult items were embedded questions and relative clause sentences (both 2.48 logits) followed by two indefinite article sentences (2.24 and 2.21 logits), another relative clause sentence (1.73 logit) and another embedded question (1.62 logit). The 326 participants encountered one of the two relative clause items and one of the two embedded question items depending on the GJT form, but only 115 judged the relative clause items as incorrect and 120 did so for the embedded question items. Even though the base form of a verb after an auxiliary verb was an easy item on the rule test, finding the error in "*Will she comes to the party?" turned out to be difficult (0.89 logit), with only 180 out of 326 participants marking the item incorrect. The easiest items included a grammatical verb complement sentence (-2.20 logits), a grammatical *go –ing* sentence (-1.96 logits), and an ungrammatical regular past sentence (-1.87 logits).

Table 5: The Easiest 10 Items in GJT Judgment (*N* = 326)

Grammar Structure	Sentence	Item Difficulty (logit)	% Correct (*n* varies)
verb complement	She finished writing the letter.	-1.09	86.7
so/because	He was sick, so he didn't go to work.	-1.13	85.9
since/for	We have lived in this town since 2010.	-1.14	88.0
adverb placement	Amy sometimes cooks dinner.	-1.28	89.5
regular past	*Takuya clean his room yesterday.	-1.38	90.2
third person *-s*	*Aya like English.	-1.64	91.7
possessive *-s*	*This is my mother car.	-1.72	91.3
regular past	*Hiroshi receive a letter yesterday.	-1.87	92.4
go + ing	Let's go camping this weekend.	-1.96	94.2
verb complement	He finished reading the report.	-2.20	95.3

3.3 Grammaticality Judgment Test (Correction)

Tables 6 and 7 list the 15 most difficult and 10 easiest items from the correction part of the grammaticality judgment test. Out of 49 sentences, 12 sentences were grammatically correct, thus this part of the test involved 37 items. As participants did not make corrections on items they thought were correct, similar grammar structures are on the list of difficult items as the judgment section of the test. The most difficult error to correct was adding "a" to "*He is riding red bicycle" (2.61 logits). Even though the general difficulty of using articles contributed for the results with many participants not noticing the error, the highest number of wrong correction was adding *on* before bicycle, occurring 16 times out of 240. Also, 15 participants changed *is riding* to *rides*, and three participants changed to *ride*. For a similar present progressive sentence with missing article, "*She is reading new book" (0.62 logits) the incorrect change to *reads* occurred only twice out of 86 participants who had this item on their version of the GJT.

Even though the *so/because* item, "*I'm going home so I'm tired" did not seem so difficult in judgment (-1.13 logits), the correction part of the test revealed that many participants judged the sentence for wrong

reasons. Although 219 out of 326 participants marked this sentence as incorrect, only 109 could change the sentence correctly, making the correction the eighth most difficult (1.33 logits). The most frequent wrong correction made by 52 participants was adding *to* before *home*, making the sentence, "*I'm going to home so I'm tired." Other common erroneous changes were changing *I'm going* to *I go*, and *going home* to *going to go home*. The base form after an auxiliary verb item mentioned earlier, "*Will she comes to the party?," appears as one of the most difficult again (0.23 logits), but this is mostly because participants did not make corrections. Out of the 180 participants who judged this sentence as incorrect, 172 made the correct changes.

The easiest errors to correct were "*This is my mother car" (-2.71 logits), "*Hiroshi receive a letter yesterday" (-2.31 logits) and "*Aya like English" (-2.28 logits). Even though the unclear plural *-s* in "*Do you like cat?" was one of the most difficult in both judging (1.51 logits) and correcting (1.87 logits), the obvious plural *-s* in "*Lisa has two dog" and "*Aki has two brother" were both easy to correct (-1.80 and -1.25 logits).

