

日本言語テスト学会 (JLTA)
第 16 回 (2012 年度) 全国研究大会発表要綱

Handbook
of
the 16th Annual Conference
of
the Japan Language Testing Association

Conference Theme: Evaluating extensive reading

日時：2012年10月27日(土) 8:30-17:45

会場：専修大学生田キャンパス 10号館2階
〒214-0033 神奈川県川崎市多摩区東三田2-1-1 044-911-0504 (代表)

主催：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



The Japan Language Testing Association

758, Wakamiya, Chikuma, Nagano 389-0813, Japan
Telephone: +81-26-275-1964 Fax: +81-26-275-1970

<https://e-learning-service.net/jlta.ac/>

E-mail: youichi@avis.ne.jp

Conference Venue:

Senshu University, Ikuta Campus

2-chome-1-1, Higashi-Mita, Tama-ku, Kawasaki-shi, Kanagawa
214-0033, Japan

Telephone: +81-44-911-0504

<http://www.senshu-u.ac.jp/english/campus/ikuta/index.html>

全 国 研 究 大 会 本 部 委 員

中村 優治	(慶應義塾大学)
中村 洋一	(清泉女学院短期大学)
小泉 利恵	(順天堂大学)
法月 健	(静岡産業大学)
片桐一彦	(専修大学)

全 国 研 究 大 会 運 営 委 員

澤木 泰代	(早稲田大学)
印南 洋	(芝浦工業大学)
木村 哲夫	(新潟青陵大学)
島谷 浩	(熊本大学)
島田 勝正	(桃山学院大学)
清水 公男	(木更津工業高等専門学校)
Soo-im Lee	(龍谷大学)
谷 誠司	(常葉学園大学)
藤田 智子	(東海大学)
金子 恵美子	(会津大学)
井上 千尋	(旭川医科大学)

全 国 研 究 大 会 実 行 委 員

委員長	
片桐 一彦	(専修大学)
委員	
上村 妙子	(専修大学)
神白 哲史	(専修大学)
滝澤 武	(専修大学)
印南 洋	(芝浦工業大学)

研 究 発 表 審 査 委 員

委員長	
片桐 一彦	(専修大学)
委員	
澤木 泰代	(早稲田大学)
島田 勝正	(桃山学院大学)
印南 洋	(芝浦工業大学)

Table of Contents (目次)

1. Overall Conference Schedule (大会スケジュール)	4
2. Conference Committee Information (学会事務局からのお知らせ)	10
3. Abstracts (発表要旨)	13
4. Workshop Information (ワークショップ情報)	42
5. Access to the Conference Venue (会場へのアクセス)	43
6. Advertisements and Commercial Exhibits (協賛企業)	48

Overall Conference Schedule: Timetable

October 26, 2012 (Friday)

14:00—17:00	Pre-Conference Workshop “Confirmatory Factor Analysis” Yo IN’NAMI (Shibaura Institute of Technology) (conducted in Japanese; Room 101B, 1st floor, Building 10, Ikuta Campus, Senshu University)
16:30—18:30	Board Meeting (Meeting room, 2nd floor, Hotel Sunroute Plaza Shinjuku.)

October 27, 2012 (Saturday), 2nd floor, Building 10, Ikuta Campus, Senshu University

8:30—	Registration (Hall in front of Room 10204)
9:00—9:10	Opening Ceremony (Room 10202)
9:15—9:45	Report from the Web Publication Committee (Room 10202)
9:55—10:55	Keynote Speech (Room 10202)
11:05—11:35	Presentation I
11:40—12:10	Presentation II
12:10—13:20	Lunch Break (JLTA Committee Meetings: Room 10214)
13:20—13:50	Presentation III
13:55—14:25	Presentation IV
14:30—15:00	Presentation V
15:10—16:40	Symposium (Room 10202)
16:45—17:05	JLTA General Business Meeting (Room 10202)
17:05—17:15	Closing Ceremony (Room 10202)
17:30—19:20	Banquet (Shidax, 4th floor, Building 10)

Commercial Exhibits: Hall in front of Room 10204 (Drinks are available for free.)

Lunch Room for Participants & Participants’ Lounge: Room 10213 (Please only use this room for eating lunch.)

Headquarters: Room 10215

大会日程表

2012年10月26日(金)

14:00—17:00	ワークショップ「確認的因子分析」(Confirmatory Factor Analysis) 講師：印南 洋 (芝浦工業大学) (専修大学生田キャンパス 10号館 1階 101B)
16:30—18:30	理事会 (ホテル・サンルート・プラザ新宿 2階会議室)

2012年10月27日(土) 専修大学生田キャンパス 10号館 2階

8:30—	受付 (10204 教室前ホール)
9:00—9:10	開会行事 (10202 教室)
9:15—9:45	Web 公開委員会報告 (10202 教室)
9:55—10:55	基調講演 (10202 教室)
11:05—11:35	研究発表 I
11:40—12:10	研究発表 II
12:10—13:20	昼食 (役員会：10214 教室)
13:20—13:50	研究発表 III
13:55—14:25	研究発表 IV
14:30—15:00	研究発表 V
15:10—16:40	シンポジウム (10202 教室)
16:45—17:05	JLTA 総会 (10202 教室)
17:05—17:15	閉会行事 (10202 教室)
17:30—19:20	懇親会 (10号館 4階 シダックス)

協賛企業展示： 10204 教室前ホール (無料の飲み物がございます)

一般参加者昼食・休憩室： 10213 教室 (昼食は、この部屋のみでお願いいたします)

大会本部： 10215 教室

Program of the 16th JLTA Annual Conference

October 27, 2012 (Saturday)

- 8:30— Registration (Hall in front of Room 10204, 2nd floor, Building 10)
Conference Attendance Fee: JLTA Members: ¥1,000;
Non-members: ¥3,000 (except Students: ¥1,000)
- 9:00—9:10 Opening Ceremony (Room 10202, 2nd floor, Building 10)
Coordinator: Yasuyo SAWAKI (Waseda University)
Greetings: Yuji NAKAMURA (JLTA President; Keio University)
- 9:15—9:45 Report from the Web Publication Committee (Room 10202)
Coordinator: Yasuyo SAWAKI (Waseda University)
Presenters: Yukie KOYAMA (Nagoya Institute of Technology), Randy THRASHER (Okinawa Christian University), Youichi NAKAMURA (Seisen Jogakuin College), Minoru AKIYAMA (e-Learning Service), Yasuhiro IMAO (Osaka University), Rie KOIZUMI (Juntendo University), Ken NORIZUKI (Shizuoka Sangyo University)
Theme: The Role of the Web Publication Committee in Language Testing in Japan
- 9:55—10:55 Keynote Speech (Room 10202)
Coordinator: Yoshinori WATANABE (Sophia University)
Introduction of the lecturer: Yuji NAKAMURA (JLTA President; Keio University)
Theme: Evaluating extensive reading
Lecturer: David BEGLAR (Temple University Japan Campus)
- 11:05—12:10 Presentations I & II (Presentation: 20 mins; Discussion: 10 mins)
- 12:10—13:20 Lunch
(JLTA Committee Meetings: Room 10214; Lunch Room for Participants: Room 10213)
- 13:20—15:00 Presentations III to V (Presentation 20 mins; Discussion 10 mins)
- 15:10—16:40 Symposium (Room 10202)
Theme: Evaluating extensive reading at school
Coordinator: Yo IN'NAMI (Shibaura Institute of Technology)
Panelist: Atsuko TAKASE (Kansai University)
Panelist: Rob WARING (Notre Dame Seishin University)
Discussant: David BEGLAR (Temple University Japan Campus)
- 16:45—17:05 JLTA General Business Meeting (Room 10202)
Selection of the chair
Reporter: Rie KOIZUMI (JLTA Vice Secretary General; Juntendo University)
- 17:05—17:15 Closing Ceremony (Room 10202)
- 17:30—19:20 Banquet (Shidax, 4th floor)
Coordinator: Yasuyo SAWAKI (Waseda University)

日本語テスト学会第15回全国大会プログラム

2012年10月27日(土)

