
JLTA Newsletter

日本語テスト学会

The Japan Language Testing Association

JLTA Newsletter No. 20 発行代表者：大友 賢二 2004年（平成16年）1月25日発行
発行所：日本語テスト学会（JLTA）事務局
〒389-0813 長野県千曲市若宮 758 TEL 026-275-1964 FAX 026-275-1970
e-mail: youichi@avis.ne.jp URL: <http://www.avis.ne.jp/~youichi/JLTA.html>



もっとアジア諸国に目を向けた JLTA に期待して

木下 正義(福岡国際大学)

10月25日(土)に JLTA 第7回全国大会熊本大会は参加者70名を超し、懇親会参加者も40名と大盛會裡に終了した。韓国からの Guest Speaker の Dr. Oryang Kwon (Seoul National University) の基調講演 “Interactive Impacts of English Testing, Teaching, and Learning in Korea” も大会を盛り上げてくれた。JLTA 大会終了後の2週間後の11月7日から9日まで小生の旧友の Dr. Hyo-Woong Lee (Korea Maritime University) が Founding President を務めた The 1st Asia TEFL International Conference 参加のために4名の同僚と韓国・プサンを訪れた。大会にはアジア諸国及び欧米から約20カ国から約1,000名の学者・英語教師の参加があり大盛會であった。会場の Nong-Shim Hotel を貸切の大会に韓国の政界・財界及び教育器機メーカーからの寄付金約1,200万円を集めた大会であった。来年度は韓国・ソウル大会が予定されているが大会実行委員長の Dr. Oryang Kwon にとっては二回目の大会は大変なご苦労であろう。

大会2日目に同僚2名とプサンから約80Km離れた韓国の京都と称される慶州を訪問した。訪問目的は Dr. Larry D. Chong (Gyeongju University) の面会と慶州大学の CALL 施設の見学であった。Dr. Larry D. Chong は12月3日から5日まで Bangkok で開催される Asia CALL International Conference の Chair である。Dr. Chong は大会参加者は約120名位だが、参加者全員がその道の専門家であり、大会会場も Bangkok の大学を無償で貸してもらい、宿泊も二食付きで3,000円と格安の Hotel を契約したとのことであった。2つの学会を比較して言えることは Asia TEFL は今後大会を受け入れる国は財政的な基盤に不安が残るが、一方 Asia CALL は地味な大会運営と専門家集団の研究會である点で JLTA に酷似している。

21世紀はアジアの時代と良く囃される。数年前に韓国・プサン国立大学校を訪問の折、Dr. Jung Hun Ahn (Director of Language Education Center, PNU) が Asia Language Testing Association (以下 ALTA) を立ち上げたら如何かと小生に提案されたことを思い出した。彼の ALTA 設立の理由は1) Testing の研究が余りにも欧米中心的で影響力が強すぎる 2) TOEFL や TOEIC 等の Testing が ETS の独占市場になり、アジア諸国の Testing 専門家がアジア諸国の風俗・文化や習慣等の内容を駆使した題材の問題作成とテスト実施が出来ないだろうかと2点を力説された。アジア諸国の Language Testing の専門家が一同に集まり、各国の Language Testing の現状と展望や各国の高校・大学の入学試験問題の比較分析等を討議・議論することも興味深く今後期待される項目ではあるまいか。

小生が代表を務める東アジア英語教育研究会(会員数:35名)では東アジアの英語教育に関する研究会を今年10月で40回の研究会を開催したことになる。韓国・台湾・中国及びマレーシア・ブルネイ等のテスト研究・教科書分析や言語政策に至る研究をしてきた。1999年度のLTRC大会後のAILA東京大会では日本・韓国・中国三カ国の大学教授による各国の Language Testing に関するシンポジウムを持った。今後のJLTAの新企画事業の中にアジア諸国の Language Testing について Research Project を提案し、将来いつか ALTA International Conference が開催されることを期待したい。

日本語テスト学会
第7回全国研究大会

第7回日本語テスト学会全国研究大会が、2003年10月25日(土)に、熊本大学(黒髪キャンパス)で開催された。以下その報告である。

日本語テスト学会
第7回全国研究大会を開催して

大会運営委員長 島谷 浩(熊本大学)

2003年10月25日(土)、秋晴れの下、熊本大学大学教育センター棟において、日本語テスト学会第7回全国研究大会が開催されました。



本学会の全国大会は、1999年7月につくば国際会議場で第21回言語テスト国際会議が開催された際に、第3回大会も同会場で開催されたのを除き、本学会理事の中村優治先生が所属されておられる東京経済大学にて開催されました。今回初めて、首都圏を離

れ、九州熊本で開催されることになったのは、昨年の研究会運営委員会において、全国研究大会については、委員長の所属校またはその近辺で開催するという内規が決まったからです。私は、本年4月より新委員長となり、大会テーマを委員間の投票により「言語学習とテストング: テスティングが言語学習に果たす役割」と決め、準備を進めてまいりました。

はじめての地方での大会開催ということで、研究発表がどれだけ集まるかという点が一番心配でしたが、20件の発表申込みがありました。2件のキャンセルがあり、最終的な発表件数は18件となりましたが、この発表件数は、本学会としては過去最多となりました。研究発表は、5室に分かれて午前中10件、午後8件行われ、18件中10件が英語を使用した発表でした。

基調講演は、韓国ソウル大学の Oryang Kwon 先生にお願いしました。韓国英語教育学会(KATE)と韓国英語テスト学会(KELTRA)の会長を務められておられるKwon先生の講演は、過去10年間の英語教育改革、特に大学修学能力試験(CSAT)における大きな変化とその波及効果が中心で、日本での英語教育改革に大変示唆あふれる内容でした。