Table 6: The Most Difficult 15 Items in GJT Correction (*N* = 326)

Grammar Structure	Sentence	Item Difficulty (logit)	% Correct (*n* varies)
indefinite article	*He is riding red bicycle.	2.61	15.8
relative clause	*The man lives in that house is very tall.	2.61	10.9
indefinite article	*There is concert tonight.	2.34	19.0
embedded question	*Does she know who are you?	2.08	16.3
plural *-s*	*Do you like cat?	1.87	25.2
indefinite article	*I bought a new shoes yesterday.	1.64	28.5
relative clause	*The woman works in that cafe is very kind.	1.41	34.6
so/because	*I'm going home so I'm tired.	1.33	33.4
embedded question	*Do you know what is this?	1.00	41.5
plural *-s*	*How many student were late for class?	0.83	42.0
go+ing	*Did you go to shopping yesterday?	0.64	45.4
indefinite article	*She is reading new book.	0.62	46.5
aux. + base form	*Will she comes to the party?	0.23	52.8
tag question	*She wasn't crying, did she?	0.18	54.7
verb complement	*Jane enjoyed to study French.	0.17	48.9

Table 7: The Easiest 10 Items in GJT Correction ($N = 326$)

Grammar Structure	Sentence	Item Difficulty (logit)	% Correct (*n* varies)
comparative	*Ken is more old than Aya.	-0.76	70.5
adverb placement	*Amy cooks sometimes dinner.	-0.90	71.3
comparative	*Mary is more tall than Tom.	-0.91	69.6
plural -*s*	*Aki has two brother.	-1.25	77.9
auxiliary + base	*I must getting up early tomorrow.	-1.34	79.1
regular past	*Takuya clean his room yesterday.	-1.74	82.5
plural -*s*	*Lisa has two dog.	-1.80	82.5
third person -*s*	*Aya like English.	-2.28	87.4
regular past	*Hiroshi receive a letter yesterday.	-2.31	88.0
possessive -s	*This is my mother car.	-2.71	91.3

3.4 Japanese to English Sentence Translation Test (Items)

The translation test used the same item pool as the GJT. However, as described earlier, each participant was given a different set of sentences from the GJT, except for the five common items. Even though participants had seen the five common items in the GJT, they were not told whether those were correct sentences until after the translation test. The result of *t*-tests conducted to compare the mean difficulties between the five common sentences and the other 20 sentences, and between 11 items on common sentences and 42 items on other sentences showed no statistically significant differences (Table 8).

Table 8: Welch's *t*-tests Comparing 5 Common Sentences and the Other 20 Sentences on the Translation Test

	n	Mean Measure (logit)	*SD*	*t*	*p*
Common sentences	5	.46	.81	1.01	.33
Other sentences	20	-.11	1.89		
Items on common sentences	11	.39	.87	1.27	.21
Items on other sentences	42	-.10	1.72		

Each participant was asked to translate 13 Japanese sentences into English, and most content words in English were listed with the Japanese sentence. Tables 9 and 10 show the difficulties of 15 most difficult and 10 easiest items out of 53 items on the Japanese to English written translation test with target grammar structures capitalized for clarity. As described in the test batteries section, each sentence had more than one target structures. For example, the sentence "He is riding a red bicycle" was checked for the present progressive and the indefinite article, and score was given for each item.

Surprisingly, the most difficult item was the present progressive in "He is riding a red bicycle" (3.43 logits). It is possible that its Japanese sentence 「彼は赤い自転車に乗っている。」 may be confused

with a simple present tense as in "He has a red bicycle," so this may have been a bad item. However, even though not among the 15 most difficult, the present progressive in 「彼女は新しい本を読んでいる。」 ("She is reading a new book") was also not easy (0.57 logits), with 75 out of 243 participants not being able to write it correctly. Also, "She wasn't crying, was she?" was checked for progressive after the initial analysis. Because the target structure for "She wasn't crying, was she?" was negative and a tag question, the initial analysis counted "She didn't cry, did she?" as correct even though the Japanese sentence 「彼女は泣いていませんでしたよね。」 clearly indicates past progressive. Even though 120 of 234 participants were marked correct for negative structures, it turned out that only 16 of them wrote "She wasn't crying", and 94 wrote "She didn't cry." Other correct negative responses included "She isn't crying" and "She doesn't cry." Confusion on the progressive form was also seen, with 12 participants writing "*She isn't cry," 7 writing "*She wasn't cry," and 5 writing "*She doesn't crying."