- 8:30— 受付 (10号館2階10204教室前ホール)
学会参加費: JLTA 会員 1,000円、非会員 3,000円 (ただし学生は 1,000円)
- 9:00—9:10 **開会行事** (10202教室)
総合司会 澤木 泰代 (早稲田大学)
挨拶 中村 優治 (JLTA 会長・慶應義塾大学)
- 9:15—9:45 **Web 公開委員会報告** (10202教室)
司会 澤木 泰代 (早稲田大学)
発表者 小山 由紀江 (名古屋工業大学)、Randy THRASHER (沖縄キリスト
教学院大学)、中村 洋一 (清泉女学院短期大学)、秋山 實 (株式会社eラー
ニングサービス)、今尾 康裕 (大阪大学)、小泉 利恵 (順天堂大学)、法月 健
(静岡産業大学)
題目 The Role of the Web Publication Committee in Language Testing in Japan
- 9:55—10:55 **基調講演** (10202教室)
司会 渡部 良典 (上智大学)
紹介 中村 優治 (JLTA 会長・慶應義塾大学)
演題 Evaluating extensive reading
講師 David BEGLAR (Temple University Japan Campus)
- 11:05—12:10 **研究発表 I・II** (発表 20分, 質疑応答 10分)
- 12:10—13:20 **昼 食**
(役員会: 10214教室 一般参加者昼食控室: 10213教室)
- 13:20—15:00 **研究発表 III—V** (発表 20分, 質疑応答 10分)
- 15:10—16:40 **シンポジウム** (10202教室)
テーマ Evaluating extensive reading at school (使用言語: 英語)
コーディネーター 印南 洋 (芝浦工業大学)
パネリスト 高瀬 敦子 (関西大学)
パネリスト Rob WARING (Notre Dame Seishin University)
討論者 David BEGLAR (Temple University Japan Campus)
- 16:45—17:05 **JLTA 総会** (10202教室)
議長選出
報告 小泉 利恵 (JLTA 事務局次長・順天堂大学)
- 17:05—17:15 **閉会行事** (10202教室)
- 17:30—19:20 **懇親会** (10号館4階 シダックス)
司会 澤木 泰代 (早稲田大学)

Presentation Overview

Time	Part	Room 10202	The 1st room (10205)	The 2nd room (10206)	The 3rd room (10207)	The 4th room (10208)	The 5th room (10209)
9:15– 9:45		Web Publication Committee		--	--	--	--
9:55– 10:55		Keynote speech		--	--	--	--
11:05 – 11:35	I	--	PELLOWE & HOLSTER	CHO	NAKA- MURA	今井・石塚・ 山田・篠崎・ 西村・酒井・ 菊地	木村
11:40 – 12:10	II	--	YOSHI- ZAWA, TAKASE, & OTSUKI	INOUE	SHIM	矢野	伊東・川口・ 太田・佐々木
12:10 – 13:20							
13:20 – 13:50	III	--	HOLSTER, PELLOWE, & LAKE	SATO	Cancelled AHMADI	徳永	谷
13:55 – 14:25	IV	--	BLAKE	--	島谷	宮崎	高木
14:30 – 15:00	V	--	GROGAN	SAITO	秋山・松下	長沼・工藤・ 根岸	静
15:10 – 16:40		Symposium	--	--	--	--	--

Presentation Details

Room 10202

Chair Report from the Web Publication Committee Yasuyo SAWAKI (Waseda University)
 Keynote speech Yoshinori WATANABE (Sophia University)
 Symposium summary Takayuki NAKANISHI (Tokiwa University)

Part	Presenter (Affiliation)	Title (Page)
	Report from the Web Publication Committee Yukie KOYAMA (Nagoya Institute of Technology), Randy THRASHER (Okinawa Christian University), Youichi NAKAMURA (Seisen Jogakuin College), Minoru AKIYAMA (e Learning Service), Yasuhiro IMAO (Osaka University), Rie KOIZUMI (Juntendo University), Ken NORIZUKI (Shizuoka Sangyo University)	The Role of the Web Publication Committee in Language Testing in Japan (p. 17)
	Keynote speech David BEGLAR (Temple University Japan Campus)	Evaluating extensive reading (p. 13)
I–V	--	--
	Symposium Coordinator: Yo IN'NAMI (Shibaura Institute of Technology) Panelist: Atsuko TAKASE (Kansai University) Panelist: Rob WARING (Notre Dame Seishin University) Discussant: David BEGLAR (Temple University Japan Campus)	Evaluating extensive reading at school (pp. 14-16)

The 1st Room (Room 10205)

Chair	Part I	Hidetoshi SAITO (Ibaraki University)
	Part II	Hideki IIMURA (Tokiwa University)
	Part III	Yuko SHIMIZU (Ritsumeikan University)
	Part IV	Ken YANO (Hitachi Daini Senior High School)
	Part V	J. LAKE (Fukuoka Women's University)

Part	Presenter (Affiliation)	Title (Page)
I	Bill PELLOWE (Kinki University), Trevor HOLSTER (Fukuoka Women's University)	Mobile System to Monitor Extensive Reading (p. 18)
II	Kiyomi YOSHIZAWA (Kansai University), Atsuko TAKASE (Kansai University), Kyoko OTSUKI (Kansai University)	Can we treat the EPER Form A and Form E as alternate forms? (p. 19)
III	Trevor HOLSTER (Fukuoka Women's University), Bill PELLOWE (Kinki University), J. LAKE (Fukuoka Women's University)	Learning by Assessing In Second Language Writing (p. 20)
IV	John BLAKE (Japan Advanced Institute of Science and Technology)	Gems of English: A tale of triangulated testing (p. 21)
V	Myles GROGAN (Momoyama Gakuin University)	A testing environment: The experiential side of testing (p. 22)

The 2nd Room (Room 10206)

Chair	Part I	Takanori SATO (Graduate School, University of Melbourne)
	Part II	Emiko KANEKO (University of Aizu)
	Part III	Chihiro INOUE (Asahikawa Medical University)
	Part V	Makoto FUKAZAWA (Ibaraki University)

Part	Presenter (Affiliation)	Title (Page)
I	Dong Wan CHO (Pohang University of Science and Technology, Korea)	Is a letter grading system appropriate for language classes? (p. 23)
II	Chihiro INOUE (Asahikawa Medical University)	Comparing Spoken Performances at Different Levels on the Same Tasks: What Linguistic Features do Japanese Learners of English Demonstrate at A2, B1 and B2 and Above? (p. 24)
III	Takanori SATO (Graduate School, University of Melbourne)	Folk Judgments of Second Language Oral Communication Ability: A Preliminary Study (p. 25)
IV	--	--
V	Hidetoshi SAITO (Ibaraki University)	What Makes Peer Comments More "Formative" in Speaking Activities? (p. 26)

The 3rd Room (Room 10207)

Chair	Part I	Shuichi TAKAKI (Graduate School, University of Tsukuba)
	Part II	Tetsuhito SHIZUKA (Daito Bunka University)
	Part III	Keita NAKAMURA (The Society for Testing English Proficiency)
	Part IV	小山由紀江 (名古屋工業大学)
	Part V	島谷浩 (熊本大学)

Part	Presenter (Affiliation)	Title (Page)
I	Keita NAKAMURA (The Society for Testing English Proficiency)	Examining the structural validity of the achievement test of English lessons at public elementary schools in Japan (p. 27)
II	KyuNam SHIM (Cheongju National University of Education, Korea)	The National Achievement Test of Primary English in Korea: Two different perceptions (p. 28)
III	Cancelled [Alireza AHMADI (Shiraz University, Iran)]	Cancelled [Language testing competence in the Iranian National University Entrance Examination (p. 29)]

The 3rd Room (Room 10207)

IV	島谷浩 (熊本大学)	英語外部テストを利用した単位認定の波及効果：英語履修免除学生の英語学習状況と満足度 (p. 30)
V	秋山實 (東北大学大学院生)・松下達彦 (東京大学)	潜在ランク理論に基づくコンピュータ適応型テストシステムの開発と日本語語彙サイズテストへの適用—シミュレーションによる評価— (p. 31)

第4室 (10208 教室)

Chair	Part I	藤田智子 (東海大学)
	Part II	宮碕啓 (慶應義塾高等学校)
	Part III	木下正義 (福岡大学)
	Part IV	松本佳穂子 (東海大学)
	Part V	今井新悟 (筑波大学)

Part	Presenter (Affiliation)	Title (Page)
I	今井新悟 (筑波大学)・石塚賢吉 (筑波大学大学院生)・山田武志 (筑波大学)・篠崎隆宏 (千葉大学)・西村竜一 (和歌山大学)・酒井たか子 (筑波大学)・菊地賢一 (東邦大学)	コンピュータによる日本語スピーキング自動採点システム—日本語教師による採点との相関による検証— (p. 32)
II	矢野賢 (茨城県立日立第二高等学校)	統合タスクにおける文章と質問内容との構造的類似性がスピーキングパフォーマンスに与える影響 (p. 33)
III	徳永美紀 (福岡大学)	大学生の文法用語認識度を調べるための基礎知識テスト (p. 34)
IV	宮碕啓 (慶應義塾高等学校)	物語文、叙述文エッセイの評価信頼性考察 (p. 35)
V	長沼 君主 (東京外国語大学)・工藤 洋路 (東京外国語大学)・根岸 雅史 (東京外国語大学)	日本人学習者ライティングコーパスに基づいた CEFR レベル別基準特性の分析 (p. 36)

