What can Latent Rank Theory contribute to SLA research?

Hiromasa Ohba*, Kenichi Yamakawa**, Yuko Shimizu ***,
Naoki Sugino****, and Michiko Nakano*****

Joetsu University of Education
hohba@juen.ac.jp
**Yasuda Women's University*
kyamakaw@yasuda-u.ac.jp
***Ritsumeikan University*
yukos@ec.ritsumei.ac.jp
****Ritsumeikan University*
gwisno@is.ritsumei.ac.jp
*****Waseda University*
nakanom@waseda.jp

Abstract
In elucidating the developmental processes of overall grammatical proficiency of Japanese EFL learners, the possibility of the application of a new test theory, Latent Rank Theory (LRT), has been investigated. In this symposium, by applying LRT, we will attempt (a) to identify levels of grammatical proficiency and differentiate them in terms of what learners can do, and (b) to characterize individual learners by utilizing achievement descriptors obtained through LRT. To this end, we will concisely explain LRT, and then present several studies which attempted to apply LRT to analyzing the results of the grammaticality judgment tests of English verbs, and *wh*-question and relative clause constructions as well as the results of the National Center Test for University Admissions.

I. Introduction

Educational and psychological measurement is always afflicted with several problems when making decisions based on a certain score, which is usually computed using a limited sample of the observations. Crocker and Algina (1986: 6-7) show several measurement problems common to all psychological assessment and emphasize the importance of the role of test theory that supports test use and test construction. In our research project to elucidate the developmental processes of overall grammatical proficiency of Japanese adult learners of English, we will focus on the potential for a new test theory, Latent Rank Theory (hereafter LRT), to serve as a better tool to clarify the hard-to-define constructs of their grammatical proficiency.

By applying LRT, therefore, we will attempt (a) to identify levels of grammatical proficiency and differentiate them in terms of what learners can do, and (b) to characterize individual learners by utilizing achievement descriptors obtained through LRT. First, Shimizu and Nakano will concisely explain LRT, and then Yanakawa, Ohba and Sugino present several studies which attempted to apply LRT to analyzing the results of the grammaticality judgment tests of English verbs and *wh*-question and relative clause constructions as well as the results of the National Center Test for University Admissions.

II. Latent Rank Theory
(Yuko Shimizu & Michiko Nakano)
A. Problems in Educational Measurement

When we try to identify our learners' competence in educational settings such as entrance examinations, placement tests and some, but not all, tests in classroom settings, we often compare the test-takers' performances to the performances of the others. Those tests are designed to disperse the performances in a normal distribution for purposes of categorizing the test-takers into levels. Although observed scores are not true scores of the test-takers, the existence of measurement error is disregarded and the scores obtained are used to make an educational decision by aligning the test-takers on a continuous scale. Assume we decide 65 as a cutoff point in an English test. What does an observed score mean if a test-taker satisfies the cutoff point? Is s/he definitely different from the one who gets 64? How much are the test users sure about the discriminating power of the test? Compared to some physical measurement such as weight and height, which we have a shared understanding of their measurement tools and units of measurement, educational measurement do not have universally accepted and well-defined scales and units. Educational measurement is much harder to do consistently than measuring the heights or weights of the same people. In addition, items in educational measurement are based on assumptive constructs. It is an undeniable fact that the observed results have some limitation in interpreting.

B. The Latent Rank Theory

When academic abilities are measured and evaluated, we assume that the test-takers will be aligned along a straight continuous line. Kojiro Shojima and his research collaborators take more realistic approach to the test score

use and advocated a new test theory: Latent Rank Theory (LRT) (Shojima, 2008, 2009, 2010). LRT is to evaluate academic achievements using an ordinal scale, and ranks the test-takers into five to 20 levels or 'latent ranks.' The number of ranks is determined by data analysts, and the higher the rank a test taker is categorized, the higher his/her ability is. Such evaluation is seemingly rough and less accurate to make crude distinction among the test-takers. If we consider the problems in educational measurement and error components of the observed scores, however, judgment based on the ordinal scale of LRT is more feasible to describe each test-taker in a certain rank.

The analysis of LRT is done using *Exametrika* (Shojima, 2011). It produces various outputs to detect latent ranks of each test-taker as well as to provide information on each test item as follows. (Graphic information of Figures 1-5 was taken from Shojima's homepage.)