シンポジウムは、中村優治先生(東京経済大学)にコーディネーターをお願いし、法月健先生(静岡産業大学)、Garold Murray先生(東海

大学)、渡部良典先生(秋田大学)をパネリストとしてお招きしました。使用言語は全て英語で、テーマの“The Role of Testing in Language Learning”について、各先生よりご発表いただき、その後、熱心な質疑応答がなされました。

昼食後の研究会運営委員会において、全国大会開催地に関する修正案が承認され、今後は、委員長の所属機関またはその近辺での開催と首都圏の大学等での開催とを交互にすることが決まりました。そして来年度の全国大会は、大坪一夫理事と片桐一彦事務局次長が所属されておられる麗澤大学で開催されることが決まり、その後、総会で了承されました。

熱心な発表、討論の後の懇親会は、学内のくすの木会館において行われ、熊本特産の海の幸、山の幸をご賞味いただきました。懇親会参加者が40名を越えたため、この学会で恒例になっていた参加者全員が輪になって自己紹介する時間が持てないほど、にぎやかで、有意義な歓談の時間を持つことができました。

今回の大会を運営して、本学会が確実に進化していることを実感いたしました。そして、会員52名、当日会員18名の参加があり、予想をはるかに越えて盛り上がった大会を運営できて、大変嬉しく思っております。いろいろと準備不足で、ご迷惑をおかけした点が多々あったことと思いますが、大友賢二会長をはじめ、多くの先生方から、暖かい感謝とねぎらいのお言葉をいただきました。

最後になりましたが、この大会のために、さまざまな形でご協力いただいた皆様方に心より御礼申し上げます。また、来年の第8回全国研究大会を、さらに発展した大会としたいと思っております。今後とも、ご協力の程、どうぞ宜しくお願いいたします。

Dr. Kwon Oryang's Plenary Lecture



Dr. Kwon Oryang, professor of English education at Seoul National University, gave a plenary lecture entitled “Interactive Impacts of English Testing, Teaching, and Learning” at the 7th National JLTA Conference, which was held at Kumamoto University on October 25, 2003.

The lecture focused major reforms in English education in Korea since 1991. The reforms are: (1) new public teacher employment test (1991), (2) the 6th National Curriculum (1992), (3) the College Scholastic Ability Test (1993), (4) reforms in university English teaching (from the early 1990s), (5) reforms in teacher training and retraining (1990s), (6) importing of native-speaker teachers (1996), (7) English as a subject in elementary schools (1997), and (8) the 7th National Curriculum. Among these, Dr. Kwon focused on the 6th National Curriculum, the College Scholastic Ability Test (CSAT), and the 7th National Curriculum. He also discussed the recent adoption of performance assessment (PA) in elementary and secondary school.

The 6th National Curriculum, which was announced in 1992, but went into effect for middle school in 1995 and for high school in 1996. The 6th National Curriculum was the first in Korean history to emphasize communicative competence and to follow a functional syllabus. These changes caused a sea change in English education that facilitated the implementation of more controversial reforms later.

To reinforce the goals of the 6th National Curriculum, the CSAT, a new university entrance examination began for the 1994 academic year. The CSAT was the first university entrance examination in Korean history to include a listening component. The test emphasized fluency over accuracy and included discourse-based questions and context-embedded questions. The adoption of a listening component, in particular, caused immediate changes in classroom teaching.

The 7th National Curriculum, which was announced in 1997, built on the positive effects of the above reforms. The 7th National Curriculum became effective for elementary and middle school in 2001 and for high school in 2003. The 7th National Curriculum included two reforms: level-sensitive classes and performance assessment (PA) and integrative testing. Related to the new curriculum, the Ministry of Education and Human Resources promoted teaching English through English (TETE) in the classroom.

Performance assessment (PA) remains one of the most controversial of the recent reforms. To prepare for the introduction of the 7th National Curriculum, the Ministry of Education and Human Resources began to encourage PA in 1998. PA is designed to test the development of communicative competence through qualitative description in elementary school and a

combination of integrative testing and discrete-point testing in middle school and high school.

At the end of this lecture, Dr. Kwon reported on results of a survey of 232 secondary school teachers on the washback of the CSAT on classroom teaching and the adoption of PA in the classroom. According to the survey, 60% of the teachers reported that the CSAT had influenced their classroom teaching. By skill area, teachers reported that the test had the greatest effect on their teaching of reading, followed by listening, grammar, and speaking, with writing being affected least. Results showed that the CSAT had more influence on high school teachers, but that high school teachers were more critical of the test's adequacy as an assessment tool.

Regarding PA, results of the survey showed that teachers were generally positive to the idea of PA and that most teachers were familiar with PA. Teachers reported that PA was difficult to use in assessing large classes. There was little difference in attitudes to PA according to number of years of teaching experience, which suggests that experienced teachers are willing to try new approaches in teaching and assessment.

Dr. Kwon concluded his lecture by reviewing how simultaneous and multi-faceted reforms worked together in reforming English education in Korea. In particular, he noted that public support for and acceptance of single reforms raises public support for other reforms, which is critical to success of the multi-faceted approach to reform.

(Reported by Robert J. Fouser
Kyoto University)

第7回全国研究大会 口頭発表内容の報告

Can-do 項目作成のための英語能力評価共通 枠組みの検討

長沼 君主(清泉女子大学)
和田 朋子(東京外国語大学大学院)
田中 敦英(桐朋中学・高等学校)
根岸 雅史(東京外国語大学)

上記4名の方々による共同研究で、当日の発表は長沼、和田の両氏によるものであった。近年、項目応答理論(IRT)を用いた絶対尺度による評価が注目を集めているが、これらのスコアはあくまでも「ものさし」であって、被験者が英語を使って何ができて、何ができないかを保障しているわけではない。本発表は、被験者の能力をテストスコアではなく、何ができるかといった視点から客観