第5室 (10209 教室)

Chair	Part I	秋山朝康 (文教大学)
	Part II	長沼君主 (東京外国語大学)
	Part III	平野絹枝 (上越教育大学)
	Part IV	島田勝正 (桃山学院大学)
	Part V	木村哲夫 (新潟青陵大学)

Part	Presenter (Affiliation)	Title (Page)
I	木村哲夫 (新潟青陵大学)	能力記述文による自己評価 (p. 37)
II	伊東田恵 (豊田工業大学)・川口恵子 (芝浦工業大学)・太田理津子 (慶應義塾大学)・佐々木裕 (豊田工業大学)	ビジネス分野の can-do statements による自己評価の分析—言語能力と業務経験の関係から— (p. 38)
III	谷誠司 (常葉学園大学)	CEFR 読解尺度に基づいた日本語読解テスト開発の試み (p. 39)
IV	高木修一 (筑波大学大学院生・日本学術振興会特別研究員)	英語教科書におけるリーディングの内容理解確認問題は何を測っているのか (p. 40)
V	静哲人 (大東文化大学・VELC 研究会)	VELC テストによる TOEIC スコアの予測: リスニングとリーディングについて示唆されるもの (p. 41)

Conference Committee Information

To All Participants

University Parking Lots

University car parks are not available for this conference. Please use public transportation to come to the venue.

Registration

1. The conference registration site is located at the hall on the 2nd floor of Building 10.
2. The conference attendance fee is ¥1,000 for members (including institutional members) and ¥3,000 for non-members (¥1,000 for non-member students). If non-members apply for membership at the registration desk, the conference attendance fee will be ¥1,000. The JLTA annual fee is ¥8,000 for a general member and ¥5,000 for a student member; the admission fee for the JLTA membership is ¥1,000.
3. Please wear your conference name card throughout the conference.
4. The banquet fee is ¥4,000. The banquet registration is conducted at the the registration desk. The banquet will be held on the 4th floor of Building 10.
5. The conference handbook is available at the registration desk to JLTA non-members.

Lunch and Participants' Lounge

1. Please use Room 10213 for eating lunch as the Participants' Lounge. Lunch can be purchased at a nearby convenience store (seven-minute walk from Building 10).
2. The following three cafeterias are open on the conference day (seven-minute walk from Building 10). See the Campus Map below for the locations.

Ikuta Hall	1st floor	Suehiro	Saturday 9:00 to 14:00
Ikuta Dining Hall	2nd floor	Restaurant VIEW	Saturday 8:30 to 14:00
Building 9	5th floor	CABIN	Saturday 10:00 to 14:00

3. Complimentary drinks are available at the hall in front of Room 10204.
4. No smoking is permitted on campus.

Accommodation

We are afraid that we provide no accommodation services through our association. Please make arrangements by yourself.

Emergency Contact E-Mail Address: rie-koizumi@mwa.biglobe.ne.jp (Rie Koizumi)

Received e-mail messages will be automatically forwarded to her mobile phone.

To Presenters

1. Presenters will have 20 minutes to present their paper, followed by 10 minutes for discussion.
2. Please register at the registration desk first. Please go to the designated room 10 minutes prior to the starting time of the presentation.
3. If you are not a member, please pay the ¥3,000 "Presentation fee" (different from "Attendance fee") at the registration desk. This rule applies to every presenter on the program.
4. You are expected to connect your computer to the projector and operate it yourself. The projector and connector cable are in the room. There is sound system and you can play sounds from your computer. LAN internet access is NOT available.
5. Please bring your handouts in case the PC or the projector does not work.
6. If you need a letter of invitation, contact Rie Koizumi at rie-koizumi@mwa.biglobe.ne.jp

To Chairs

1. One chair is assigned to each presentation.
2. Please make sure that the presentation does not exceed the allotted time.
3. Please start the presentation at the time designated in the program. Please do not change the starting time or the order of the presentations.

学会事務局からのお知らせ

大会参加者へのご案内

■駐車場

学会用に駐車場は準備しておりません。公共の交通手段をお使いください。

■受付

1. 10号館2階の10204教室前ホールで行います。
2. 学会参加費は、会員1,000円（個人・賛助会員を含む）、非会員3,000円（ただし学生は1,000円）です。非会員の方でも、受付で入会手続きを行えば学会参加費は1,000円となります。JLTA年会費は一般会員は8,000円、学生会員は5,000円、入会費は1,000円です。
3. 学会中は、名札をお付けください。
4. 懇親会費は4,000円です。受付でお支払いください。懇親会は10号館4階 シダックスにて開かれます。
5. 非会員の方には、『JLTA 第16回（2012年度）全国研究大会発表要綱』を受付で配布します。

■昼食・休憩室

1. 昼食は受付では販売いたしません。ご持参ください。もしくは会場から徒歩7分の場所にコンビニエンスストア（セブンイレブン）があります。昼食・休憩室として、10213教室をご利用ください。
2. 大会当日、学生食堂3店が営業予定です。徒歩7分程度の場所にあります。キャンパス・マップをご参照ください。

生田会館	1階	スエヒロ	土曜 9:00-14:00
食堂館	2階	レストラン VIEW	土曜 8:30-14:00
9号館	5階	CABIN	土曜 10:00-14:00

3. 無料の飲み物は10204教室前ホールにございます。
4. 大学内はすべて禁煙です。

■宿泊

宿泊の斡旋はいたしておりません。

■緊急連絡先のEメールアドレス rie-koizumi@mwa.biglobe.ne.jp（小泉利恵）
携帯電話のEメールアドレスに転送されます。

発表者へのご案内

1. 20分の発表と10分の質疑応答の時間があります。
2. 受付を済ませ、発表開始10分前には、発表会場にお越しください。
3. 非会員の方は、「研究発表費(参加費とは別)」の3,000円を、受付時に支払いをお願いいたします。これは、プログラムに掲載する共同研究者すべてに適用されます。
4. ご発表にプロジェクタを使われる方は、PCとの接続作業は各自で行ってください。発表会場にはプロジェクタとともに接続ケーブルもごさいます。PCからの音声を会場のスピーカーから出すことができます。インターネットへの直接接続はできません。
5. 予測できない不具合に備え、ハンドアウトのご持参をお勧めします。
6. 出張依頼状などが必要な方は、rie-koizumi@mwa.biglobe.ne.jp (小泉利恵) までご連絡ください。

司会の方へのご案内

1. 1発表につき1人の司会をお願いしています。
2. 時間オーバーにならないようにお気をつけください。
3. 決められた時間に発表を始めてください。発表時間や順番を変えないでください。

Abstracts

Keynote Speech (Room 10202)

Evaluating extensive reading

David Beglar (Temple University Japan Campus)
dbeglar@mac.com

Approaches to evaluating extensive reading should support and promote the fundamental aims of extensive reading programs, which I take to be the following: Students will engage in large amounts of pleasurable reading that will produce cognitive, affective, and social benefits. Stated differently, the goal of extensive reading programs should be to create students who are competent readers, who enjoy reading, and who can effectively communicate with others about what they are reading. If approaches to assessment are to support these goals, they will predominantly be alternatives to traditional assessment, and they will involve students in using the other three major language skills of listening, speaking, and writing. They will also take advantage of behavior students engage in naturally outside of class, as this will make it more likely that they will engage in and benefit from the assessment tasks. When assessment is viewed in this light, it becomes possible to outline general principles for assessing extensive reading that are designed to promote students' cognitive, affective, and social development. Although the principles might be implemented quite differently in different classrooms, the fundamental goals of the assessment tasks will be similar, even in different educational contexts. The purpose of this presentation is to outline and briefly discuss a set of assessment principles that are designed to contribute to the achievement of cognitive, affective, and social goals.

Bio

David Beglar has used extensive reading for over 20 years in a variety of contexts in Japan. He is currently involved in research investigating the effects of extensive reading on the development of reading fluency and approaches to extensive reading that yield substantial fluency benefits.