Item Reference Profile (IRP): IRP is an item-wise statistical information to provide a set of correct response ratios from test-takers at each latent rank. Figure 1 is an example of an IRP. Those who are ranked lower latent ranks are less successful and have a lower probability to give a right response to the given item than those in the upper latent ranks.
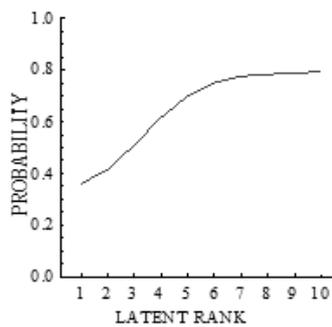

Figure 1. Example of Graphic Representation of IRP

Test Reference Profile (TRP): Each test item has its information of IRP. By combining them we will obtain another output: the Test Reference Profile (TRP). TRP shows an expected score for the test-takers at each latent rank. In Figure 2, the test-takers at a latent rank of 8 will get about 20 items right. Since LRT is based on the ordinal scale, TRP is expected to make a upward-sloping curve.
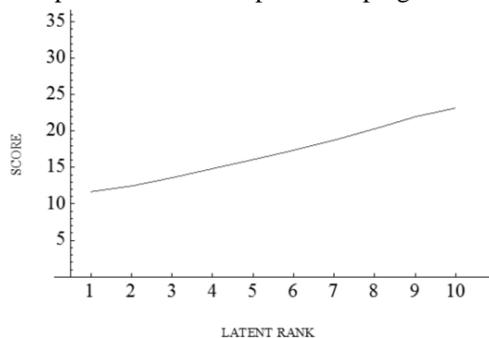

Figure 2. Example of Graphic Representation of TRP

Rank Membership Profile (RMP): Another useful output obtained from LRT is the Rank Membership Profile (RMP), which indicates each test-taker's membership probability for each latent rank. In the example of Figure 3-a, the probability that this test-taker belongs to Rank 10 is as high as .6 As to the example of 3-b, s/he will be at Rank 4 and the probability is around .3.
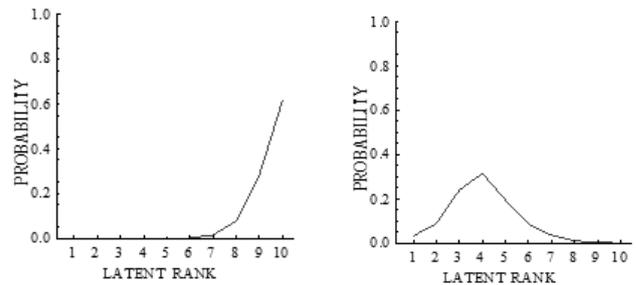

Figure 3-a. Example of RMP (1)   Figure 3-b. Example RMP (2)

Rank Membership Distribution (RMD): As TRP is the sum of IRPs, the Rank Membership Distribution (RMD) is the sum of RMPs. RMD shows the distribution of the test-takers and plots frequencies of them under each latent rank.
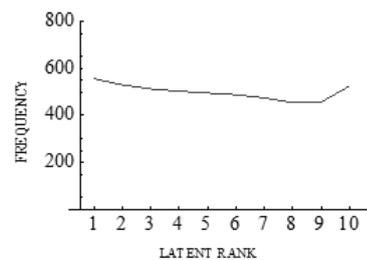

Figure 4. Example of Graphic Representation of RMD

Item Category Reference Profile (ICRP): As used in Yamakawa's and Ohba's studies below, graded polytomous data can be analyzed using an extended model of the LRT. Results of the Item Category Reference Profiles (ICRPs) indicate the probability of selecting each item category, where test-takers at higher latent ranks generally select higher categories.
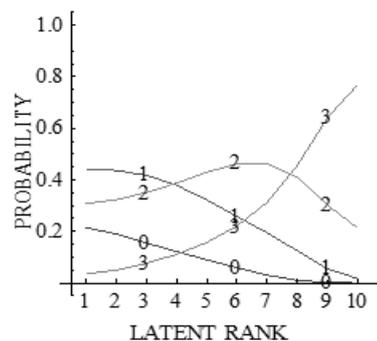

Figure 5. Example of Graphic Representation of ICRP

C. LRT and Can-Do Chart

An LRT analysis makes it easy to explain the relationship of test scores and abilities since each individual ability changes in stages on an ordinal scale. By closely analyzing

characteristics of items and response patterns which belong to each stage or each latent rank based on IRPs, we can develop a Can-Do Chart/Statement to explain abilities corresponding to each rank. In our study, we are experimenting to identify levels of grammatical proficiency and describe what Japanese learners of English can do in each developmental stages.
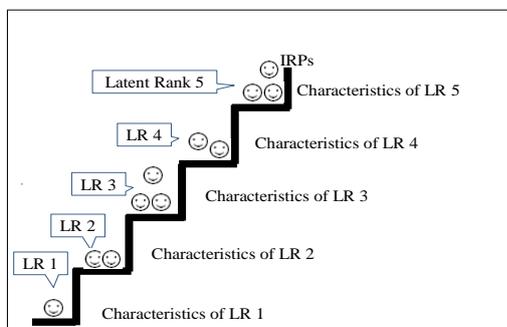