Table 9: The Most Difficult 15 Items in the Written Sentence Translation Test ($N = 333$)

Grammar Structure	Sentence (Target Structures Capitalized)	Item Difficulty (logit)	% Correct (n varies)
progressive	He IS RIDING a red bicycle.	3.43	25.0
embedded question	Do you know what THIS IS?	3.05	20.0
negative	You AREN'T angry, are you?	2.73	36.8
indefinite article	He is riding A red bicycle.	2.55	39.8
indefinite article	There is A concert tonight.	2.28	39.3
preposition	The man who lives IN that house is very tall.	2.04	47.5
preposition	The woman who works IN/AT that cafe is very kind.	1.94	39.4
relative clause	The woman WHO WORKS in that cafe is very kind.	1.64	45.7
tag question	John likes beer, DOESN'T HE?	1.62	57.3
embedded question	Does she know who YOU ARE?	1.56	55.7
negative	She WASN'T crying, was she?	1.56	51.3
relative clause	The man WHO LIVES in that house is very tall.	1.55	56.5
tag question	She wasn't crying, WAS SHE?	1.10	64.9
indefinite article	She is reading A new book.	1.09	59.7
third person -s	Amy sometimes cookS dinner.	0.88	69.7

Table 10: The Easiest 10 Items in the Written Sentence Translation Test (N = 333)

Grammar Structure	Sentence (Target Structures Capitalized)	Item Difficulty (logit)	% Correct (*n* varies)
comparative	Ken is oldER than Aya.	-1.36	92.6
aux. + base form	I must GET up early tomorrow.	-1.82	94.7
possessive pronoun	This is MY mother's car.	-2.11	95.7
adverb placement	Amy SOMETIMES cooks dinner.	-2.26	96.6
adverb placement	Tom SOMETIMES plays tennis.	-2.51	97.1
comparative	Mary is tallER than Tom.	-2.56	97.0
plural -*s*	Lisa has two dogS.	-2.67	97.7
comparative	Ken is older THAN Aya.	-2.70	97.9
possessive pronoun	Jack is MY older brother's friend.	-2.71	97.9
comparative	Mary is taller THAN Tom.	-2.93	97.9

Other grammar structures with high difficulties included embedded questions (3.05 and 1.56 logits), negative (2.73 and 1.56 logits), indefinite articles (2.55, 2.28 and 1.09 logits), prepositions (2.04 and 1.94 logits), relative clauses (1.64 and 1.55 logits) and tag questions (1.62 and 1.10 logits). The high difficulty of the negative form in "You aren't angry, are you?" was surprising, with only 32 out of 87 participants writing the correct form. Wrong responses varied from "You don't angry," "You not angry," to "Don't you angry?" For indefinite articles and tag questions, avoidance was more common than writing wrong forms. Although adding *to* before *home* lead to wrong corrections in the GJT, being able to write *go home* instead of **go to home* was not one of the 15 most difficult items, with 274 out of 333 participants writing correctly (-0.29 logits). The easiest structures included *than* in comparative sentences (-2.93 and -2.70 logits), possessive pronoun *my* (-2.71 and -2.11 logits), obvious plural -*s* (-2.67 logits), adverb placements (-.2.51 and -2.26 logits) and the base form of a verb after an auxiliary verb (-1.82 logits).