Symposium (Room 10202)

Evaluating extensive reading at school

Coordinator	Yo In'nami (Shibaura Institute of Technology) innami@mre.biglobe.ne.jp
Panelist	Atsuko Takase (Kansai University) atsukot@jttk.zaq.ne.jp Rob Waring (Notre Dame Seishin University) waring_robert@yahoo.com
Discussant	David Beglar (Temple University Japan Campus) dbeglar@mac.com

Introduction

Coordinator: Yo In'nami (Shibaura Institute of Technology)

Extensive reading is intended to increase learners' language skills (e.g., reading comprehension, reading fluency, vocabulary) by having them read large amounts of text for extended periods. Although the concept has been in existence from the 1920s, the increasing interest in it in Asia, and particularly in Japan, in recent years is noteworthy. Nevertheless, in contrast to the practice of and research into extensive reading, relatively little attention has been paid to assessment issues in extensive reading, including the aspects of reading that need to be measured and to the ways to measure them (e.g., the tests/instruments to use, the amount of time needed for measurement). Many commercial tests exist to measure language proficiency, but the degree of their usefulness for measuring gain from extensive reading requires meticulous research and discussion.

In this symposium, the topic of extensive reading from an assessment perspective is examined and illustrative examples of how extensive reading has been or can be practiced and assessed in various school settings are presented. The first paper, by Takase, provides an overview of the types of tests used to measure gains from extensive reading and focuses on the Edinburgh Project on Extensive Reading (EPER) placement test. The second paper, by Waring, explores the several ways for practitioners to conduct action research into extensive reading in a school setting. Issues in the two studies and future directions for assessing extensive reading in school contexts will be discussed.

Bio

Yo In'nami has been interested in extensive reading (particularly in terms of testing and assessment) since his undergraduate supervisor used it in class almost 15 years ago. He practices extensive reading in his own classes in universities where he teaches. Additionally, he is currently interested in exploring the effectiveness of extensive reading through reviews of the literature in this area.

Paper 1: Evaluating extensive reading program

Atsuko Takase (Kansai University)

The principle purpose of this presentation is to provide information about evaluation tools used in an extensive reading (ER) program, and the results of the tests. In the last decade, extensive reading has gained popularity both in secondary schools and universities in Japan, and the number of ER classes has increased immensely. Although this change has provided many benefits to students, many teachers are still sceptical about the effects of ER on improving learners' English proficiency mainly using easy reading materials.

In order to improve English proficiency, learners need to read an abundance of fully comprehensible materials at the beginning of ER, as this leads to the automatic recognition of vocabulary and syntax, which leads to greater reading fluency, facilitates a larger amount of reading in a shorter time, which is followed by further fluency improvements (virtuous circle).

According to Day and Bamford (1998), the purposes of evaluating an ER program are "to see whether a program has achieved its goal," to identify outcomes other than intended ones, and to identify necessary changes or improvement (p. 157). However, most sceptical teachers' primary concern is whether ER improves learners' English reading ability.

The best way to examine the effects of ER on learners' reading ability is to administer pre- and post-cloze tests (Day & Bamford, 1998). The major cloze test utilized for investigating learners' general English proficiency is the Edinburgh Project on Extensive Reading (EPER) placement tests. Among several versions, the easiest (Version A) is used most frequently in Japan. It is a semi-fixed-ratio cloze test consisting of twelve paragraphs (963 words) arranged in an ascending order of difficulty. It contains 141 blanks mainly for every 5th to 7th word. To identify the correct words in the deleted parts of a passage, the ability to understand context and vocabulary, and grammatical knowledge in context are necessary, as most of the missing words are related to grammar. Day and Bamford suggested that the pretest score should be less than 35% to see the gains on the posttest.

Other available tests are: STEP Eiken for junior high school students, SLEP (Secondary Level of English Proficiency) test, ACE (Assessment of Communicative English), the GTEC (Global Test of English Communication) for high school students, and the TOEIC for National College of Technology and university students.

Day, R. R. & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.

Bio

Atsuko Takase is interested in extensive reading and listening, ER motivation, and vocabulary acquisition through extensive reading. For over 15 years, she has practiced extensive reading in a variety of ways in universities and a high school in Osaka.

Paper 2: Extensive Reading Research in Action
Rob Waring (Notre Dame Seishin University)

Many practitioners balk at the idea of doing research into extensive reading as it can seem complex and hard to fathom given all the potential variables at play. This is often compounded by the apparent complexity of many published ER research designs which need to control multiple variables. These include knowing what to test; what reading to administer; who and how to test, which instruments to use and how to ensure there is no contamination of the data given that extensive reading is usually conducted alongside other types of language instruction, among others. This is further compounded by a myriad of assessment methods and instruments.

This presentation will suggest several ways in which practitioners can indeed conduct small-scale action research projects on their own students and classes. This might involve looking at various aspects of extensive reading including attitudes to ER; perceptions of language gains; classifying and rating reading materials; efficiency of feedback and so on. This will be followed by presenting an action research agenda and guidelines for conducting your own research in your classes and possibly getting it published.

The presentation closes with an analysis of some of the reasons why little action research is being done in extensive reading at the moment and what we can do about it, both individually and collectively.

Bio

Dr. Rob Waring is an acknowledged expert in Extensive Reading and second language vocabulary acquisition. He is on the Executive Board the Extensive Reading Foundation and is also author or series editor of 200 graded readers.

Report from the Web Publication Committee (Room 10202)

The Role of the Web Publication Committee in Language Testing in Japan

Yukie KOYAMA (Nagoya Institute of Technology)

koyama@nitech.ac.jp

Randy THRASHER (Okinawa Christian University)

Youichi NAKAMURA (Seisen Jogakuin College)

Minoru AKIYAMA (e Learning Service)

Yasuhiro IMAO (Osaka University)

Rie KOIZUMI (Juntendo University)

Ken NORIZUKI (Shizuoka Sangyo University)

After the discussion in the steering committee of 2011 JLTA annual conference, JLTA decided to form a task force, the Web Publication Committee, for the purpose of dissemination of language testing information and JLTA activities. There are specifically two areas of the dissemination mentioned above, (A) Registration of the JLTA Journal on CiNii (Citation Information by NII [National Institute of Informatics]) and (B) Publishing useful materials on the JLTA web site.

This presentation explains the tasks of each area, and reports the progress of the committee's activities.

(A) Registration of JLTA Journal on CiNii

The committee has already started asking authors of all the past JLTA Journal articles for copyright permission in order to register the material on CiNii.

(B) Publishing useful materials on the JLTA web site

In this category, the following three types are included; 1) digitalized materials useful for language testing such as booklets, slides, handouts, pictures and video clips of workshops, lectures, online tools and their manuals, 2) links to useful sites such as DIALANG, and 3) JLTA original on-line tutorials.

Work on the on-line tutorials has begun by trying to find what materials would be useful to help teachers develop a deeper understanding of testing and good testing practice by first carrying out a needs analysis. A preliminary questionnaire survey was designed to discover what teachers believe they need to know about measurement and test development. It is expected that the data from this survey will be useful in deciding what teaching modules to include in the web tutorial.

For this purpose, a 14 item questionnaire was prepared in Japanese and English and data was collected on-line from 31 subjects and from 42 others using a pencil and paper format. Respondents included elementary, junior high, and senior high school language teachers in Japan as well as university students in training to become teachers. Respondents were asked to report their level of need to understand or interest in 14 different aspects of language testing on a 4-point Likert scale.

The results indicate that teachers and teacher trainees are more interested in practical (how to test language elements) issues than they are in understanding measurement theory or the issues currently of interest to language testing specialists.

The presentation presents a brief overview of the on-line tutorial program that the JLTA task force is developing, the data from the questionnaire research, and concludes with a discussion of how the tutorial might be constructed to encourage teachers to develop an understanding of the test theory underlying the 'how to' issues they seem to be most interested in.

Paper Presentation

The 1st Room (Room 10205) Part I

Mobile System to Monitor Extensive Reading

Bill PELLOWE (Kinki University)

pellowe@fuk.kindai.ac.jp

Trevor HOLSTER (Fukuoka Women's University)

Extensive reading (ER) requires comprehension of large quantities of interesting text of appropriate difficulty, raising the question of how to monitor students' reading. This is often addressed through quizzes to check that students did read the claimed books, book reports, or through word count targets. However, quizzes and book reports may result in students engaging in intensive rather than extensive reading, while performance goals such as word counts may be harmful for motivation and thus be counterproductive. Data from a pilot study using paper survey forms to ask about book completion and book difficulty was analyzed using many-faceted Rasch measurement, finding sufficient reliability for effective measurement. However, the data collection and input from paper survey forms is too time consuming for large scale administration. In order to automate the process, the presenters developed an open-source module for the MOARS audience response system (Holster & Pellowe, 2011; Pellowe, 2010) to allow data collection using mobile devices such as an iPod Touch, mobile phone, or any other internet connected device. The MOARS system is open source and freely available. When students answer the survey questions, the data is stored in an online database. The system provides simple tables of book counts to allow teachers to monitor students' reading, and can also output a data and control file for immediate analysis using the Facets software for many-faceted Rasch measurement. Pilot introduction of the system found it to be easy to use while providing sufficient measurement reliability to inform text recommendations and future purchasing decisions.