Figure 6. Ordinal Scale and LRT

III. Developmental Processes of English Verbs
(Kenichi Yamakawa)

This section attempts to illustrate the potentials of LRT in research on developmental processes of English verbs. We will approach this issue in two stages: an analysis of a single verb type (unaccusative / unergative verbs), and an analysis of multiple verb types (unaccusative / unergative verbs, datives, and psych-verbs).

A. Unaccusative / unergative verbs

In Yamakawa et al. (2010), we used the data of one of the grammaticality judgment tasks we had developed (see Yamakawa et al, 2008). The task aimed to explore the acquisition of English unaccusative and unergative verbs, and it was administered to 369 university-level Japanese EFL learners. The task was based on a format of grammaticality judgment task with a five-point (1 to 5) scale, consisting of 48 items including 12 filler items. Targeted 36 items contained six unaccusative and six unergative verbs, which were used in the NP+V, *the NP+be+p.p., and *the NP+V+NP constructions. Samples are provided below:

(1) Category A (unacc. in NP+V construction)
  Your letter arrived yesterday.
(2) Category B (unerg. in NP+V construction)
  Her father cried at her wedding ceremony.
(3) Category C (unacc. in *NP+be+p.p. construction)
  *Because of the rain, the train was arrived late.
(4) Category D (unerg. in *NP+be+p.p. construction)
  *He was cried when he heard of his mother's death.
(5) Category E (unacc. in *NP+V+NP)
  *Finally the waitress arrived the salad to us.
(6) Category F (unerg. in *NP+V+NP)
  *The boy hit his little sister and cried her.

In the process of analysis, we measured the "distance" of learners' judgments from the correct answers and converted their judgments to points (0 to 4) accordingly. For example, if a learner judged a grammatical sentence as "5", s/he was given four points, and if s/he judged it as "1", s/he was given no point.

As the results of the analysis using the statistical software, *Exametrika*, we found that the learners could be categorized into four latent ranks as in Table I below. Yamakawa et al. (2010) discussed the characteristics of the learners belonging to each latent rank, and compared the results with the findings of our prior research based on Classical Test Theory.

Table I. Item Reference Profiles (IRP)

Item No.	Cat.	Verb	Item Reference Profile (IRP)			
			Rank 1	Rank 2	Rank 3	Rank 4
			n=88	n=99	n=95	n=87
46	B	sing	2.991	3.365	3.672	3.822
42	B	cry	2.890	3.232	3.475	3.592
32	B	dance	2.830	3.203	3.380	3.517
06	A	die	2.949	3.086	3.299	3.460
23	B	laugh	2.817	3.123	3.367	3.383
04	E	arrive	2.813	3.013	3.217	3.282
36	A	happen	2.674	2.946	3.137	3.278
26	F	work	2.549	2.865	3.143	3.236
09	B	work	2.621	2.696	2.808	3.010
24	E	die	2.571	2.706	2.915	*2.931
03	B	play	2.405	2.554	2.888	3.169
16	A	appear	*2.539	2.607	2.775	3.048
18	D	sing	2.414	2.661	2.819	3.005
28	A	exist	2.324	*2.477	2.846	3.142
08	F	sing	*2.566	2.637	2.601	2.745
31	F	cry	2.416	2.525	2.688	2.907
20	A	arrive	2.421	2.613	2.805	2.751
13	A	fall	2.409	2.583	2.675	2.799
10	E	happen	2.345	2.524	2.755	2.818
17	F	dance	2.348	2.423	2.678	2.936
07	C	exist	2.358	*2.582	2.694	2.804
48	F	play	2.268	2.395	2.643	2.858
34	C	arrive	2.243	2.328	2.627	2.866
45	E	appear	2.310	*2.521	2.612	2.661
35	E	exist	2.126	2.380	2.594	2.841
29	D	play	2.125	2.293	2.627	2.874
14	D	cry	2.209	2.315	2.561	2.588
47	D	work	2.206	2.287	2.442	2.665
38	D	dance	2.152	2.214	2.367	2.581
39	F	laugh	2.270	2.347	2.316	2.300
12	D	laugh	1.937	1.982	2.073	2.171
02	C	appear	1.822	1.899	2.041	2.343
41	E	fall	1.933	1.938	1.873	2.085
43	C	die	1.790	1.780	1.916	2.116
27	C	fall	1.928	1.831	1.841	1.847
22	C	happen	1.736	1.623	1.685	1.848