3.5 Japanese to English Sentence Translation Test (Sentences)

For this part of the translation test, each sentence as a whole was scored as correct (scored 1) or incorrect (scored 0) regardless of how many target structures were involved in the sentence. Figure 2 shows the Winsteps variable map of the written sentence translation test. As in Figure 1, persons are ranked on the left and items are on the right, with each "X" representing one item and "." and "#" representing one and two participants. Compared with the rule test (Figure 1), it is clear that sentence translation was more difficult, with abilities of participants more spread out and some participants with lowest abilities having no items easy enough to match their ability.

```
INPUT: 345 Person  25 Item  REPORTED: 333 Person  25 Item  2 CATS WINSTEPS 3.93.
--------------------------------------------------------------------------------

Higher ability                                        More difficult
  <logit> -------------------- Person -+- Item  -------------------- <logit>
    4                          ###     +                                4
                                       |  X
                              #### T|T
    3                          .### +                                   3
                                       |
                         .######### |  XX
    2                        .## + X                                    2
                     ############## S|S X
                        ############# |  X
    1            .################# +  XXX                              1
                      .###### |  XX
                      .###### M|  X
    0            ############### +M X                                   0
                  ############### |  XX
                   ############### |  XX
   -1               .###### S+  X                                      -1
                       ###### |  X
                         ### |S X
   -2                 ##### + X                                        -2
                                       |  XX
                      .##### T|  X
   -3                        +  X                                      -3
                                     |T X
                          #           |
   -4                     #  +                                         -4
            -------------------- Person -+- Item  --------------------
Lower ability                                           Less difficult
EACH "#" IN THE Person COLUMN IS 2 Person: EACH "." IS 1
```
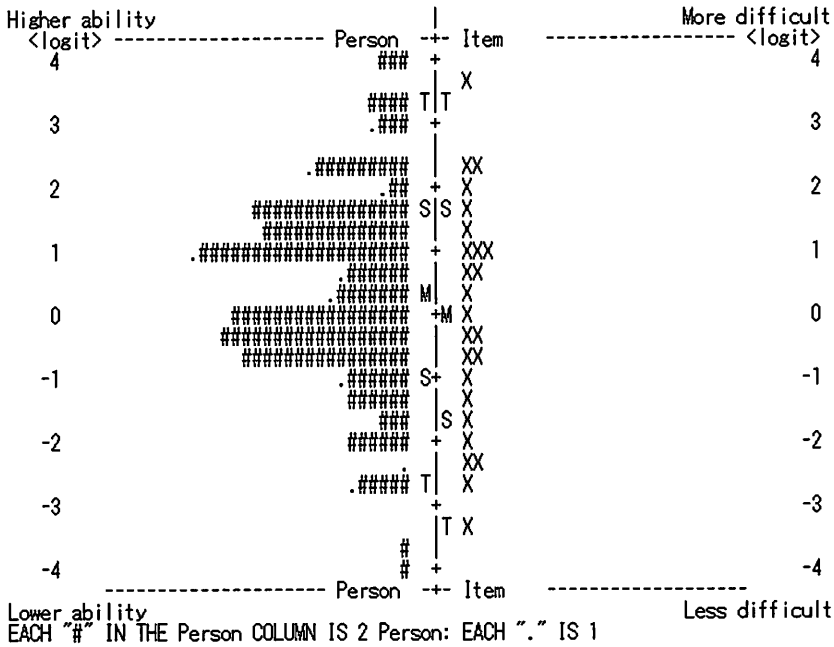
Figure 2.   Variable map of the written sentence translation test showing abilities of 333 participants on the left and difficulties of 25 sentences on the right on an equal interval scale. Each "#" represents two persons and each and "." equals one person, and each "X" is one sentence.

Table 11: Item Difficulties of the Written Sentence Translation Test in Order of Difficulty ($N = 333$)

Sentence	Item Difficulty (logit)	% Correct (n varies)
He is riding a red bicycle.	3.65	10.2
The woman who works at that cafe is very kind.	2.45	12.8
The man who lives in that house is very tall.	2.36	22.9
You aren't angry, are you?	1.9	31.0
Do you know what this is?	1.82	20.0
She wasn't crying, was she?	1.17	35.0
There is a concert tomorrow.	1.14	37.4
John likes beer, doesn't he?	1.06	39.1
He has been sleeping for 12 hours.	0.91	41.5
She is reading a new book.	0.73	42.0
Does she know who you are?	0.72	47.8
He should finish his homework.	0.17	59.6
Amy sometimes cooks dinner.	-0.03	62.9
I'm going home because I'm tired.	-0.24	61.3

Tom sometimes plays tennis.	-0.43	62.6
Jane enjoyed studying French.	-0.51	66.7
Did you go shopping yesterday?	-0.58	67.2
Jack enjoyed watching baseball.	-0.85	68.8
This is my mother's car.	-1.43	79.1
Aki has two brothers.	-1.63	80.2
Jack is my older brother's friend.	-1.94	84.4
I must get up early tomorrow.	-2.22	86.4
Lisa has two dogs.	-2.26	88.8
Ken is older than Aya.	-2.63	90.5
Mary is taller than Tom.	-3.33	93.7