References

- Holster, T. A., & Pellowe, W. R. (2011). Classroom peer assessment using mobile devices. Paper presented at the 37th JALT International Conference, National Olympics Memorial Center, Yoyogi, Tokyo.
- Pellowe, W. R. (2010). MOARS (Version 0.8.3) [Audience response system]. Retrieved from <http://moars.com>

The 1st Room (Room 10205) Part II

Can we treat the EPER Form A and Form E as alternate forms?

Kiyomi YOSHIKAWA (Kansai University)
yoshizaw@kansai-u.ac.jp
Atsuko TAKASE (Kansai University)
Kyoko OTSUKI (Kansai University)

Many extensive reading classes use the placement/progress test developed by the Edinburgh Project on Extensive Reading (henceforth EPER). Three forms of the placement/progress test are available: Form A, Form B, and Form E. Among them, the former is most often used. These forms have been treated as alternate or parallel forms of the EPER placement/progress test. They have been used to decide appropriate reading levels for learners and to analyze their progress. The EPER placement/progress test is basically a cloze test with different numbers of passages and items. For example, Form A consists of 12 passages with 141 items. However, the instructors who use these forms describe that EPER E is more difficult than EPER A. A conversion table is available to convert the scores on Form A or Form E to standard scores, but we cannot convert the scores on Form E to the scores on Form A or vice versa without using the standard scores. Although the conversion table is available, we feel it necessary to confirm the difficulty levels of the two forms when those forms are administered to the EFL learners for assigning appropriate reading materials and assessing their progress.

The present study has two aims. The first aim is to conduct a content analysis of the EPER Form A and Form E, including the length and readability of the passages used in each form and the types of deletions. The second aim is to examine the difficulty levels of the two forms. To this end, two forms were linked using a common-person design. A total of 200 students at two four-year universities in the western Japan participated in the study. There were three groups of students: one group took Form A, another group took Form E, and the last group took both Form A and E. Rasch analysis was conducted to analyze the data files of the two forms. Those data files were analyzed separately to confirm that they fulfilled the conditions to use the Rasch measurement model. Then, they were equated.

The presentation includes the results of the content analysis, those of equating, and educational implication.

The 1st Room (Room 10205) Part III
Learning by Assessing In Second Language Writing

Trevor HOLSTER (Fukuoka Women's University)

trevholster@gmail.com

Bill PELLOWE (Kinki University)

J. LAKE (Fukuoka Women's University)

Recent studies have used many-faceted-Rasch-measurement to investigate the potential of peer assessment in second language classrooms (Hirai, Ito, & O'ki, 2011; Mok, 2011; Saito & Fujita, 2004). However, "learning by assessing", where the assessment process results in learning by raters (Topping, 1998), has not been widely investigated. This study therefore investigated learning by assessing in second language writing by testing the research hypothesis that gains measured on a peer assessment rubric would exceed gains on a secondary rubric measuring general academic writing. A convenience sampling of 26 first year university students in an academic English program participated as part of a compulsory academic writing course. Pre and post treatment writing was rated by teachers on the rubric used for peer assessment and a general academic writing rubric, using the Facets software package. Students were exposed to the peer assessment rubric, but not to the secondary rubric. The MOARS audience response system (Holster & Pellowe, 2011; Pellowe, 2010) allowed students to rate each others' work and to access their own ratings instantaneously using internet connected mobile devices. Substantive gains were observed on the peer assessment rubric but not on the secondary rubric, providing evidence of learning by assessing. Response patterns suggested holistic rating by peer raters, casting doubt on the potential of peer assessment to provide formative feedback and supporting the view that the gains observed arose from learning by assessing. The holistic response patterns displayed by peer raters discounted peer assessment as a source of diagnostic feedback. Instead, the results suggest that the major benefit of peer assessment lies in improving peer raters' knowledge of the target construct through repeated engagement with the rubric during the rating process.

References

- Hirai, A., Ito, N., & O'ki, T. (2011). Applicability of peer assessment for classroom oral performance. *JLTA Journal*, 14, 41-59.
- Holster, T. A., & Pellowe, W. R. (2011). Classroom peer assessment using mobile devices. Paper presented at the 37th JALT International Conference, National Olympics Memorial Center, Yoyogi, Tokyo.
- Mok, J. (2011). A case study of students' perceptions of peer assessment in Hong Kong. *ELT Journal*, 65(3), 230-239. doi: 10.1093/elt/ccq062
- Pellowe, W. R. (2010). MOARS (Version 0.8.3) [Audience response system]. Retrieved from <http://moars.com>
- Saito, H., & Fujita, T. (2004). Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching Research*, 8(1), 31-54. doi: 10.1191/1362168804lr133oa
- Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276. doi: 10.3102/00346543068003249

The 1st Room (Room 10205) Part IV

Gems of English: A tale of triangulated testing

John BLAKE (Japan Advanced Institute of Science and Technology)

johnb@jaist.ac.jp

This study discusses the creation and execution of an end-of-course test for the first cohort of one hundred showroom sales staff in the world's largest jewelry store in Thailand. The showroom sales staff had just completed a seventy-hour English for Specific Purposes course held over twenty weeks. The test candidates were mainly ethnic minorities from various hill tribes in the Golden triangle. The majority spoke various Chinese dialects but there was no *lingua franca* as there were also monolingual Thai and Korean speakers. A significant number were illiterate as they had received no formal education.

The remit of the test was to provide a comprehensive evaluation of the candidates' ability to sell jewelry to English-speaking customers in the showroom in order to allocate the staff into sales teams. Those in the higher-ranked teams would be assigned the first coachload of English-speaking customers, and the lower-ranked teams would only be called on should a second coachload arrive at the same time. This meant the earning potential was substantially higher for the higher-ranked teams.

The end-of-course test consisted of elements that were proficiency tests and other sections that were pure achievement tests. This difference was ameliorated, however, by the fact that the vast majority of the candidates were absolute beginners at the start of the course although there were a few false beginners.

In consultation with the store management, a battery of tests was used with each one testing different aspects of the sales process. A knowledge-focused written test making use of visual prompts was used to assess the candidates' ability to recognize key lexical items and functional phrases. A one-minute spoken presentation of a self-selected jewelry item was used to assess persuasive presentation technique and pronunciation. Secret (but quickly discovered) customers were used to evaluate the effectiveness of the sales staff in the showroom. The sales process was observed *in situ* by the teacher, the floor manager and the staff supervisor. The teacher listened for language usage, the floor manager assessed personality and manner while the staff supervisor assessed sales technique.

The problems and solutions in creating and carrying out these tests in terms of validity and reliability will be examined from the standpoints of various stakeholders in the process.

The 1st Room (Room 10205) Part V

A testing environment: The experiential side of testing

Myles GROGAN (Momoyama Gakuin University)
mgrogan@andrew.ac.jp

When getting started in testing, there are resources for design (Davidson and Lynch, 2002; Bachman and Palmer, 1996) and how to analyze results (Brown 2005; Carr, 2011). There are fewer materials on the qualitative experience of working with testing within an institution, and how the balance of a normal working environment may affect how those resources are brought to bare.

This presentation is a qualitative look at the history of a norm-referenced placement instrument in use in a medium-sized university Language Centre. It will be presented by the person charged with maintaining the test—a contract lecturer who took on the role with limited testing experience—and will look at how the test has been analyzed, changed, and maintained. Rather than being a model or a suggestion for best practice, however, the focus will be on the experience of developing the necessary testing skills, spreading awareness of testing ideals within the institution, and blending testing with everyday teaching and institutional realities. The main object of the presentation is to try to compare the ideals presented in common texts with the learning curve and buy-in necessary to make such ideals work.

After describing the context of the test, the presenter will look at how the test has been perceived over time. The test has changed from being a resource-heavy instrument, simultaneously administered to well over 1,000 students at a sitting, to being a relatively simple-to-administer take-home or class-based test. The former and latter systems will be discussed, along with the relative merits of each position. Next, attempts to improve the test will be reviewed. Despite a lack of heavy-hitting statistical software, systems were developed to look at Item Facility, Item Differentiation, and test reliability. A brief demonstration of how Excel has been used to get such data, and a review of the recent literature and resources will be included here. Following this, the test will be described in terms of the institutional “fit”, and the buy-in the test gets from the various stakeholders. In particular, test development and quality assessment will be examined, with reference to trying to create a set of test specification or blueprints, and getting involvement from teachers and administrative stakeholders.