的に自己評価させる can-do 項目に着目し、既存のフレームワークを概観しながら、能力評価の共通の枠組みを作成し、言語能力の発達段階を記述することを試みたものであった。その際参考にした資料は、主に Common European Framework、UK National Curriculum、ACTFL Proficiency Guidelines、Canadian Language Benchmarks、それに英検であったが、能力評価の観点、それぞれの評価基準が作成されたバックグラウンドにかなり影響されているようである。TOEIC や英語コミュニケーション能力テストなどでも can-do 項目を用いた分析をもとに被験者にフィードバックするコメントが作成されているが、今後は今回作成したような共通枠組みを利用して、被験者の能力評価に適した観点を選択していく必要があると結論づけている。

(報告:大津敦史 福岡大学)

日本人英語学習者の英語メンタルレキシコン — 感情を表す形容詞の定着度

折田 充(八代工業高等専門学校)

本発表は、日本人英語学習者のメンタルレキシコン形成における、感情を表す形容詞4種類(HAPPY、AFRAID、SAD、ANGRY という4つのカテゴリー)30個の定着度を Haastrup and Henriksen (2000)で用いられた3つのタスク(Primary Sorting Task=課題1、Card-Sorting Task=課題2、Situation Task=課題3)を用いて調査・分析したものである。調査の対象となったのは、九州内の大学工学部の1年生30名であった。結論としては、以下の3つがあげられている:

- 1) 上記4つのカテゴリーにおいて、HAPPY、AFRAID、SAD、ANGRY の順に定着度が下がる。
- 2) 課題3は英単語のネットワーク構造化の程度以外のもの(リーディング)を tap している可能性がある。
- 3) 課題の結果を左右する大きな要因は、単語の出現頻度、接辞情報、カタカナ語の有無、多義語などであるが、その関連性の度合いは今後の研究課題である。

(報告:大津敦史 福岡大学)

A Possibility of Norm-Referenced Testing for Educational Purposes in Japan—A TOEIC Case Study

Mizumoto Atsushi (*Graduate School of Language Education and Information Science, Ritsumeikan University*)

In Mizumoto's study, he looked at whether TOEIC and the testing preparation course used for the test can be used for educational purposes

in school. His study is based on the assumption that a norm-referenced test like TOEIC can be regarded as a criterion-referenced test. Norm-referenced tests in language learning are designed to measure the learner's general language ability. Mizumoto pointed out that a TOEIC test is a norm-referenced test. On the other hand, he pointed out that a criterion-referenced test, which has been traditionally used in schools, tests a particular body of knowledge i.e. an achievement test. In Mizumoto's presentation, he explored, "the possibility of regarding a norm-referenced test, the TOEIC, as a criterion-referenced test..." First, Mizumoto collected data from a questionnaire to determine the test takers knowledge of norm-referenced tests and criterion-referenced tests they took. In his analysis, he reported that, "Test takers revealed their unawareness toward the differences among the tests administered to them". Next, students were given a 12-hour TOEIC preparation course. According to Mizumoto, the course focused on questions that would most likely be asked. Mizumoto concluded that, "The scores of the students, with the idea that the TOEIC can be studied for, after the course showed clear score gains compared with the those before treatment". Mizumoto concluded that the results of his study suggest that conventional tests practices used by TOEIC can also be introduced into an English curriculum for educational purposes. A comment from a member of the audience questioned whether a test designed to be a norm-referenced test could actually be regarded as criterion-referenced test. Mizumoto considered this comment.

(Reported by Terry Laskowski
Kumamoto University)

Enhancing the Learning, Teaching and Testing of New Vocabulary by Systematically Using a Depth of Lexical Processing Scale and a Vocabulary Learning Strategies Taxonomy

John Paul Loucky (*Seinan Women's University*)

In the past few years, we have all watched the use of electronic dictionaries expand in our classes as now a majority of students have access to them. At the same time we also have seen how students incorrectly use an accessed word, for example, where they opt for the noun

usage of a word when the adjective or adverb form should have been used. Moreover, we have seen awkward semantic uses of words e.g. *my little brother molests me everyday* instead of more contextually appropriate verb choices such as *bothers*, or *annoys*. How do we monitor these unproductive uses of the dictionaries? Gratefully, in his presentation, Loucky has addressed problems that can occur when students access dictionaries. He believes that as better software, advanced electronic pocket bilingual dictionaries and Internet mediated machine translation become more available, problems for learners will intensify if left unchecked. In this environment, Loucky posits that "...maximising [students] retrieving, recording, rehearsal and recycling functions for most rapid TL access and acquisition now seem to be far more important skills for low proficiency language learners to develop...". Loucky offered an in depth model framed around a taxonomy that "...could help improve both the learning and teaching and testing of various aspects of L1 and L2 vocabulary learning...". In the discussion, Loucky provided a working definition of what successful access, i.e. acquisition of a word, means. He stressed that learners need to demonstrate competencies in understanding the *form*, *meaning*, and *use* of a word in order to achieve successful acquisition of an accessed word. In the framework Loucky presented, he provided categories of knowledge that involved five strategic steps as effective strategies to help students develop good guessing skills for successful acquisition when they access a word.

