

JLTA Newsletter

日本語テスト学会
The Japan Language Testing Association

JLTA Newsletter No. 15 発行代表者：大友 賢二 2002年(平成14年)9月30日発行

発行所：日本語テスト学会 (JLTA) 事務局

〒389-0813 長野県埴科郡戸倉町芝原 758 TEL 026-275-1964 FAX 026-275-1970

e-mail: youichi@avis.ne.jp URL: <http://www.avis.ne.jp/~youichi/JLTA.html>



第6回(2002年度)全国研究大会のご案内

大会テーマ： Language Testing and Second Language Acquisition Research

言語テストと第二言語習得研究

日時：2002年10月20日(日) 8:30～18:50

会場：東京経済大学6号館、〒185-8502 東京都国分寺市南町 1-7-34

参加費：会員1,000円、一般3,000円(事前申し込み不要)

プログラム：

10月19日(土)

16:00～17:30 理事会・委員会 (6号館7階小会議室)

10月20日(日)

8:30～ 受付 (6号館1階)

8:50～9:00 開会の挨拶 (7階大会議室) 総合司会 中村 優治(東京経済大学)
会長 大友 賢二(常磐大学)

9:00～10:20 研究発表(発表30分、質疑10分) I 9:00～9:40、II 9:40～10:20

第1室 (7階中会議室1) 司会 藤田 智子(東海大学)

発表 I 「複数の語彙レベルによるリーダビリティの測定の試み」

長沼 君主(東京外国語大学大学院)

発表 II Peer-, self-, and teacher-assessment in group presentation: A pilot study

神前 陽子(武庫川女子大学)

第2室 (7階中会議室2) 司会 Laura MacGregor(学習院大学)

発表 I Validating Scores in University-wide Group Oral Tests

Alistair Van Moere (*Kanda University of International Studies*)

発表 II Development of multiple-choice grammaticality judgement test types

AMMA Kazuo (*Tamagawa University*)

(以下に続く)

第3室 (7階中会議室3)

司会 Jeff Hubbell (法政大学)

発表 I Entrance Examination Archaeology: The Case of Kyoto University. Robert J. Fouser (Kyoto University)

発表 II A Case Study on the Influence of Environment and Motivation on Second Language Learning in Taiwan. Chih-hui Chang (Da-yeh University, TAIWAN)

第4室 (7階中会議室4)

司会 大坪 一夫 (麗澤大学)

発表 I 「語彙習得の測定におけるテスト」 山崎 朝子 (武蔵工業大学)

発表 II 「TOEFL Summary Reports (1993-2001) から見えるもの」 池田 央 (教育測定研究所)

10:20~10:40 休憩

10:40~12:30 基調講演 (7階大会議室)

司会 木下 正義 (福岡国際大学)

紹介 会長 大友 賢二 (常磐大学)

「『実践的コミュニケーション能力』の達成度設定と評価のために」

講師 青木 昭六 (愛知学院大学)

12:30~13:30 昼食

(役員会: 6号館7階小会議室)

13:30~14:50 研究発表 (発表30分, 質疑10分)

I 13:30~14:10、II 14:10~14:50

第1室 (7階中会議室1)

司会 櫻井敏子 (神戸松蔭女子学院大学)

発表 I Reliability/Validity of "Invisible-Blank Filling" Items. Shizuka Tetsuhito (Kansai University)

発表 II Validating Speaking Test for Japanese Junior High School Students. Rie Koizumi (Doctoral Course, University of Tsukuba)

第2室 (7階中会議室2)

司会 塩川 晴彦 (北海学園大学)

発表 I Inferring Japanese Learners' English Ability by way of Measuring Their Vocabulary Size. KATAGIRI Kazuhiko (Reitaku University)

発表 II Rater Training Effects on On-line Peer Assessment of EFL Individual Presentations. Hidetoshi Saito (Hokusei Gakuen University)

第3室 (7階中会議室3)

司会 清川 英男 (和洋女子大学)

発表 I 「日本語と英語のライティング能力の実証的、統計的、比較研究」 中村 優治 (東京経済大学)・飛渡 洋 (都立農業高等学校)

発表 II 「高校生のための英語学力テスト分析—学校英語テストの改善に向けて」 齊田 智里 (茨城県立並木高等学校)