0-2.0	2.0-2.5	2.5-3.0	3.0-4.0

As for the item/category features, the order of overall difficulty among all the categories was B < A < F < E < D < C, from easy to difficult. Also, the difficulty orders of the verbs were different according to the categories (i.e., the construction types). Next, as for the rank features, as learners progressed from Rank 1 to Rank 2, they obtained higher scores for Categories A and B. As they further proceeded from Rank 3 to Rank 4, their judgments became closer to perfect with regard to the two categories. Moreover, as learners progressed from Rank 2 to Rank 3, they gained higher scores for Categories E and F. As learners progressed from Rank 1 to Rank 4, the range for 2.0-2.5 became narrower, which could indicate learners made more correct judgments with confidence.

One interesting observation was that some learners in higher ranks seemed to regress on an acquisition path when they judged incorrect passive sentences in Categories C and D. This was reflected in ICRPs. ICRPs, which are also by-products of the analysis using LRT, express the probability of learners at each latent rank selecting each item category. In our particular case, ICRPs show what judgment learners made on each test item. Learners in higher latent ranks generally made a more correct judgment of a test item (i.e., they were more likely to select "definitely possible" for a grammatical item). However, some learners in higher ranks tended to make an incorrect choice ("definitely possible") about an ungrammatical item. This type of profile can be found with relatively more difficult items, especially items in Categories C and D.

B. Unaccusative / unergative verbs, datives, and psych-verbs

We also combined the data of unaccusative / unergative verbs with the data that had been obtained from two other grammaticality judgment tasks targeting dative alternation and psych-verbs. We finally obtained the data of 741 participants altogether.

The task targeting dative alternation has six categories, each consisting of eight items, four to-datives and four for-datives, totaling 48 items. Samples of the task are provided below:

(7) Category A (well-formed ditransitives)
 Mr. Jones gave me some money. (to-dative)
 John found me a new dress. (for-dative)
(8) Category B (ill-formed ditransitives)
 *Mr. Jones reported me the accident. (to-dative)
 *King Arthur fought the queen the monster. (fo-dative)
(9) Category C (well-formed prepositional datives)
 Mr. Jones gave some money to me. (to-dative)
 John found a new dress for me. (for-dative)
(10) Category D (well-formed prepositional datives)
 Mr. Jones reported the accident to me. (to-dative)
 King Arthur fought the monster for the queen. (for-dative)
(11) Category E (well-formed passive sentences)
 I was given some money by Mr. Jones. (to-dative)
 I was found a new dress by John. (for-dative)
(12) Category F (ill-formed passive sentences)

 *Some money was given me by Mr. Jones. (to-dative)
 *The new dress was found me by John. (for-dative)

English psych verbs are categorized into two classes: ES verbs and EO verbs. ES verbs take an Experiencer NP as the subject and a Theme NP as the object. This alignment is in accordance with the thematic hierarchy and uniform assignment of specific theta-roles. EO verbs, on the other hand, take an Experiencer NP as its object, a Theme NP filling the subject position. Therefore, EO verbs are generally recognized as more problematic for ESL/EFL learners. The task targeting psych-verbs has 48 items in total, with 10 types of psych-verb each consisting of four items, plus 8 filler items. Samples of the task are provided below:

(13) ES1 (Human Experiencer/ES verbs/Human Theme)
 e.g. Most people hate Harry but they don't dare to say so.
(14) ES2 (Human Experiencer/ES verbs/Inanimate Theme)
 e.g. They enjoy anything that breaks the dullness of their routine life.
(15) ES3 (Inanimate Theme/ES verbs-passive/PP with Human Experiencer)
 e.g. His beautiful house is envied by the neighbors.
(16) ES4 (*Inanimate Theme/ES verbs/Human Experiencer)
 e.g. *Your skills have fully appreciated our boss.
(17) ES5 (*Human Experiencer/ES verbs-passive/PP with Inanimate Theme)
 e.g. *She will not be enjoyed by parties or dining out.
(18) EO1 (Human Theme/EO verbs/Human Experiencer)
 e.g. A handsome man bored me with stories of the Navy.
(19) EO2 (Inanimate Theme/EO verbs/Human Experiencer)
 e.g. The sharp tone of her voice amazed her boss.
(20) EO3 (Human Experiencer/EO verbs-passive/PP with Inanimate Theme)
 e.g. We all were bored with his lecture.
(21) EO4 (*Human Experiencer/EO verbs/Inanimate Theme)
 e.g. *After the competition, I disappointed the results..
(22) EO5 (*Inanimate Theme/EO verbs-passive/PP with Human Experiencer)
 e.g. *Their performance was impressed by the spectators.