(Reported by Terry Laskowski
Kumamoto University)

プレイスメントテスト記述課題の量的・質的分析

西村よしみ (筑波大学留学生センター)

筑波大学留学生センターでは、日本語授業に参加する新入留学生に対し、各学期プレイスメントテストを実施している。テスト内容は、多肢選択式テスト(文法・聴解・語彙)・SPOT (Simple Performance Oriented Test)・記述課題で、所要時間は2時間半程度である。

記述課題は、制限時間 15 分の間と与えられたテーマ(グラフの説明・4コマ~6コマ漫画の説明)について記述する。本研究の目的は、記述課題を量的・質的観点から分析し、特に多肢選択テストとの相関を検討することにより、記

述式テストを評価するための有効な指標を探ることである。

研究方法としては当初、記述課題を総文字数、T-unit 数、T-unit/総文字数、正解 T-unit/総 T-unit 数、Idea unit などの指標と多肢選択式テストの下位テストとの相関を検討し、更に質的分析の指標として Idea unit 群の質的分類を中心に行う予定であったが、評価を行う上で効率的な観点として「従属節の使用」が浮上してきたため、そこに焦点を絞ることにした。従属節の使用には、能力に応じて階層性と深まりが観察されるためである。

記述課題の評価項目は、1) 量(文字数)、2) 複雑さ、3) 正確さ、4) 豊かさ(語彙)、5) 構成・内容と5項目があるが、「従属節の使用」は、2) 複雑さの指標となる。そして、プレイスメントの判断材料としては、1) 量(文字数)と 2) 複雑さ(従属節の使用)が特に有力な目安となる。総文字数で上・中・下の3グループに分類し、その後従属節の使用状況に応じて6段階に分類する。

上記の記述課題の評価と多肢選択式テスト結果には相関性が観られるが、特に従属節の分析にはどうしても主観性が反映するため、テストの信頼性と妥当性と実用性の観点からテスト形式と評価方法に関する研究を今後も進めていく予定である。

(報告:石井和仁 福岡大学)

TOEIC で見る新カリキュラムの成果

吉川勝正(熊本学園大学)

熊本学園大学経済学部国際経済学科では、2000 年度より新カリキュラムを導入し、学生の卒業時の英語力を入学時よりも高めることを目標とした。学生の英語力の客観的評価基準として TOEIC の得点を用いた。

新カリキュラムの特徴を述べると、英語履修希望者は1年次の春に TOEIC を受験し、300 点未満の学生は後期週2コマの基礎英語の履修が義務づけられる。300 点以上の学生も履修可能。1年前期には英語の授業は無い。英語1~4は全て半期制で、それぞれ2単位で週2コマ。4単位コースの学生は2年後期で終了し、8単位コースの学生は3年後期で終了する。8単位コースの終了目標として TOEIC400 点を掲げている。基礎英語と英語1と3は日本人教師が担当し、英語2と4は英語母国語話者の教師が担当。3年生と4年生の演習の中に、英語をさらに勉強したい学生のための演習(週1回通年)があるが、8単位コースの学生は3年次でも週2回の授業があるためか、英語の演習を取る学生は 10 名程度である。

上記カリキュラムにおける 2000 年度入学生 85 名の、1年次春、1年次冬、2年次春、2年次冬、3年次春、3年次冬の TOEIC の平均点を見てみると、292 点から 356 点まで徐々に伸び

ている。また、2001年度入学生50名の1年次春、1年次冬、2年次春、2年次冬、3年次春、(3年次冬は受験予定)のTOEIC平均点を見ると、285点から320点まで徐々に伸びている。大学全体のTOEIC平均点の方は、2000年度入学生の場合318点から329点までの伸びで、途中342点も記録しているが、3年次での伸びが見られない。2001年度入学生の場合は313点から313点と1年次と3年次の点数が同点となっている。こちらも途中342点を記録しているが、学年が進んでもやはり伸びが見られない。

今後の検討課題としては、学生の学力に対するTOEICの難易度が適切かどうかという点や、学生個人の学習状況の影響度や、各担当教師の評価時における定期試験とTOEICの比重のかけ方の問題等がある。

(報告:石井和仁 福岡大学)

Quantitative and Qualitative Analysis of Writing Task in Placement Test

Yoshimi Nishimura (*Center for Foreign Students of Tsukuba University*)

The new students of the Center for Foreign Students of Tsukuba University, who are planning to have the Japanese classes, are to take a placement test in each semester. The test includes multiple-choice questions about grammar, listening comprehension and vocabulary, and SPOT — Simple Performance Oriented Test, and the writing task.

In the writing task the testee has to describe what a graph shows and make a story out of a comic strip within 15 minutes. The purpose of this research was to analyze the writing task from the quantitative and qualitative views, and find an efficient way to evaluate the task especially by examining the correlation between the multiple-choice questions and the writing task itself.

The approach that was employed at the first stage of analyses was to examine the correlation between the multiple-choice test result and the analysis of the writing task in terms of the number of letters used in it, the number of T-units, T-units observed in the total number of letters and the correct T-units observed in the total number of T-units, with focus on the qualitative analysis of the idea units. However, as an efficient way of evaluating the writing task, the analysis of the use of the subordinate clause turned up because the proficiency of the Japanese language was reflected in the use of subordinate clauses; therefore, our focus of approach shifted a little accordingly.

There were five evaluation items for the writing task: 1) quantity or the number of used letters, 2) complexity, 3) exactness, 4) richness or

the amount of vocabulary, 5) structure and content; and the use of the subordinate clause can be the index of complexity of the writing. In the process of placement 1) quantity or the number of used letters and 2) complexity or the use of the subordinate clause can be used as very effective standards.

The writing examinees were to be divided into three different groups according to the number of letters they use, and subdivided into six groups in view of how they use the subordinate clauses.

Significant correlations were observed between the above approach of evaluation of the writing and the result of the multiple-choice test, but the examiner's subjective commitment may not be eliminated. Therefore, research for making more reliable, appropriate, and practical test and the better evaluation method should be continued.

(editor's report)

The Outcome of the New Curriculum Examined by TOEIC

Katsumasa Yoshikawa (*Kumamoto Gakuen University*)

Since 2000 the Department of International Economics of the Faculty of Economics of Kumamoto Gakuen University has introduced the new curriculum, and aimed to enhance the junior and senior students' English ability more than the freshmen's and sophomores' counterpart. As an objective standard to examine the outcome, TOEIC points have been used.