14:50~15:10 休憩

15:10~16:40 パネルディスカッション

(7階大会議室)

「言語テスト研究 (LT) と第2言語習得研究 (SLA) の相互作用」

司会・発表 伊藤彰浩 (愛知学院大学)

発表者 白畑知彦 (静岡大学)

齋藤英敏 (北星学園大学)

16:40~17:10 総会 (7階大会議室)

司会 中村 優治 (東京経済大学)

報告 中村 洋一 (常磐大学)

17:10~17:20 閉会の挨拶 (7階大会議室)

副会長 Randolph Thrasher (国際基督教大学)

17:20~18:50 懇親会 (7階ラウンジ)

司会 中村 優治 (東京経済大学)

会場近辺のホテルは、次の2つが便利です。

ホテルメッツ国分寺 TEL: 042-328-6111

9,500円前後 国分寺駅南口徒歩2分

ビジネスホテルダイワ TEL: 042-324-5221

8,000円前後 国分寺駅南口0分

XX

第3回 JLTA

言語テストワークショップ報告

XX



連日30度を越す猛暑の中、各地から集まった意欲的な参加者を迎え、第3回 JLTA 言語テストワークショップが7月30日(火)、31日(水)に、茨城県水戸市の常磐大学で行われた。参加者は総勢27名、関東圏のみならず、北海道や海外からの参加もあり、この会に対する期待の大きさを感じさせた。各セッションは全て常磐大学総合情報センター4階のCALL教室で行われた。各座席にコンピュータが備えられ、しかし圧迫感のない形のCALLラボで、講演でのプレゼンテーションやデータ実習などにもCALLラボの機能が多く活用された。

[第1日目講演1]

言語テストの基礎(大友 賢二)

開会行事や説明の後、大友賢二教授による講義が行われた。タイトルは「言語テストの基礎」。言語テストを作成・実施するうえで根幹に関わる課題について3つのセクションに分け説明された。

1. 言語能力のモデル

言語テストを考えるには、『言語能力』『言語学力』の構造を検討することが必須である。歴史上、1) 技能と構成要素に基づくモデル、

2) 語用論に基づくモデル、3) 意思伝達能力に基づくモデルという流れがある。

1) は、Lado を中心とした4技能+翻訳などの skill と発音、文法、語彙、文化的知識などの component を統合的に考える取り組みで、Spolsky (1978) では psychometric-structuralist trend、Oller (1979)では言語能力分解可能仮説と称されたものである。2) で Oller (1979) は、言語能力は分解不可能な一つの能力であると考へた。しかし、この主張は Farhady, Carroll, Upshur との議論の後撤回され、partial divisibility を認める事となった。3) としては、Canale and Swain (1980) の communicative competence の概念を Bachman and Palmer (1996)が扱ったものだが、これについてもまだ議論は尽きない。

2. テストの質と測定準拠

テストの質を高めるための検討課題として、1) 妥当性の検討、2) 信頼性と一般化可能性 (generalizability theory, g-theory)、3) 集団基準準拠測定 (norm-referenced measurement) と目標基準準拠測定 (criterion-referenced measurement) の3つがある。従来は様々な種類の妥当性があるとされてきたが、近年は妥当性が unitary concept であるとする考えが主流である。また妥当性が何を指すのかについても多くの議論調査が行われている。2) については、古典的理論における信頼性の算出法のあと、一般化可能性理論 (g-theory) が紹介された。一般化可能性理論では、エラーに

複数の構成要素を捕らえるため、その情報を利用してより効率よく正確な問題作成を可能にする。3) では、この後の根岸雅史先生の講演にも関連する目標基準準拠測定についての問題提起がされた。やはり問題となるのは cut score をどのように、どこに設定するかであり、その代表として Popham (1981) の contrasting-group model が紹介された。これは特徴を良く知った被験者のテストデータをもとにその performance curve を検討し、cut score を定めるというものであり、今後大いに参考にされるべきであろう。

3. 新しい動向

最後に1) 項目応答理論の特徴と2) CAT、3) performance standard についての説明があった。項目応答理論は、test-free person measurement, sample-free item calibration, multiple reliability estimation を特徴とし、古典的テスト理論の問題点を補完する。現在、項目応答理論は TOEFL - CBT などの CAT (computerized adaptive testing) に応用されているが、これは先に述べた特徴の上に成り立つ項目応答理論の不変性 (invariance) により可能になるものである。学習者の実力に合わせて問題を提示し、しかし共通の尺度で正確な評価を効率よく行う、理想的な評価である。