The analysis using LRT yielded the results indicating that the learners were best categorized into four latent ranks. It also enabled us to illustrate the characteristics of each latent rank as well as to compare the levels of difficulty of the test items from across the three tasks targeting different grammatical features. As for the rank features, there were less test items than expected that could characterize different ranks; on the other hand, however, as for the item features, it seemed that the difficulty levels of the test items were determined by two elements: construction-based and verb-based elements.

IV. Developmental Processes of *Wh*-question and Relative Clauses (Hiromasa Ohba)

In this section, we analyze the data obtained from grammaticality judgment tests on how adult Japanese EFL learners interpret English relative clause and *wh*-question

constructions, using LRT. By using this data analysis procedure, we could expect to elucidate a real picture of their developmental stages of English relative clause and *wh*-question constructions.

Participants were 740 university-level Japanese learners of English who were from a variety of departments. They could be regarded as at least low-intermediate level students as they were required to pass the entrance examinations of the universities. They were given untimed grammaticality judgment tests on *wh*-question and relative clause constructions. These tests consisted of 16 grammatical and 18 ungrammatical relative clauses, and 8 grammatical and 18 ungrammatical *wh*-questions with some fillers. In addition, there were five different types of violations on *wh*-movement in ungrammatical sentences (i.e., extraction out of complex NP, subject island, relative clause, *wh*-island, and adjunct island). Sample sentences of the tasks are supplied as follows:

(1) Grammatical relative clauses (*wh*-operator, complementizer *that*, null operator（or complementizer)) and ungrammatical ones (*who(m) that / which that* (doubly-filled complementizer), resumptive pronoun):
The boy who(m) I kicked yesterday broke the window.
The picture that you are looking at was painted by Picasso.
The friend they lent money to bought a very big house.
*The woman who that is singing on the stage is my wife.
*The classmate that you don't like him is very unkind.
(2) Grammatical *wh*-questions and ungrammatical ones (no subject-auxiliary inversion):
What did your girlfriend want to talk about?
*Whose house Sandy's father is going to build?
(3) Ungrammatical relative clause and *wh*-question sentences violating the Subjacency condition in the following 5 construction types:
 (a) Extraction from complex NP
 *This is the car which we heard the news that Toru bought.
 *Which car did he believe the claim that John stole?
 (b) Extraction from embedded question, *wh*-island
 *This is the CD which Peter knows where Tom bought.
 *Which book did she ask John when he read?
 (c) Extraction from relative clause
 *This is the bicycle which the police caught the man who stole.
 *What did he interview the teacher who wrote?
 (d) Extraction from sentential subject
 *This is the ghost which a picture of frightened the children.
 *What did a discussion of occur during the meeting?
 (e) Extraction from adjunct (adjunct island)
 *This is the homework which Ann went to school before she did.
 *Who did the earthquake occur while you were talking with?

Test items were randomized, and the participants were asked to read the test sentences and rate their grammaticality on a 5-point scale indicated. They were told that 5 meant that the sentence was 'completely grammatical', 1 meant that it was 'completely ungrammatical', and 4, 3 and 2 were gradations between the extremes to be used if they thought the sentence was more or less grammatical. We also converted their judgments to points 0 - 4 accordingly.

In the analysis using *Exametrika*, considering the number of items and the sample size, the number of latent ranks was set in 10. Results are given in Table II