The audience are invited to share their own experiences during the question and answer session at the end, with a view to improving the testing experience for all involved in the process.

The 2nd Room (Room 10206) Part I

Is a letter grading system appropriate for language classes?

Dong Wan CHO (Pohang University of Science and Technology, Korea)

dongwanc@postech.ac.kr

Language courses taught at universities in Korea have usually adopted a letter grading system without much consideration of its effects on learning outcomes and psychological burden on learners. A letter grading system, which has its base on norm-referenced testing, aims to differentiate students' achievement in relation to others and to boost their motivation for higher grades. In a case, a strict proportion of each grade has been mandated by some universities, causing irrational competition among students. In contrast, an S/U grading system, which bases its rationale on criterion-referenced testing, measures to what extent students have achieved goals set up by a class and what they can and cannot do. In the field of language teaching and testing, little research has been done on the effects of the grading system on teaching and learning outcomes. This research is to show how a change in a grading system from a letter grading to an S/U grading at a university in Korea has impacted the motivation of students for learning, attitude in classes, and achievements. With the survey results from 632 students, interviews from 30 students, and an open discussion, this research recommends that an S/U grading system be more appropriate to language teaching and learning. Suggestions to make up the drawbacks of an S/U grading system will also be presented.

The 2nd Room (Room 10206) Part II

Comparing Spoken Performances at Different Levels on the Same Tasks: What Linguistic Features do Japanese Learners of English Demonstrate at A2, B1 and B2 and Above?

**Chihiro INOUE (Asahikawa Medical University)
c-inoue@asahikawa-med.ac.jp**

Since its publication in 2001, the Common European Framework of Reference (CEFR) has increasingly attracted the interest of stakeholders involved in language education and assessment. The CEFR itself is claimed to be a reference framework that can be applied to describe the proficiency levels in various languages, however, there has been an increasing demand for a language-specific set of linguistic criteria, especially for the English language, which can serve as a benchmark to illustrate the linguistic performance that learners are expected to demonstrate at each level. A leading project that has been started to respond to this situation is the English Profile Programme, and one of the main studies so far focuses on analysing the Cambridge Learner Corpus (CLC) (the results have recently been published in Hawkins and Filipovic (2012)). While the research may be quite informative, the results need to be carefully interpreted. As the CLC is a corpus which stores data from the Cambridge Suite exams, the examinees at different levels are not given the same tasks. However, it should not be forgotten that the language produced by the examinees is sensitive to the demands of the tasks, and it is potentially erroneous to draw conclusions in relation to language proficiency based on elicited language structures without knowing what structures and functions were required by the tasks.

Consequently, this study aims to examine the differences in language structures used by learners at different proficiency levels performing on the same tasks. The performances of 65 Japanese university students on two spoken monologic tasks of different difficulty were qualitatively analysed. In addition, the elicited performances were rated by 7 raters as 'Below A1' to 'C1' based on the CEFR Oral Assessment Grid on both tasks. The rater training was conducted following the procedures recommended by the Council of Europe (2009). The rating criteria used in this study were Range, Accuracy, Fluency, Coherence, Sustained Monologue and Considered Judgement (CJ), all of which were moderated with the use of MFRM. Focusing on how the learners described the same events in the pictures of the two tasks, this study manually identified the structures present at different levels of proficiency. Specifically, the analyses were based on mean length of clauses and clause-based syntactic structures, following the taxonomy presented by Hawkins and Filipovic (2012). The results will be discussed in the presentation along with the issues of task comparability and task effects.

References

- Council of Europe. (2009). Manual for relating Language Examinations to the Common European Framework of Reference for Languages (CEFR). Retrieved November 24, 2009, from http://www.coe.int/t/dg4/linguistic/Manuel1_EN.asp
- Hawkins, J. A. & Filipovic, L. (2012). *Criterial features in L2 English: Specifying the reference levels of the Common European Framework*. Cambridge: Cambridge University Press.

The 2nd Room (Room 10206) Part III

Folk Judgments of Second Language Oral Communication Ability: A Preliminary Study

Takanori SATO (University of Melbourne)

tsato@student.unimelb.edu.au

In the field of applied linguistics and language testing, language specialists, including applied linguists or language teachers, have conceptualized second language (L2) communication ability and defined the construct of general-purpose oral performance tests. Although the eventual receivers of L2 speakers' oral performance in the real-world context are those who are not trained professionals in applied linguistics or language teaching (i.e., folk), their perspectives on L2 communication ability have not been incorporated into language assessment. In this presentation, I will show and discuss some findings of a preliminary study in my doctoral research in progress, which explores the features contributing to the folk judgments of L2 oral communication ability.

Four graduate students of disciplines other than applied linguistics and TESOL (three high-proficiency non-native English speakers and one native English speaker) participated in this study. First, the participants watched the video-recorded oral performances of 10 test-takers of the national College English Test-Spoken English Test of China and indicated their impression of each test-taker's communication ability without being given any pre-determined rating criteria. Subsequently, the participants verbalized features of the performances that influenced their impressions. Afterwards, I conducted post-session semi-structured interviews and solicited their thoughts on the influential features in their impressions of the performances. I analyzed their verbal reports by (a) segmenting reports into ideas units, (b) developing a coding scheme, and (c) coding each segment.

The analysis of the verbal reports and interview data indicates that the participants paid relatively strong attention to fluency, pronunciation (accents), and content. On the other hand, linguistic resources (grammar and vocabulary) affected participants' impressions to a much lesser extent. Although there were a few comments on non-verbal behaviors, two participants underscored the role of body language in their impressions. Participant comments also mentioned features related to test-takers' demeanor or interactional behaviors. In summary, the results appear to show that participants' judgments of communication ability do not necessarily correspond to the ideal performance defined by most oral performance tests, whose primary concern is with assessing exclusively linguistic factors.

Since this is a preliminary study, the findings cannot be generalized to other contexts. However, the study tentatively suggests that communicative language ability models developed by language specialists might fail to explain folk perspectives on "good L2 communication." In this presentation, I will emphasize the importance of investigating folk real-world criteria for L2 communication ability in the field of applied linguistics and language testing.

The 2nd Room (Room 10206) Part V

What Makes Peer Comments More “Formative” in Speaking Activities?

Hidetoshi SAITO (Ibaraki University)
cldwtr@mx.ibaraki.ac.jp

The purpose of this study is to characterize peer comments among 74 third-year Japanese junior high school students that were exchanged in a series of speaking activities in a course. This is a preliminary report of a larger-scale study on the effects of formative assessment on students’ achievements.

During a year-long regular English course at a junior high school in Japan, two classes of students were involved in speaking activities. Lessons from which the data were derived were team taught by the same male Japanese teacher of English and male native speaking assistant teacher of English. During each lesson, students were divided into groups of four, two of which were engaged in a speaking activity while the remaining two observed the performance and filled out a feedback form that contained four performance items with a rating scale and an open-ended comment space. Then, the students switched their speaker-observer roles. Time permitting, the second round was implemented in the same group but with alternated pairings. All students received more than 10 opportunities to give and accept comments from peers over a 7-month period of the academic year. At the end of the year, the students looked back at these comments and responded to a questionnaire concerning the comments they received and their own learning.

The present study reports characteristics of the student peer comments and what students felt was learned from those comments. To analyze the data, exploratory text data software—SPSS Text Analytics for Surveys—was used for categorization and compilation of comment data. One characteristic is that student comments consist overtly of “praises” rather than “advice” or “critical comments.” Another characteristic is that the commenting categories are limited mainly to those of rating item categories on the same feedback sheet. What this indicates is that although those rating scale items on the sheet possibly suggested items on which to comment on, they simultaneously seemed to limit students from finding their own performance categories. Analyses of the final reflection sheet show that the majority understood the utility of the commenting practice and accept peers’ comments, though there are individuals who feel threatened and awkward about assessment and comments. The study discusses what makes student comments more formative in speaking activities and what needs to be done to facilitate learning.

The 3rd Room (Room 10207) Part I

Examining the structural validity of the achievement test of English lessons at public elementary schools in Japan

Keita NAKAMURA (The Society for Testing English Proficiency; STEP)

ke-nakamura@eiken.or.jp

This study investigates the structural validity of the newly developed achievement test for young learners of English in Japan based on the result of pilot test, which was conducted in 2011. The test was designed to measure the basic English vocabulary, phrases, and Q & A expressions learned at classrooms of public elementary schools. The test, consisting of 44 listening items, was conducted to both fifth graders ($N = 1039$) and sixth graders ($N = 988$) from 17 public elementary schools around Japan in October, 2011. The Rasch model based screening method of both item and person was employed, thus making it possible to compute a linearly transformed scaled score from person ability measures. The rationale of employing unidimensional Rasch model analysis and subsequent scoring was investigated using item-level factor analysis. Weighted least square mean and variance adjusted estimation method (WLSMV) was employed in order to conduct both exploratory and confirmatory factor analysis using Mplus. Exploratory factor analysis found plausible one-factor solution, which indicated the existence of an overall listening component in the test. Subsequent confirmatory factor analysis compared the unifactor model, the correlated factor model and the higher-order model, and found the correlated and the higher-order models were superior to the unifactor model in terms of model-data fit indices. The higher-order model best explained the structure of the test design and the Rasch model based analysis and scoring, thus making it the accepted model in this study. The result and its implication to the current theory of the structural validity in the field of language testing are also discussed.