The new curriculum requires the freshmen to take TOEIC in the spring, and those who scored less than 300 points have to take two English classes a week in the second semester — there are no regular English classes in the first year. English classes categorized into four courses from 1 to 4 are all semester classes, and every two classes in the same course are a pair for a week. Each course is given two credits; therefore, those who want four credits in English start their courses 1 and 2, and finish them in sophomore. If they want eight credits, they have to take two more courses 3 and 4 in junior. The final target points in TOEIC is 400. The basic course and the courses 1 and 3 are instructed by the Japanese staff, while the courses 2 and 4 by the native speakers.

Juniors and seniors can also take the seminar type course, but only about ten of them do so now.

The average TOEIC points 85 students scored under the new curriculum were 292 points in the spring in 2000 and 356 points in the winter in 2002. They took TOEIC in every spring and winter for three years. Their average points gradually escalated. And those average points 50

freshmen in the same department scored in the spring in 2001 were 285 points, and they scored 320 points in the spring in 2003 — they are to take the test in the winter in 2003. Their average points gradually escalated again. On the other hand, the average points all freshmen of Kumamoto Gakuen University scored in the spring in 2000 were 318 points, and they scored 329 points in the winter in 2002. Although they scored 342 points half way through between 2000 and 2002, their average points showed less increase than those of the students in the Department of International Economics did. Almost the same phenomena can be seen about the average points of all the students and those in the Department of International Economics who entered the university in 2001.

The problems which have to be considered are 1) if TOEIC is appropriate for the students in difficulty, 2) how much individual learning affects the result of TOEIC, and 3) how the teaching staff should evaluate the English ability of each student, taking TOEIC points into consideration.

(Reported by Kazuhito Ishii
Fukuoka University)

語彙知識によるスピーキング能力の予測可能性

小泉 利恵 (筑波大学大学院博士課程)

語彙知識はスピーキング能力の一部と考えられ、スピーキング能力で大きな役割を果たすと考えられている。しかし、「語彙知識が、スピーキング能力や発話のどの部分にどの程度影響するか」については、あまり調べられてきていない。本研究では、(1) その点を検証し、(2) 「語彙知識からスピーキング能力をある程度予測し、適切なスピーキングテスト (ST) のタスクを選択することが可能か」を探索することを目的とする。本研究は、ST の実用性向上や、スピーキング能力の実証的なモデル構築に向けての第一歩として位置付けられる。

参加者は約170名の日本人英語学習者 (中学3年生150名・高校1年生20名) で、2種類のテスト (ST と筆記テスト) を受験した。ST は、面接官と受験者が対面して行われた。本研究では、5つのタスクの中で長めの発話を引き出した3つ (自己紹介・好きなものについて・2枚の絵の比較) を分析対象とした。書き起こした発話は、語彙・流暢さ・正確さ・複雑さの観点で分析し、評価を行った。筆記テストには、発表語彙・受容語彙・文法問題が含まれていた。分析では、項目応答理論・共分散構造分析などが用いられ、発表語彙知識がスピーキング能力に強く影響することなどが示された。

(editor's report)

日本語 Can-do-statements の IRT 尺度化

野口 裕之 (名古屋大学)
和田 晃子 (北海道大学)
熊谷 龍一 (名古屋大学)
脇田 貴文 (名古屋大学)

日本語 Can-do-statements は、日本語能力試験の妥当性を検証するための外的基準として、開発された自己評価方式による、日本語能力の評価尺度である。「読む」「書く」「話す」「聞く」の4技能につき各15項目、全部で60項目の質問から構成される。例えば、「郵便局・銀行での説明がわかりますか」などの質問に、「1 全くできない」から「7 問題なくできる」までの7段階で回答させる。これまでに項目分析、因子分析、信頼性分析などが実施され、1) すべての項目が、一定程度以上の識別力を持ち、2) 質問項目群は、全体として一因子構造を示し、3) 技能別・全体ともに高い信頼性を示し、4) 大学のプレイスメント・テスト (大部分の項目が日本語能力試験から抜粋) とは高い相関を持つ、などの結果が報告されている。本報告では、この日本語 Can-do-statements に対して、項目応答理論による分析を実施した結果について示す。具体的には、1) 複数の IRT のモデルの中から最適なモデルを示す、2) 多値型の IRT モデルを用いて、評定尺度の等間隔性について確認する、3) 尺度に含まれる項目の難易度の分布の適切性を検討する、4) テスト情報量を用いて測定精度をきめ細かく評価する、5) IRT 特性尺度値と従来の加算得点との関係を検討する、6) 回答者の属性によって回答傾向の異なる (DIF) 項目を同定する、などである。IRT 尺度化によって、日本語 Can-do-statements をより洗練度の高い測定尺度にすることができる。

(editor's report)

Vocabulary Level Test at the 1000-Word Level

Paul Westrick (Kyushu University)
J. Lake (Fukuoka Jo Gakuin and Kyushu University)

The Vocabulary Levels Test (VLT) (Nation, 1983; 1990) is widely used in language education research for estimating vocabulary sizes of ESL/EFL students (Nation, 2001; Beglar, 1999; Schmitt, N, Schmitt, D. & Clapham, C, 2001; Read, 2000). The VLT is used to correlate vocabulary size with other linguistic research measures, as part of a test battery in placement tests, to gauge the appropriate range of graded readers for extensive reading, to gauge vocabulary size for materials development and other pedagogical purposes. Traditionally the levels have been at the 2000, 3000, 5000, 10,000 and "University" word levels. The 2,000-word level tests include

items from the 1-1,000 and the 1,001-2,000 levels; however, in an EFL context or with lower ability students there may be a need to assess more precisely below the 2,000-word level. The Westrick (2001) 1,000-word level test was designed to do this. In this presentation, the development, analysis and validity of this 1,000-word level test was explained and several suggestions were made for future research directions and uses.