Table II.   Item specifications at each latent rank

Rank 1	**RC (null)**: The house you can see on the corner was built ten years ago. **RC (*wh*-operator)**: The job which I wanted to apply for was very popular. **WH (question)**: What did you do for your son last week? **WH (question)**: What did your mother want to talk about? **WH (question)**: When did he go to London to study English? ***WH (sub-aux inversion)**: *Who your favorite movie stars are? ***WH (sentential subject)**: *Who(m) did that she went out with make him sad? ***WH (relative clause)**: *What did they visit a shop which sold?
Rank 2	**RC (that)**: The picture that he is looking at was painted by Picasso. **RC (*wh*-operator)**: The woman who helped me with my homework is Hanae. **RC (*wh*-operator)**: The boy who(m) I met yesterday broke the car. **WH (question)**: Why was the man surprised to read a letter? ***WH (sub-aux inversion)**: *Where your friend got such a great idea? ***RC (relative clause)**: *This is the soup which Mary visited a restaurant which served.
Rank 3	**RC (*wh*-operator)**: The man who(m) I employed as my assistant works hard. ***WH (relative clause)**: *What did he interview the teacher who wrote?
Rank 4	**RC (that)**: The car that you can see over there caused this accident. **RC (null)**: The magazine they are always talking about is very useful.
Rank 5	**RC (that)**: The student that has written this letter must be crazy. ***RC (resumptive pro)**: *The classmate that you don't like him is very unkind. ***RC (doubly-filled comp)**: *The woman who that is singing on the stage is my aunt.
Rank 6	***RC (resumptive pro)**: *The town that my mother came from it is far from here ***RC (resumptive pro)**: *The friend that I lent the book to her studied very hard. ***RC (sentential subject)**: *This is the meeting which that Taro attended shocked his parents. ***WH (sub-aux inversion)**: *Which city you believe that they attacked? ***WH (sub-aux inversion)**: *Whose car your father is going to drive? ***WH (sub-aux inversion)**: *Which medal it is difficult for him to win?
Rank 7	***RC (sentential subject)**: *This is the ghost which a picture of frightened the children. ***RC (adjunct island)**: *This is the homework which Ann went to school before she did.
Rank 8	**RC (*wh*-operator)**: The girl whose handbag was

	stolen is suffering from shock.
	RC (wh-operator): The boy to whom I talked yesterday seemed very nervous.
	RC (wh-operator): The girl for whom I have bought a computer is my sister.
	*RC (doubly-filled comp): *The woman whom that we talked with was our teacher.
	*RC (doubly-filled comp): *The dogs which that I gave the milk to were very small.
	*WH (sub-aux inversion): *Why she was worried about her daughter?
Rank 9	***RC (doubly-filled comp)**: *The glasses which that Judy broke were very expensive.
	*RC (adjunct island): *This is the girl who(m) the bell rang while I was thinking of.
Rank 10	***WH (sentential subject)**: *What did a discussion of occur during the meeting?
	RC (that): The student that you gave a present to looked very happy.
	RC (null): The friend they lent money to bought a very big house.
	RC (wh-operator): The girl from whom I received a letter is pretty.
	RC (wh-operator): The man used a word whose meaning I don't know at all.
	WH (question): Whose books did you borrow yesterday?
	WH (question): Who came to see your father yesterday?
	WH (question): Who(m) does she know that Mary loved?
	WH (question): What books is it necessary for you to read?
	*RC (resumptive pro): *The building that it stands near the lake is our school.
	*WH (sub-aux inversion): *What you and your son looked at?
	*WH (sub-aux inversion): *What you and your friend wanted to do?
	*RC (relative clause): *This is the book which John interviewed the man who criticized
	*WH (adjunct island): *Which car did they cross the street when John stopped?
	*WH (adjunct island): *Who(m) did a fire occur while you were talking with?
	*RC (complex NP): *This is the boy who(m) Jack described the way that Bill hit.
	*RC (complex NP): *This is the car which we heard the news that Toru bought.
	*WH (complex NP): *Who(m) did they know the fact that David hit?
	*WH (complex NP): *Which car did he believe the claim that John stole?
	*RC (embedded question): *This is the house which Peter knows when Tom bought.
	*RC (embedded question): *This is the story which Kyoko wondered who believed.
	*WH (embedded question): *Which book did she ask John when he read?
	*WH (embedded question): *What did you wonder who would believe?

Results showed that the learners were categorized into 10 latent ranks, sorted by the average item scores of 2.5 or

above, and a learner at a higher rank judged correctly in more test items. However, there were several items in which, as the learners' ranks went up, parallel increase was observed in the probabilities of both correct and incorrect responses. This implies that even for advanced Japanese learners of English (i.e., learners at Rank 10), acquisition of some types of relative clause and *wh*-question constructions poses unique difficulty. Moreover, advanced Japanese learners of English did not seem to reach the native speakers' level in judging sentences violating some subjacency constraints as ungrammatical.

V. The National Center Test for University Admissions
(Naoki Sugino)

Reported in this section are the results of LRT-based analysis on the actual test-takers' performance in the National Center Test for University Admissions (NCT, henceforth), developed and administered by National Center for University Entrance Examinations (NCUEE, henceforth) in 1990, 1997, and 2004. These three test administrations were chosen in order to investigate if the revision of the *Course of Study* in 1989 had any effect on the contents of the NCTs, the knowledge and skills measured, and test-takers' overall English language proficiency.