The 3rd Room (Room 10207) Part II

The National Achievement Test of Primary English in Korea: Two different perceptions

KyuNam SHIM (Cheongju National University of Education, Korea)

knshim@cje.ac.kr

This study aims to investigate how Korean stakeholders perceive the development and implementation of the national achievement test of primary English (hereafter NATPE). This study also set out to uncover, and to obtain in-depth understanding of the issues which emerged from the author's earlier work in 2011 which had investigated washback effect of NATPE on teachers' beliefs and practices. Following a discussion of national curriculum of primary English and assessment policy, this study introduces the overall structure and implantation system of the NATPE. This study also examines a number of mainstream theories of language testing or assessment which are relevant to NATPE and considers how NATPE confirms to these theoretical principles. To carry out the research, 691 students from 9 middle schools in 3 different cities participated in the study; they took the NATPE in July 2011. 75 head teachers or deputy head teachers from 50 primary schools in the same areas took part in the study. In addition, 10 NATPE item writers joined the research. Regarding research instrument, 3 questionnaires were developed; the questionnaire for the students consists of close-ended questions, but the questionnaires for head teachers and test developers are comprised of both close-ended and open-ended questions. Analysis of the quantitative data was carried out using SPSS (15.0 for Windows) and the qualitative data were analyzed through classification and listing.

The results of the study are as follows: First, generally all the stakeholders hold their own firm beliefs regarding NATPE; two-thirds of all the participants accepted that the NATPE was worthy of implementation. However there are still considerable number of respondents who regarded the test as unnecessary. Second, the impact of the test was not very big, contrary to expectation. Also the test had more influence on what teachers do than what students do in the schools; teachers were usually heavily concerned with the test and its results, but the students were not. Third, with reference to the future direction of the test, NATPE is required to revise or improve the curriculum prior to developing the test mainly because the curriculum is so finely tuned, which makes it difficult to develop various and creative test items. In addition, the data base of test items should be constructed, through which the stakeholders can get some useful information about the testing such as topics, vocabulary lists, and sentence structures. Finally, certain areas needing further investigation were identified.

Cancelled

Language testing competence in the Iranian National University Entrance Examination

**Alireza AHMADI (Shiraz University, Iran)
arahmadi@shirazu.ac.ir**

This study aimed at investigating the construct of "language testing competence" in the Iranian National University Entrance Examination (INUEE) which is designed to screen candidates for studying at higher education as MA students of TEFL. To do so, the INUEE questions were reviewed in an interval of more than two decades to see how the construct has developed. The results indicated that only very recently the construct has been broadened to include more hot topics of the day in language testing. Some content areas, however, have maintained their status in this high-stake test for over two decades. The results further indicated that the questions have gradually evolved into more qualitative and thought-provoking questions demanding higher levels of analysis, though questions of definition and description type which were characteristic of the earlier versions of the test still exist in less frequency.

The 3rd Room (Room 10207) Part IV

英語外部テストを利用した単位認定の波及効果：英語履修免除学生の英語学習状況と満足度

島谷 浩 (熊本大学)

hishima01@ybb.ne.jp

近年、多くの高等教育機関において、外部テストとして英語能力テストが頻繁に利用されるようになった。外部テストとは、学校外のテスト専門機関によって作成されたテストで、TOEIC、英検、TOEFLなどが大学でのプレースメントや単位認定によく利用されている。しかし、英語能力テストに対する社会的評価がある程度認知されているためか、それらの妥当性や波及効果に関する研究はほとんど行われていない。

本研究は、英語外部テストによる単位認定の妥当性・波及効果を探るために、単位認定を受けて英語履修を免除された学生のその後の大学での英語学習状況を調査するものである。単位認定によって英語履修が免除された学生に対して、卒業年度の2月から3月にかけてアンケート調査を行い、43名から得られた回答を分析した。認定後の英語力を維持するための学習内容、単位認定制度利用の効果、単位認定制度への意見、大学教養英語教育への意見などをまとめて報告する。

The 3rd Room (Room 10207) Part V

潜在ランク理論に基づくコンピュータ適応型テストシステムの開発と日本語語彙サイズテストへの適用—シミュレーションによる評価—

秋山 實 (東北大学大学院生)
akiyama@ei.tohoku.ac.jp
松下 達彦 (東京大学)

潜在ランク理論 (荘島, 2007) は、アイテムバンクを構築するためのプレテストに必要な受験者データが少なくても、モデルにフィットしたアイテムパラメータをデータ数に応じた精度で推定することができる (秋山, 2012 印刷中) という特長を持っている。秋山は潜在ランク理論に基づくコンピュータ適応型テストシステム (LRT-CAT) をオープンソースソフトウェアの Moodle のモジュールとして開発した。このモジュールは、過去の受験データを用いてアイテムバンクを構成した場合のシミュレーションを行なうことができ、アイテム数が少ないアイテムバンクであっても最適な実行条件をシミュレーションによって確認し、設定することができる。また、LRT-CAT は、その処理の大半をクライアント側 (受験者のパソコン等) で実行するので、全ての処理をサーバ側で処理する従来の方式に比べサーバへの負荷が同時に受験する受験者の数に大きく左右されないという特長を持ち、プレースメントテストを一斉に受験させて実施する場合、Moodle の小テスト機能を使う場合と比べてほぼ同じ程度のサーバ側の負荷で実施できる。

松下は、日本語を読むための語彙量テスト (アイテム数 150) を開発し、プレテストによって検証した (松下, 2012 印刷中)。リニアテスト (全員が同じ問題に回答する従来のテスト) では半数程度で十分な精度で語彙サイズの推定が可能であるとの感触を得ているが、さらに受験アイテム数を少なくできれば、利便性が高まると考え、LRT-CAT による語彙サイズ推定を試みることにした。

秋山は、予め日本語語彙サイズテストの回答データを EXAMETRIKA5.3 (荘島, 2008-2011) を使用して分析し、その結果を含む EXCEL ファイルを LRT-CAT に読み込ませ、LRT-CAT のシミュレータ機能を使用してシミュレーションを行なった。

その結果を報告する。

References

Shojima, K. (2007) Neural test theory. *DNC Research Note*, 07-02.

秋山實 (2012、印刷中) . 潜在ランク理論のパラメータ推定に必要なアイテム数と受験者数. 日本テスト学会第 10 回発表論文抄録集.

松下達彦(2012、印刷中). 「日本語を読むための語彙量テスト」の開発. 2012 年日本語教育国際研究大会予稿集.

第4室 (10208 教室) Part I

コンピュータによる日本語スピーキング自動採点システムー日本語教師による採点との相関による
検証ー

今井 新悟 (筑波大学)

imai@intersc.tsukuba.ac.jp

石塚 賢吉 (筑波大学大学院生)

山田 武志 (筑波大学)

篠崎 隆宏 (千葉大学)

西村 竜一 (和歌山大学)

酒井 たか子 (筑波大学)

菊地 賢一 (東邦大学)

日本語非母語話者の「話す」能力をコンピュータで自動推定するテストを開発中であるが、コンピュータによる能力推定値を日本語教師による評価と比較することにより、信頼性を検証した。本テストは以下を経て開発されている。マルチメディアを使った問題項目の作成、WEB 上で回答させて音声データを収集するシステムの構築、約 100 問×約 200 人分の音声による回答データの収集、日本語教師による採点、音声認識精度を向上させるための言語モデル・音響モデルのチューンナップ、能力推定モデルの構築、能力推定精度向上のための特徴量の調整、能力推定値と日本語教師の採点の比較、など。本テストには、文読み上げ、選択肢読み上げ、空所補充・文生成、自由回答の形式がある。能力推定モデルに使えるような以下の特徴量の有効性について調べた。音響尤度、母語話者と受験者の発話タイミング距離、母語話者と受験者の基本周波数距離、キーワード数、語彙多様性、発話量、発話速度 1 (発話全体に対する音素数)、発話速度 2 (音声区間に対する音素数)、発話速度 3 (発話全体に対する息継ぎ無音区間の長さ)、発話速度 4 (録音時間に対する音素数)、発話速度 5 (音素内の発話速度)、発話速度 6 (単語内の発話速度)。結果、発話速度 2 と発話速度 6 は使用する必要がないこと、問題形式ごとに使用する特徴量を変えた方がよいことが判明した。以上により、問題形式ごとに使用する特徴量を変えて、サポートベクター回帰によって学習データを使ってモデルを構築した後、評価データにおいて能力推定値を産出した。その値と日本語教師による採点平均値との相関係数は、文読み上げ、選択肢読み上げ、空所補充・文生成、自由回答において、それぞれ、0.72, 0.84, 0.72, 0.88 となった。これは日本語教師間における相関に匹敵するものであり、コンピュータによる能力推定に十分な実用性があることを示している。