3) ではまず、基準に関する用語の混用について指摘があった。standard, cut score, passing score 等の用語はまだ混用が見られる。また test specification の作成も performance standard の設定には欠かせない。Davidson and Lynch (2002) 等が参考文献として挙げられた。

言語テストの歴史から現在の問題点までを一つの流れで見ることが出来た。大変意義のある、中身の濃い講義であった

[第1日目講演2]

新学習指導要領における

測定と評価の課題 (根岸 雅史)

初めに、現在の評価の状況についてお話があった。1) 今回の新指導要領導入に伴い、相対評価から絶対評価への大きなパラダイムシフトが起きていること、2) 絶対評価には観点別評価も必要となり、中学ではその土壤があるが、高校にはないためにダブル導入となり、非常に大きな変化とそれに伴う負担が予想されるということ、そして3) 絶対評価を行うために学校内のテストも大きく変わる必要があること、が提言され、このような大変革が一時に導入されるには、余程の計画性と準備が必要であることを改めて考えさせられた。

その後、評価、アセスメント、テスト、NRT と CRT 等の基礎的な概念の説明があり、それでは絶対評価と相対評価でテスト作りがどのように変わるかについてお話があった。

<変わる点>

- ① 指導目標が必須
- ② 到達目標が必須
- ③ 評価の複眼化
- ④ テスティング・ポイントの明確化が必須
- ⑤ 総合問題は不可
- ⑥ 直接的測定が必須
- ⑦ 高い信頼性が必要
- ⑧ テストの合計点は不要
- ⑨ 「ごった煮」採点は無意味
- ⑩ 説明責任が重くなる



これを実現するためのステップとして、1) 指導目標と到達目標の設定、2) 評価計画の作成(テストや他の評価法の採用など)、3) 評価の出し方の決定の3段階が必要である。

次に今年2月に作成された「評価規準の作成、評価方法の工夫改善のための参考資料」

(国立教育政策研究所教育課程研究センター)より「内容のまとめりごとの評価規準」についての説明があった。これは4技能を内容のまとめりとしてそれぞれに評価規準を定めたもので、各技能に対し、ア)コミュニケーションへの関心・意欲・態度、イ)表現の能力、ウ)理解の能力、エ)言語や文化についての知識・理解、の4つの観点についての規準が一覧表となって示されているものである。例えば、「聞くことの評価規準」ア)コミュニケーションへの関心・意欲・態度は、「聞くこと」の言語活動に積極的、意欲的に取り組んでいる、さまざまな工夫をすることで、コミュニケーションを続けようとしているかなどとなっている。

この後、この表に基づいて、accuracy や appropriateness の見方、音読、知識と運用のバランスなどについて解説があった。特に注意すべきは、「規準」(のりじゅん)と「基準」(もとじゅん)の違いである。「規準」は評価の枠組みであり、「基準」は各枠組みの内、行われるべき内容をより具体的に示したものである。本来導入に際して「基準」を示したかったが、教科書に学年指定がなく、指導内容が様々なため「規準」のみを示し、各校で「基準」を決めてもらうことになったそうである。

この他にも、絶対評価においてテストは何回必要か、一斉テストは必要か、またペーパーテストは必要ないのではないかなどの問いも投げかけられた。

参加者からの質疑応答でも、さらに突っ込んだ議論がなされた。例えば「基準」作成に

ついて、学会等が積極的に作ったらどうかという意見もあった。根岸先生は、「基準」作成の為に全県で取り組んでいる例や海外の例を出され、熱心に応じられた。重要な問題ばかりであったが、先生の分かりやすく丁寧な説明で、あっという間に時間が過ぎた

【第2日目講義と実習1】

受容技能の測定と評価(佐久間 康之)

初めに、佐久間先生は受容技能であるリーディングとリスニングの類似点と相違点を挙げられた。両者は共に bottom-up と top-down を同時に行う parallel processing を伴うが、リスニングは言語情報が瞬時に消失していくという点でリーディングとは異なる。主観的テスト、客観的テストのさまざまな種類のテスト方法が紹介、解説された後、実践的コミュニケーションに対応するには Faster Reading 以上(150-200wpm)が必要であるとの指摘があった(安藤、1979)。