Prior to the analysis, items in each year's test set were described in terms of what knowledge or skills the item requires of test-takers to answer correctly. Although there are differences in formats and weightings, the aspects of knowledge/skills being measured in the NCTs are roughly the same, as are shown in Table III.

40,000 cases are randomly sampled from each year's "first-time" test-takers and are submitted to the LRT-based analysis. Test-takers are divided into 10 latent ranks, and the test items, sorted by the overall correct response ratios, with IRP of .7 or higher by the learners at a specific latent rank are regarded as items within their reach.

Table III. Aspects of knowledge/skills and the number of items

Test Administration Aspects	1990	1997	2004
Reading skill	19	17	17
Vocabulary	16	11	7
Grammar	5	7	6
Phonology	7	6	6
Function	6	3	4
Discourse	0	3	7

Table IV shows the ranks of the 1990 NCT test-takers and the specifications of the items learners at a specific rank can cope with. Note that learners at a higher rank can handle those items that are below their rank with better IRPs.

Table IV. Items with IRP of .7 or above at each latent rank: 1990

Rank	Item Number and Item Specification	
1	No items reached IRPs of .700.	
2	40	matching understanding of the visual prompt and the text
	34-5	*mind if S V*

	5	word accent
	7	sentence stress in a dialogue
3	42	fact finding in an expository test
	25	requesting additional information
	23	asking/sympathizing
	39	matching understanding of the visual prompt and the text
	6	sentence stress in a dialogue
	41	matching understanding of the visual prompt and the text
	38	matching understanding of the visual prompt and the text
4	30-1	*afford+to*-inf., *of one's own*
	45	fact finding in an expository test
5	49	fact finding in a narrative
	44	fact finding in an expository test
	43	fact finding in an expository test
	46	fact finding in a narrative
	51	fact finding in a narrative
	52	fact finding in a narrative
6	48	fact finding in a narrative
	53	fact finding in a narrative + paraphrase
	56	fact finding in a narrative + paraphrase
	55	fact finding in a narrative + paraphrase
7	36-7	*think of* + relative pronoun
	19	*afraid / doubt / think / wonder*
	4	word accent
	11	*gas / money / order / traffic*
	1	word accent
	15	perfective x continuous x past/present
8	50	fact finding in a narrative
	13	*swim*
	3	word accent
	16	*gave up / managed / put off / succeeded*
	10	*advised / said / suggested / warned*
	14	*Because of/ Besides/ Despite/ Instead of*
	32-3	*look down on*, relative pronoun, *differ from*
9	28-9	*provide* + O + *with* …
	47	fact finding in a narrative
	22	agreeing
	21	comparative, word order
	54	fact finding in a narrative + paraphrase
	12	*message / photo / picture / scene*
10	9	*is weigh / is weight / weighs / weights*
	2	word accent
	24	adding new but relevant information
	8	*leave* + O + Vpp
	18	*grow up / raise up*
	27	changing suggested time arrangement
	26	requesting clarification
	20	*discount*
	17	*participated / played in / represented / took part in*

As can be seen in Table IV, what differentiates test-takers at Rank 7 (or above) from those at Rank 6 (or below) is their fine-tuned knowledge in vocabulary items and grammatical features. This tendency, however, is not observed in the 1997 and 2004 administrations, as are displayed in Tables V and VI.

Table V. Items with IRP of .7 or above at each latent rank: 1997

Rank	Item Number and Item Specification	
1	9	reproving/scolding
	5	sentence stress
	40	matching information in a dialogue with the visual prompt
2	3	pronunciation; consonant /z/
3	39	matching information in a dialogue with the visual prompt
	42	making inference + paraphrasing
	14	relative adverbs
	10	subjunctive
	8	requesting/inducing opinion
	34	sentence completion with fact-finding
4	43	fact finding + paraphrasing
	35	sentence completion based on the figures in the table
	46	making inference
	2	word accent
	33	matching information in the text to the figures in the table
	19	*discovering / finding out / looking for / searching*
5	41	identifying referent across several turns in a dialogue
	31	coherent sequencing of sentences in one paragraph
	7	offering
	47	making inference + paraphrasing
6	45	fact finding in a paragraph
	22	*importance / meaning / sense / truth*
	1	word accent
	13	perfective x continuous x past/present
7	48	making inference
	21	*as much/ as possibly/ as quickly/ as well*
	16	*a chance/ a change/ the lead/ the place*
	17	*have / know / make / turn + better*
	6	sentence stress
8	38	*amazing / easy / scary*
	28-9	relative pronoun, word order
	20	*large / narrow / small / wide*
	49	making inference
	37	fact finding
	50	fact finding + paraphrasing
	32	coherent sequencing of sentences in one paragraph
	18	*for / from / of / with*
9	44	fact finding
	24-5	*leave* + O + *back to* …
	30	logical connectors
	12	*almost / any / most / none*
	23	*command / direct / order / persuade*
10	11	comparative