Table 11 shows the 25 sentences in order of difficulty. Requiring both the present progressive and the indefinite article, the most difficult sentence was "He is riding a red bicycle" (3.65 logits). This was an unexpected result at the item planning stage, but not surprising given the results of the analysis of difficulty of each grammar structure in this section and in GJT (Tables 4, 6 & 9). The progressive sentence was followed by two relative clause sentences (2.45 and 2.36 logits) for which relative pronoun plus verb and prepositions were among the 15 most difficult structures. Tag questions and embedded questions were difficult throughout the test battery, thus the higher rankings of "You aren't angry, are you?" (1.90 logits), "Do you know what this is?" (1.82 logits) and others were as expected. The easiest sentences were two comparative sentences, "Mary is taller than Tom" (-3.33 logits) and "Ken is older than Aya" (-2.63 logits), followed by "Lisa has two dogs" (-2.63 logits), which had third person -s and a clear plural -s following a plural number.

4. Discussion

This research attempted to explore which English grammar structures were understood and could be used accurately by Japanese university students. Their knowledge was measured with a multiple-choice grammar rule test and a GJT which involved judging and correcting ungrammatical sentences, while production was measured with a Japanese to English written translation task.

Relative clauses, embedded questions, tag questions and indefinite articles were difficult for many of the participants throughout the battery of tests, suggesting that learners may benefit from some explanations or instruction on these grammar structures to help gain declarative knowledge which will help them practice them more accurately. However, it may be necessary to consider whether these structures are needed at the current English proficiency of learners. For learners who lack declarative and procedural knowledge of easier structures, explanations of more difficult structures may not make much sense. For example, as simple negative sentences ranked high on difficulty on the translation task (Table 9), it seems reasonable to help learners acquire procedural knowledge of how to make negative sentences before moving on to explaining tag questions which requires the knowledge of whether a be verb, *do*, *does*, *did*, or an auxiliary verb should be used to make a certain sentence negative.

The regular past, third person -s, and comparative's –er and *than* were among the easiest throughout the test battery. However, the percentages correct were lower on the production task, indicating that even those easy structures could use more production practice in class for stronger procedural and automatized knowledge. Even though using the base form of the verb after an auxiliary verb was easy in the rule test and the production test, finding and correcting the error in "*Will she comes to the party?" was difficult in the GJT. This may be the result of lack of input or practice of question forms of the structure.

There were a few structures that had interesting results. One was the plural -s and the other was progressive sentences. Even though the plural -s following a clear plural number, such as "two brothers" and "two dogs" was easy, the unclear plural -s in "Do you like cats?" and "How many students were late for class?" was difficult. This means that even though learners understood that plural nouns take s at the end of the noun, they didn't know in what situations plural nouns are used. The sentence "*He is riding red bicycle," whose target structure in the GJT was the missing indefinite article *a*, resulted in many correction of its tense to "rides" and "ride," ignoring the article. This sentence was also the most difficult in the written translation task. Although not as difficult, similar confusion of the tense was seen in "She is reading a new book," and "She wasn't crying, was she?" Because there were more errors using different tense than erroneous forms such as "*She isn't cry," this also seems that learners know the structure, but do not know when they are used. These structures may benefit from some explicit explanations on when they are used, followed by examples in reading and listening materials.

These results, including some unexpected findings, suggest the importance of understanding which grammar structures are understood and can be used by learners in order to best support their language acquisition in class. Even though promoting grammatical accuracy too much for low-proficiency learners would be counterproductive and discouraging, assuming that learners would acquire basic grammar through other activities is unrealistic when the majority (84.0%) of the participants in this study answered they have little or no contact with English outside school. Those students need enough declarative knowledge of basic English grammar structures so they can understand other activities in class and start developing procedural knowledge. Also, learners with higher proficiency and motivation could benefit from some explicit instructions on difficult grammar structures, declarative knowledge of which could help them notice those structures in reading or listening to authentic materials and develop into procedural and automatized knowledge. To help learners become equipped with enough knowledge to take advantage of real life practices when opportunities or necessities arise, understanding learners' current knowledge and providing appropriate instructions or activities is essential.