(editor's report)

Examining the Rationale of Vocabulary Size Tests Based on Word-Frequency Levels

KATAGIRI Kazuhiko (*Reitaku University*)

Mochizuki (1998) developed one form of the Vocabulary Size Test (VST), which consists of seven sections (1,000, 2,000, 3,000, 4,000, 5,000, 6,000, and 7,000 word-frequency levels). Since then, Mochizuki has developed the two more forms of the VST so far. Katagiri and Mochizuki (2002) confirmed the parallelism of the three forms. In this presentation, a report was made about whether the rationale of estimating vocabulary size of Japanese learners of English through each section (each word-frequency level) would be appropriate from the psychometric view point.

University students whose major is English took six sections (2,000 - 7,000 word-frequency levels) in each of the three forms of Mochizuki's VST. Since six variables (subtotals of correct answers in each section) were obtained in each form, Pearson's product-moment correlation coefficients among the six variables were calculated in each form, and explorative factor analyses were conducted.

The results showed that six sections (2,000 - 7,000 word-frequency levels) had high correlation with one another (especially the stronger correlations with the closer word-frequency-level sections) and that only one factor was extracted in each of all the three forms. This may indicate that no section (no word-frequency level) is heterogeneous and that each word-frequency-level section is sequentially connected. Theoretically no section (no word-frequency level) should be heterogeneous and that each word-frequency level section should successively be connected; empirically, in this study, it may partially have been confirmed from the psychometrical results obtained through the interaction between the examinees and the VSTs.

(editor's report)

An Investigation of Method Effects in Reading Comprehension Tests: Test Structure and Response Format

Miyoko Kobayashi (*Kanda University of International Studies*)

This presentation gave us very beneficial hints to find out the correlation between ESL learners' performance and input styles: text organization and response format. She also investigated the types of text structures of reading passages of Center Examinations.

The variables are: 1. text organization (problem-solution, causation, description, association) 2. response format (cloze tests, open-ended questions, summary writing) 3. learners' English language proficiency levels (low, middle, high).

The subjects were 754 Japanese university students and they were given two topics in four text types in three test formats. An English proficiency test was also conducted to measure their language ability.

The findings are: 1. Reading comprehension test performance is affected by both text organization and response format. 2. Test performance in open-ended questions and summary writing was better in more tightly-organized texts. 3. The effects of better text organization were clearer with higher proficiency learners especially in summary writing. However, less proficient students did not show much variation.

The investigation of the Center Examination was also very suggestive. According to the presenter, there was no clear text structure in the passages used in the Center Examination, which seems inappropriate to identify a student's performance.

実践報告—理科系の大学生に求められる英語力とテスト、およびその対策

河野円 (*星薬科大学*)

薬学部の学生が入学試験から大学を離れるまでに対応しなければならない英語のテストを調査し、それぞれのテストの内容と大学の英語教育の内容の関係について報告したものである。大学で受ける英語のテストは 1. 定期試験 2. TOEIC 3. TOEFL 4. 大学院入試 の4種類であるが、それぞれまったく異なる形式・内容のものに対応するわけである。理科系の学生は特に他の授業が多く課外での自学自習時間の確保も困難な中、ターゲットとするテストによって、多読や精読、訳読、など異なるスキルが求められ、適切な指導をすることの難しさが指摘された。

(報告: 小山由紀江 名古屋工業大学)

Introducing Exit Competencies into a University English Course

Brendan Rodda (*Kanda University of International Studies*)

The presentation provided information about the introduction of exit competencies into BEPP, a general skills course for first-year students at KUIS. A report was made about the research into the issue of how this process would be influencing the course curriculum in general.

The presenters are currently working on specifying clear and detailed exit competencies and plan to trial assessment tasks based on these later this year. Full implementation of the exit competencies assessment system is scheduled for the 2005-06 academic year.

This system will allow students who satisfy the exit competencies to move out of BEPP early and will require students who do not satisfy them to remain in the course beyond their first year. Students will be assessed on their performance on a series of tasks that require the integration of a receptive and a productive skill (e.g. listening-speaking), reflecting the course's focus on integrative skills. Tasks will be similar to communicative tasks used in the classroom and will take 20-30 minutes to complete. Students will be aware of the exit competencies they need to satisfy but will not have specific knowledge of the tasks, so will not be able to prepare or practice their actual performance.

(editor's report)

Standardized Testing as a Graduation Requirement: The Case of the CET (College English Test) in China

Robert J. Fouser (*Kyoto University*)

This paper serves as a comprehensive introduction to the CET (College English Test), which has been required for all students of four-year universities to receive a bachelor's degree since 1987. The test is developed by the National College English Testing Committee of China and is taken by 2.5 million students annually. The test covers listening, reading, and writing and has two versions: Band 4 and Band 6. Band 4 covers material taught in the first four semesters of university and is a graduation requirement for all students. Band 6 covers material taught in three semesters of university and is required for English majors and, graduate students, depending on the university. Students who have not passed Band 4 cannot

take Band 6 and students who fail Band 6 cannot retake it. The CET Spoken English, which was developed in 1999, is used as a supplemental test to Band 4 and Band 6.

Through a careful examination of examples from Band 4 and Band 6 tests from previous years, the two tests were compared with each other and with other commonly used standardized tests in terms of organization, item characteristics, and possible difficulty. Using the JACET 8000 list of basic words, the vocabulary and readability of reading passages were analyzed to provide a quantitative basis for comparison. Finally, by referring to the literature and translations of official information from Chinese, discussion was made on the pedagogical and social implications of the CET on English teaching in higher education in China.