	36	sentence completion based on the figures in the table
	15	*far* / *long* / *many* / *much*
	4	pronunciation; vowel /i:/
	26-7	*however* + adverb; word order

Table VI. Items with IRP of .7 or above at each latent rank: 2004

Rank	Item Number and Item Specification	
1	No items reached IRPs of .700.	
2	No items reached IRPs of .700.	
3	10	*mind* + O + Vpp
	5	sentence stress
	33	matching information in an expository text to visual prompt
	19	making a promise
	39	matching understanding of a dialogue to visual prompt
4	45	making inference of a character's feeling
	17	responding to a negative prospect
5	7	*make sense*
	14	subjunctive
	3	sentence stress
	4	sentence stress
	43	making inference of a character's feeling
	18	responding to an unexpected schedule change
6	12	*unless* / *if* / *whether* / *since*
	48	fact finding
	36	making inference based on the visual prompt
7	41	overall understanding of a dialogue
	34	fact finding
	15	*before any- / every- / no- / some- thing*
	44	making inference of a character's feeling
8	42	overall understanding of a dialogue
	27	logical connector: *therefore*
	20-1	relative pronoun; word order
	30	coherent sequencing of sentences across paragraphs
	38	back channeling markers
	13	*mind* Ving
	31	coherent sequencing of sentences across paragraphs
	47	making inference of a character's feeling
9	46	making inference of a character's feeling
	29	coherent sequencing of sentences in one paragraph
	32	coherent sequencing of sentences across paragraphs
	35	fact finding
	22-3	*used to*, *with the light on*, word order
	26	logical connector: *as a matter of fact*
	2	accent patterns; *comments* and *particular*
	49	fact finding
	16	tense and aspect, psych verb
	50	fact finding
	11	past perfective
	6	sentence stress
10	37	making inference in an expository text
	28	coherent sequencing of sentences in one paragraph
	1	accent patterns; *network* and *fascinating*
	40	identifying referent
	8	*devote* + O + to+ Ving
	24-5	*what it is that is making the noise*
	9	*common* / *familiar* / *popular* / *broad*

VI. Conclusion

By applying LRT to analyzing the results of grammaticality judgment tests of English verbs, and *wh*-question and relative clause constructions as well as the results of the NCTs, we have demonstrated how Japanese learners of English can be characterized in terms of what they can do at each latent rank. Thus, we believe in the high potential of LRT in SLA research by explicating L2 developmental processes of grammatical features with knowledge descriptors gained through it..

References
Crocer, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadworth Group.

Shojima, K. (2008). Neural test theory: A latent rank theory for analyzing test data (DNC Research Note, 08-01). Retrieved from http://www.rd.dnc.ac.jp/~shojima/ntt/Shojima2008RN08-01.pdf

Shojima, K. (2009). Neural test theory. In K. Shigemasu, A. Okada, T. Imaizumi, & T. Hoshino (Eds.), *New trends in psychometrics* (pp. 417-426). Tokyo: University Academic Press.

Shojima, K. (2010). Neural Test Theory. In M. Ueno & K. Shojima, (Eds.), *Gakushu Hyouka no Shin-chouryuu* [*New trends in evaluation of learning*] (pp. 83–111). Tokyo: Asakura Shoten.

Shojima, K. (2011). *Exametrika* 5.3 [Computer software]. Retrieved from http://www.rd.dnc.ac.jp/~shojima/exmk/index.htm

Yamakawa, K., Sugino, N., Ohba, H., Nakano, M., & Shimizu, Y. (2008). Acquisition of English grammatical features by adult Japanese EFL learners: The application of Item Response Theory in SLA research. *Electronic Journal of Foreign Language Teaching (e-FLT), 5*(1), 13–40. Retrieved from http://e-flt.nus.edu.sg/

Yamakawa, K., Sugino, N., Shimizu, Y., Ohba, H., & Nakano, M. (2010, March). *The possibility of the application of Neural Test Theory to SLA Research*. Poster presented at the meeting of American Association for Applied Linguistics, Sheraton Atlanta Hotel, Atlanta, GA.