その後、信頼性、妥当性、実用性の定義、パラグラフ構成の知識、推測、signal word による構成パターンの認識の影響などに触れられ、実習となった。実際に自分たちの reading speed を測定したり、同じ問題でも main idea を問う問題、inference を要する問題、自由記述問題とで大きく違うことを体験した。



[第2日目講義と実習2]

発表技能の測定と評価 (中村 優治)

午後の部ではまず、writing の holistic evaluation を体験した。

- ① 短い3つの writing のサンプルが配られ、そのうちの3つをランク付けした。
- ② それを周りの参加者の方と比較。
- ③ さらにもう一つのサンプルをその3つのどこかにランク付けする。
- ④ その過程で何を評価の基準としたか考え、criterion-setting を行う。
- ⑤ 作った criterion を Cohen (1994) と照らし合わせる。



次に Speaking の評価として speech のサンプルを評価し、設定した criterion で holistic evaluation を行った。そして今度は analytic evaluation として同じサンプルを fluency と pronunciation の観点から評価した。それぞれどこにあたるかを討議した。先生によると分析的評価の結果は相対評価の結果と異なることがよくあるとのこと。次に第4のサンプルを聞き、fluency, pronunciation について考察した。他にも grammar, vocabulary が、あるいは会話のときは conversational strategies も評価項目となる。criterion 設定の難しさを実感したが、他の参加者との交流が楽しくかつ大変有意義な実習であった。

[第2日目テストデータ処理実習]

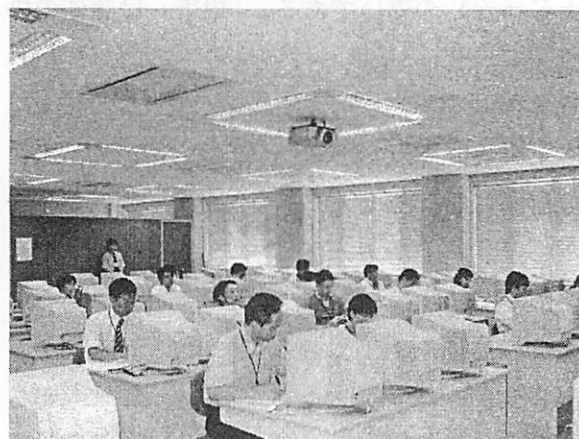
実習 (中村 洋一)

実習は、資料と共に送られてきた20項目のテストデータをもとに行われた。データ処理はTDAPの最新版であるTDAP Ver. 2.0を用いて、各自コンピュータ操作をして行った。

講習は古典的テスト理論に依る分析から始まった。基礎統計量、標準得点等の説明を経て、TDAPを用いた項目分析を試みた。項目困難度、項目弁別力、実質選択肢数(AENO)を算出した。実質選択肢数は、各選択肢が被験者によって実際にどのくらい選ばれているかを示すもので、良い錯乱肢を作る目安になる。

つづいて、信頼性係数を算出、解説がなされた。古典的テスト理論で出されたこれらの数値は個々のグループの特性も含んでおり、またテストフォームにも依存しているなどの欠点があるが、項目応答理論を用いることでこれらの問題点は解決される。項目応答理論の以下のような長所が提示された。

- 1) 絶対的な原点を持つ
- 2) 等間隔の目盛を持つ
- 3) テストに依存しない受験者特性値
- 4) 受験者グループに依存しない項目特性値
- 5) 各項目、各受験者の情報関数の提供



再びTDAPを用いて項目応答理論による分析が行われた。PROX法による計算で、まず

は項目困難度パラメータ、能力推定値、モデルとの適合度を算出した。項目困難度、能力推定値共に 0 を平均的数値とし、項目適合度は+2.000 以上をモデルに対する不適合の可能性の目安とすることができる。

項目困難度パラメータを入力し、項目特性曲線も表示された。テストの特性曲線、情報関数はこの項目毎の曲線の合計ということになる。この特性曲線と情報関数を受験者の特性やテストの目的に合わせて検討することでより良い測定が可能となる。簡単な操作で、すぐ数値が出、きれいな曲線が姿をあらわしたが、プログラムではかなりの計算が瞬時に行われている。この TDAP の新バージョンは、まもなく出版される。