(editor's report)

非母語話者の発話に対する日本語教師の評価—教育経験の多少はどのような影響を及ぼすか— 鈴木 秀明 (神田外語大学)

日本語教師が非母語話者の発話を評価する際に教育経験はどのような影響を及ぼすかを調査した。先行研究では、発話能力の評価において、教師と一般人では評価の仕方に違いがあることは報告されているが、教師の教育経験の多少による評価の違いについての研究はまだ見られない。

調査協力者は、日本語教師 40 名で、平均教育経験年数は 5.6 年であった。調査では ACTFL-OPI のサンプルテープを使用し、発話テープを調査協力者に関して評価してもらった。調査では、全体的評価、分析的評価、発話評価を行う際の教師の自己基準を記述してもらうというタスクを実施した。全体的評価では発話を全体的印象で5段階評価してもらい、分析的評価では、発話をテキストタイプ、文法、音声、語彙、流暢さ、機能、内容の7項目について ACTFL のガイドラインに基づいて作成された基準で5段階評価してもらった。さらに、発話能力を評価する際の着目点や評価基準について、記述形式で回答してもらった。調査で得られた結果を教育経験の多いグループと少ないグループに分け、グループごとに分析した。本発表ではグループごとの分析結果、教育経験が発話能力の評価に及ぼす影響について報告された。

(editor's report)

WEB上の英語語彙力テストの開発 風斗 博之 (東北学院大学)

発表者は一般英語の授業履修者を中心に、ネット上に作成した英語単語テストを受験させ、各受験者の語彙力を測定する試みを行っている。受験者からは、同時にTOEICやTOEFLのスコアを提供してもらい、それらのスコアをもとに語彙力を表示する。テストはIRT(項目応答理論)に基づいて作成し、問題項目や受験者が増えるほどその測定の正確さも増えるように作られている。受験者の能力・項目の難易度や識別度の計算には最尤推定法を使用している。問題の形式は現在、英和および和英の記述式であるが、選択式、絵や例文、音声を用いたものなど様々な形式を開発中である。

テスト問題は再編集し、別のテストセットとして再利用することができる。語彙力の正確な評価が受験者の学習意欲をかきたてる。また、テストのデータから得られた、問題項目の難易度から英単語の難易度を導き出すことができる。たとえば、よく使われる頻度順のリストとは異なる、英日方向の英単語の難易度順(知名度順)のリストなどを作成し、語彙学習教材で利用することなどが可能となる。

発表では、個々の問題の難易度の計算法を含め、CGIプログラムでの実際の処理方法なども解説された。

Self-Assessment, Peer-Assessment, and Teacher-Assessment on Japanese University Students' Writing an English Composition Sumie Matsuno (Aichi Prefectural University)

Professor Matsuno performed a series of multi-faceted Rasch analyses to explore ways in which self-assessment, peer-assessment and teacher-assessment methods worked interactively. A six-point rating scale was employed for students to assess their own and their peers' essay compositions and for Professor Matsuno herself as a teacher-assessor. Sixteen performance criteria were rated as items. It was found that students in general were positioned high on the common logit scale. Rater severities varied a lot. Self-raters tended to be harsher than peer-raters, indicating how difficult it would be to analyze such data reliably in classical statistics alone. Item difficulties also had considerable variations. Grammar tended to be the hardest item and spelling the easiest. Infit statistics revealed that teacher assessment had some problems, with quite a few items identified as misfitting. Multiple teacher raters might have been needed in order to make a fair comparison with other groups of raters. A number of rater-writer bias interactions were exhibited. These bias

estimates can be analyzed qualitatively and used as a frame of reference for raters to improve their assessment practices. Finally, category statistics were discussed. Because of the infrequent lowest possible scores observed, scores of one and two could be merged as a single category, and the whole data could be reanalyzed on a five-point scale. This line of study merits continued analysis and discussion from both theoretical and pedagogical perspectives, and is expected to encourage learners to get actively involved in the process of assessment as well as learning.

(Reported by Ken Norizuki
Shizuoka Sangyo University)

シンポジウム The Role of Testing in Language Learning

Coordinator & Panelist:

Yuji Nakamura (Tokyo Keizai University)

Panelists:

Ken Norizuki (Shizuoka Sangyo University)

Garold Murray (Tokai University)

Yoshinori Watanabe (Akita University)



本年度の全国大会テーマである「言語学習とテストング:テストングが言語学習に果たす役割」について、4人のパネリストにより英語で発表がなされた。

まず、中村氏は、このシンポジウムの調整役兼最初の発表ということで、impact, backwash, influence, evaluation, assessmentといった術語に関する定義づけの説明から始まり、次に勤務校でのテストに関連した教育実践を紹介した。TOEICを2回受験させて英語力の伸びを測定しているが、このことによって学生の自己評価が促進された。また、TOEIC対策コースは、全般に学生のアンケート結果がよいという実践報告が今まで複数なされている。といったことについて述べ、結論として、全員には(特に低得点者には)必ずしもやる気を出させることはできないけれども、目標を学習者にもわかりやすくはっきりと定め、かつ、その

第18回研究例会報告

目標が学生にとって到達可能なレベルで設定されれば、学習者のやる気を引き出すことができると主張した。

法月氏は、昨年教えたクラスにどのような授業をしどのように測定をしたかについて発表した。テストで大切なことは、(大学の英語教員としては)テストの専門家である前に、教育者として、学生をやる気にさせたり、動機づけおこなったり、そしてその結果学生の英語力が向上していくことが大切であると述べた。

Murray氏は、learners' autonomy がキーワードであり、日記や、チェックリストや、ポートフォリオなどを活用するなどして、学習者に学習を構造化して責任を持たせることが大切であると述べた。また、言語学習の道のりは長く、教室での学習を卒業してからの道のりがまた長いことから、meta-cognitive awarenessを発達させることも重要である。テストは、フィードバックをして、良い効果をもたらす内発的動機づけにも貢献するのが望ましい。「学習者の自己測定」と「教師の測定」には、高い相関関係が見られる。つまり、学習者は自分で自分の学習を測定できるということである。よって、言語学習過程には、常に測定が組み込まれている。といったことについて述べた。