Because this study only used an untimed written translation task for production, it is not possible to distinguish between learners' procedural knowledge and automatized knowledge. For further research, spontaneous production tests, both written and spoken will be necessary to examine which structures, if any, are automatized by Japanese university students with limited exposure to English.

References

Benesse (2014). Chukousei no eigo gakusyu ni kansuru jittai chousa 2014 [Survey on English studies of junior and senior high school students 2014]. Retrieved from http://berd.benesse.jp/global/research/detail1.php?id=4356

DeKeyser, R. (2007). Study abroad as foreign language practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208-226). New York, NY: Cambridge University Press.

DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition; An introduction* (2nd ed.; pp. 94-112). New York, NY: Routledge.

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

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning, 54,* 227-275. doi: 10.1111/j.1467-9922.2004.00255

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*(2), 141-172. doi: 10.1017/S0272263105050096

Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, J. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 31-64). Bristol: Multilingual Matters.

Godfroid, A., Loewen, S., Jung, S., Park, J., Gass, S. & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition, 37*(2), 269-297. doi: 10.1017/S0272263114000850

Green, P., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics, 13*(2), 168-184. doi: 10.1093/applin/13.2.168

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measurers of implicit and explicit knowledge. *Studies in Second Language Acquisition, 35*(3), 423-449. doi: 10.1017/S0272263113000041

MEXT (2002). *Japanese government policies in education, culture, sports, science and technology 2002.* Retrieved from http://www.mext.go.jp/b_menu/hakusho/html/ hpac200201/index.html

MEXT (2011). *Five proposals and specific measures for developing proficiency in English for international communication.* Retrieved from http://www.mext.go.jp /component/english/__icsFiles/afieldfile/2012/07/09/1319707_1.pdf

MEXT (2015). *Heisei 27 nendo Eigo kyoiku jisshi joukyou cyousa no kekka ni tuite.* [Results of English education implementation survey 2015]. Retrieved from http://www.mext.go.jp/a_menu/kokusai/gaikokugo/1369258.htm.

MEXT (2017) *Koutou kyouiku no shourai kousou ni kansuru kiso deta.* [Basic data on future structures of higher education]. Retrieved from http://www.mext.go.jp/ b_menu/shingi/chukyo/chukyo4/gijiroku/__icsFiles/afieldfile/2017/04/13/1384455_02_1.pdf

Linacre, J. M. (2017a). Winsteps (Version 3.93.2)[Computer software]. Retrieved from http://www.winsteps.com

Linacre, J. M. (2017b). *A user's guide to Winsteps Ministep Rasch-model computer programs 3.92.0.* Retrieved from http://www.winsteps.com

Ono, H. (2005). Placement test kara akiraka ni natta nihon no daigakusei no kisogakuryoku kouzou. [Japanese university students' academic abilities based on placement tests]. *NIME research report, 6,* 1-6.

Saita, T. (2003). Koko nyugakuji no eigo nouryokuchi no nenji suii. [Annual transitions of English ability at entering high schools]. *Step Bulletin, 15,* 12-22.

Sakai, H. (2008). Implicit and explicit grammatical knowledge of L2 English: Identification, correction, and provision of rules. *Annual Review of English Language Education in Japan, 19,* 91-100.

Shimada, K. (2010). Bunpousei hanndan tesuto ni okeru monndai teiji jikannseigen no umu to meijiteki annjiteki chishiki. [Timed / timed grammaticality judgment test and implicit /explicit knowledge]. *St. Andrew's University English Review, 24,* 41-53.

Shite, K. (2007). Gakuryokusaikou: gakusei no eigoryoku no genjou to kanren sasete no ichi kousatsu. [Reconsidering academic ability: in relation to the present state of university students' English proficiency.] *Kokusai Keiei Ronshu, 34,* 109-118.

Tokunaga, M. (2013). Gakusei no bunpousei handan ryoku: Shikenteki chosa. [Students' ability in grammaticality judgment.] *FULERC Annual Review of Language Learning and Teaching, 12,* 151-162.

Tokunaga, M. (2016). Effect of time pressure on grammaticality judgment tests with L1 translation. *Annual Review of English Learning and Teaching, 21,* 1-12.

Zhang, R. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition, 37*(3), 457-486. doi: 10.1017/S0272263114000370