今後の課題と展望として、item banking の必要性、テストのマルチメディア化、コンピュータ化に伴う更なる項目応答理論の発展可能性が挙げられた。100 分という限られた時間の中に、2 つのテスト理論のエッセンスを凝縮した非常に中身の濃いセッションであった。

閉会の言葉

Randy Thrasher 教授から ILTA で作成している Code of Ethics について、JLTA でも Code of Practice を作っているのに興味のある方は参加してほしいというお話があった。日本では、まだテストの信頼性、妥当性に関する認識が低く、様々な試験がいろいろな問題を抱えたまま実施されている。JLTA のこの試みは大きな貢献となると思われる。

参加者、講演者ともに非常に熱心で、また高度かつ広範囲な情報が一挙に提供された充実した 2 日間であった。参加者の中にリピーターの方も多しとお聞きしたが、自分自身も是非次回も参加したいと思う。また、このような機会は稀であると思うので、もっとこの輪を広げ、現状に変化を起こしていくことをより多くの先生方と考えていきたいと感じた。



参考文献

- Spolsky, B. (1978). *Approaches to Language Testing: Advances in Language Testing Series, vol. 2*. Center for Applied Linguistics.
- Oller, J. (1979). *Language Tests at School*. Longman.
- Canale, M. and Swain, M. (1980). 'Theoretical bases of communicative approaches to second language teaching and testing.' *Applied Linguistics* vol.1, pp.1-47.
- Bachman and Palmer (1996). *Language Testing in Practice*. Oxford University Press.
- Cohen, A. (1994). *Assessing Language Ability in the Classroom*. Heinle and Heinle.
- Davidson, F. & Lynch, B. K. (2002). *Testcraft: A Teacher's Guide to Writing and Using Language Test Specifications*. Yale University Press.
- Popham, W. J. (1981). *Modern Educational Measurement*. Prentice-Hall.
- 安藤昭一編著. (1979). 『読む英語』. (研究社出版).

報告者：土平泰子 (茨城大学)

@@

Information

JLTA 第 16 回研究例会

JLTA 第 16 回研究例会は、2002 年 11 月 30 日(土) 午後 1 時～5 時まで愛知学院大学日進キャンパスの学院会館で開催されます。内容は、従来どおり、研究発表 20 分、質疑応答 10 分です。現在までのところ、数名の方から発表希望の申し込みがありました。最終的に何件の発表が行われるかは、まだ未定ですが、できるだけ多くの発表を設定し、多くの皆様が参加されますよう、お待ちしております。

会場は、<http://www.aichi-gakuin.ac.jp/> のアクセスガイドをご覧ください。

問い合わせ先：愛知学院大学 伊藤彰浩
(akito@dpc.aichi-gakuin.ac.jp)

大学入試センター主催国際シンポジウム 「問題作成からみる大学入試」 開催のお知らせ

大学入試センター主催で、本年 11 月 16 日(土)、17 日(日)、目黒区駒場 NTT DATA 駒場研修センターにおいて、日、中、韓、英、米の 10 人の試験問題作成の専門家(5 名は英語、5 名は数学)によるシンポジウムを開催します。参加申し込み締め切りは 10 月 10 日。

詳細と申し込み方法につきましては www.rd.dnc.ac.jp/~istd/ をご覧下さい。

事務局より

◎ 2002 年度の会員名簿を発行いたしました。ご確認いただき、訂正等が必要な場合は、直ちに事務局までご連絡ください。

◎ 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。銀行引き落としが不可になった場合、その年度のみ、郵便振込による納入をお願いしています。ご理解の上、ご協力をお願いいたします。

◎ 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。

Post script from Editor

“The *test* of a vocation is the love of the drudgery it involves.” (Logan Pearsall Smith, *Afterthoughts*, 1931) Well, we’re always being tested then. But is the test really valid, is it really reliable, is it really fair? (Wt).

日本言語テスト学会事務局

〒389-0813 長野県埴科郡戸倉町芝原 758

TEL 026-275-1964 FAX 026-275-1970

e-mail: youichi@avis.ne.jp

URL: <http://www.avis.ne.jp/~youichi/JLTA.html>