渡部氏は、テストの波及効果について論じた。波及効果は、単一の単純な現象ではなく、複雑な現象であり、テスト自体以外に BAK (Belief, Assumption, Knowledge) といった要素にも大きな影響を受ける。前勤務校の国際基督教大学の入学試験にはリスニング・テストが課せられているのだが、入学した学生に聞いてみるとあまりリスニングの勉強をしてきていないということがわかった。理由は、国際基督教大学の入試リスニング・テストはあまりにも難しすぎるから、ということらしい。一方で、国立大学の入学二次試験でリスニング・テストが課せられている大学の学生に聞いてみると、これまたリスニングの勉強をしてきていないということがわかった。理由は、国立大学の入学二次試験のリスニング・テストは易しすぎるから、ということらしい。つまり、「リスニング・テストを入学試験に課せば、受験生はリスニングの勉強をする。」といった論理は実際の教育現場では必ずしも成り立つわけではないという、「波及効果」の一筋縄ではいかない複雑性を示す具体例なのである。次に、中学校・高校で導入された絶対評価の波及効果に関して話も話が及んだ。今のところ、よい効果は見られていないとのことであった。最後に、波及効果の研究は “ethnographic rather than experimental” であり、波及効果研究に必要なのは「スタミナ」と「一緒にやる仲間」として「ほんのちよっとの知性」であると締めくくった。

4人の発表後、小休憩を挟んで、フロアーから質問や意見が出され、活発な質疑応答がおこなわれた。

(報告:片桐一彦 麗澤大学)

2003年12月20日、静岡産業大学国際情報学部(静岡県藤枝市)にて第18回研究例会が開催された。今回は静岡産業大学と近隣の小・中・高・大・一般の英語教師と英語教師を目指す学生が会員となっている「英語教育をおもしろくする会」が共催するイベントとなり、中・高・大の連携と言語テストのすそ野を広げることを意識した企画となった。

大友賢二先生のご挨拶のあと、議論の口火を切って中村洋一先生の評価の体系を包括的に捉えるご講演が始まった。途中からジャケットを脱がれるほどの熱弁で、あとに続く発表の熱気を予感させるものとなった。

中学の発表者は伊東市立北中学校の杉本博昭先生。教科書に対応する具体的な目標基準を設定することの実務的な難しさを強調され、ポートフォリオなど継続的評価のあり方について豊富な資料に実例を添えて説明された。続いて静岡県立掛川西高校の織田敦先生が、実践的な言語使用能力の育成を目指したパワーポイントによるインターアクティブな授業の展開についてご説明され、定期試験の作問例も提示された。熱気を増した杉本先生と織田先生のご発表は「評価は評価の問題だけで議論が終わらない」ことを如実に示すものとなり、大学の部の発表が始まる頃には、すでに開会から3時間を越えようとしていた。

大学の部は、私(法月)のリーダビリティ新公式の開発を目指した取り組み、そして同僚で共同研究者の宮崎佳典先生とゼミ学生工藤良一氏の新公式のプログラム化の試案と続いた。情報数学がご専門の宮崎先生は過去数年間ゼミ学生とともにリーダビリティ公式のプログラム化について研究されてこられたが、今回、新公式の実装を想定して辞書サイトをリンクさせたプログラムのデモを披露された。

発表の後には、フリーディスカッションを予定していたが、すでに終了予定時間を30分近く過ぎていたため、やむを得ず割愛し、講演と発表全般の質疑応答、大友先生のご講評で閉会とした。自由な意見交換の場としてのフリーディスカッションや個別の質疑応答の時間が取れなかったことは、短い時間内に中身の濃いプログラムを詰めすぎた企画者の私のミスであるが、できることなら何時間でも議論を続けていたい(私だけかもしれないが)不思議な興奮に包まれていた。

大友先生がご講評の中で「JLTAとおもしろくする会に共通点がある」とおっしゃっていたが、そのお言葉が意味することは、教育指導や学習を縛ることなく、英語学習への意欲を駆り立て、教育と学習の効果を高める評価システムの確立にあると思う。それは、言語テストの専門家が集まるJLTAの果たす役割の一つとして、本年度（2003

年）の全国大会のテーマにもつながる重要な課題である。

地方の私立大学に勤務して、地域に言語テストの根を広げることの意義はかねがね感じていた。今回の研究例会でその思いはますます強くなった。

（報告：法月 健 静岡産業大学）

@@

事務局よりお知らせ

◆ 会員の皆様には、このニューズレターと一緒に2003年度の総会報告を送付いたしますので、ご査収ください。

◆ JLTA 研究助成の応募締め切りは毎年1月末日です。詳しくは研究紀要の「研究助成規約」をご覧ください。

◆ 転勤、転居等、JLTA の名簿記載事項に変更が必要な場合は、速やかに、事務局までご連絡下さい。また、銀行引き落としによる会費納入を利用している会員で、吸収・合併などにより、銀行名、支店名、口座番号等が変更になった場合は、必ず事務局まで、その旨をお知らせ下さい。

◆ JLTA の活動に対するご意見やご要望、Newsletter等への掲載希望記事などがありましたら、事務局までご連絡ください

◆ The JLTA office would be grateful if you could update us on your recent achievements relevant to the field of language testing and evaluation. Any information on your presentations, publications, awards, and so forth would be greatly appreciated. The relevance of the information will be evaluated by the office and given in the newsletter in due course.

◆ 会員の皆様の当該分野での近況をお伝えください。テスト・評価関係本を出版した、論文を発表した、賞を受けた、博士論文を提出した、など。随時報告してゆきたいと考えております。

◆ 市町村合併により、事務局の住所表示が以下になりました。今後、事務局への連絡には、以下の住所をご使用いただきますよう、よろしくお願いいたします。



日本言語テスト学会事務局
〒389-0813 長野県千曲市若宮 758
TEL 026-275-1964 FAX 026-275-1970
e-mail: youichi@avis.ne.jp
URL: <http://www.avis.ne.jp/~youichi/JLTA.html>