第4室 (10208 教室) Part II

統合タスクにおける文章と質問内容との構造的類似性がスピーキングパフォーマンスに与える影響

矢野 賢 (茨城県立日立第二高等学校)

kendrixyano@gmail.com

H25 実施の新学習指導要領 (高等学校編) では、4 技能を有機的に関連付けた指導が求められている。そのため、複数の技能を含むタスク(Integrated Task, 以下 IT) に対する適切な評価が不可欠である。IT においては情報を選択し統合するスキルが必要であると言われている(Brown et al., 2005)。複数の領域間に類似性があるときに一方の領域の知識構造をもう一方の領域に適用することをアナロジという (鈴木, 1996)。IT における選択-統合の過程で、アナロジを利用する可能性があると考えられる。

IT を含んだテストとして、英検準 2・2 級の二次面接試験がある。この試験では「読んだこと」および「聞いたこと」を踏まえた「話す」ことのパフォーマンスを評価している。本研究では、この面接試験に実際に使用された英文及び質問項目を用い、特に IT に焦点をあて、文章と質問の構造的類似性が L2 言語熟達度の異なる受験者のパフォーマンスに与える影響を明らかにすることを目的とする。

本研究では、文章理解と口頭質問に基づいたスピーキング試験の形式を採用し、文章読解、音読、内容確認質問、関連質問を順に行う。類似について対象・関係の 2 レベル(Gentner et al., 1993)に分類し 2 タイプの問題を出題する。予想される結果として、文章以外の既有知識を利用した答えと、文章と設問との類似性を利用した答えの 2 つが考えられる。RQ は 2 つ: (1)どのくらいの割合で類似性を利用して答えるのか; (2)類似性を利用した答えのうち、類似レベルの異なるトピック間においてスピーキングパフォーマンスがどのように異なるのか。この 2 点を、L2 熟達度、回答の適切さ、複雑さ、発話量と類似レベルの違いなどの関係を、相関分析を用いて検証する。

英検準 2 級面接問題集の出題例 10 問について、文章と質問項目との比較から「構造的類似あり」と「構造的類似なし」に分類。この中から 2 タイプの問題を使用し、高校生 36 名を対象に予備実験を実施した。その結果、各タイプの問題において類似性を利用した者が各 4 名。内容に混乱のあった者が各 5 名いた。「構造的類似なし」条件では、「類似性の利用」と、「内容の混乱」の相関が優位に高かった($r = .878$)。一方「構造的類似あり」条件においてはこの相関が低かった($r = -.147$)。

さらにこの結果をもとに実験を実施しており、当日はこの結果の報告を行う。

第4室 (10208 教室) Part III

大学生の文法用語認識度を調べるための基礎知識テスト

徳永 美紀 (福岡大学)

tokunagamiki@fukuoka-u.ac.jp

近年、英語の基礎力が定着しないまま大学に進学する学生が増加する中、文法用語などの認知度が低く、授業や教科書の日本語での解説が理解できないという問題も生じている。本研究は、こうした大学生の英語文法用語の理解度を調査することで、彼らが直面する問題を明らかにし、リメディアル教育などに役立てることを目標としている。

本研究で作成した基礎知識テストは、簡単な単語からなる英文の下線部の品詞や文法事項の名称を、選択肢から選ぶというものである。説明および選択肢はすべて日本語で作成し、ランダムな回答を避ける為、「わからない」という選択肢を設けた。英語を専門としない私立大学の1、2年生705名を対象にテストを実施し、留学生などを除いた657名のデータを分析した。その後TOEIC Bridge テストを受験した640名に関しては、基礎知識テストとTOEIC Bridge テストのスコアの相関関係も検証した。

Winsteps を使用してデータを分析した結果、基礎知識テストの受験者信頼性は.86、項目信頼性は.99であり、素点による平均正答数は36問中21.9問、標準偏差は7.2問であった。上記640名のTOEIC Bridge テストのスコアと基礎知識テストの受験者能力推定値を比較した結果、TOEIC Bridge のリーディングパートとの相関関係が一番大きく、 $r = .651$ であった。結果として、「slowly」が副詞と認識できた対象者が40%弱、「the」が冠詞であると認識できた対象者も60%にとどまるなど、多くの対象者の文法用語認識度が低いことが明らかになった。項目難易度の低かった「teacher = 名詞」でさえ、15% (96名)の対象者が認識できなかった。当日の発表では詳しい分析結果を紹介する。

第4室 (10208 教室) Part IV

物語文、叙述文エッセイの評価信頼性考察

宮崎 啓 (慶應義塾高等学校)

kei@hs.keio.ac.jp

エッセイライティングテストにおいて使用される談話モード (テキストタイプ) は、説得文や説明文といったアカデミックなものを目指したものが多く、リサーチの分野でもアカデミックライティングに焦点を当てた信頼性や妥当性の研究が数多く見られる。しかし実際、初級や中級学習者の指導には物語文や叙述文を使用する場合が多く、パーソナルなエッセイの調査にも注意を向ける必要がある。少ない先行研究の中では、特定の文法事項や特定の能力カテゴリーからの分析がなされているが、信頼性の考察は極めて少ない。

本発表は、物語文と叙述文のエッセイの信頼度をラッシュ分析を用いて検証し、データから評価者間の厳しさの度合いや、特定の構成要素に厳しさの度合いが見られるかどうかを分析する。中級レベル学習者 30 名の物語文、叙述文エッセイを、4名の評価者(日本人2名、ネイティブ2名)で8つの構成要素(Overall Impression, Content, Organization, Cohesion & Coherence, Creativity & Originality, Grammar, Vocabulary, Mechanics)をもとに採点して分析した。評価スケールの信頼性は高く、評価者間相関も高かったが、FACETS を用いたラッシュ分析から、評価者間や、構成要素間に厳しさの度合いの差が見られた。採点後の評価者のコメントから、ある構成要素における評価スケールの記述子解釈が評価者間で異なったり、評価者の文化的背景によりエッセイの評価が変わることがわかった。

物語文や叙述文はアカデミックライティングよりもパターン化されていないため、評価者間に差が出ることは予測できたが、商業テストや教育現場での使用頻度の高さを考えると、ある程度パターン化したスケール記述子や評価基準を設定し、信頼度の高い評価を目指すことを示唆する。

第4室 (10208 教室) Part V

日本人学習者ライティングコーパスに基づいた CEFR レベル別基準特性の分析

長沼 君主 (東京外国語大学)

n.naganuma@tufs.ac.jp

工藤 洋路 (東京外国語大学)

根岸 雅史 (東京外国語大学)

CEFR では個別言語によらない行動ベースのレベル記述がなされており、それを補完するため、各言語でレベル別のプロファイルとして語彙や文法の基準特性リストが開発されつつある。英語でも Cambridge ESOL を中心として、EnglishProfile の開発が現在進められているが、例えば、定冠詞の脱落を見ても、冠詞のある言語の母語話者ではレベルを問わず数%と、誤り率が非常に低いのに対し、冠詞のない言語では高く、日本語母語話者では、A2 で 25% を超えており、C1 や C2 になっても 10% 前後と高いままであるなど、学習者の母語の影響が見られることも分かっている (Hawkins and Buttery, 2010)。また、語彙でも、JACET8000 では 5,743 位の diploma が A2 なのに対し、1,001 位の determine が C1 であるなど、逆転も見られる。

そこで、本研究では、日本人学習者に特有の英語ライティングにおける CEFR 各レベルの基準特性を明らかにするため、E メール、物語文、論説文の 3 つのジャンルの作文をコーパス化し、学習者コーパス分析を行なった。作文課題は、EnglishProfile Programme のライティング課題を修正の上で利用し、2009 年度及び 2010 年度の東京外国語大学 1、2 年次生を対象として、2 学期始めの主・副専攻英語授業内で実施した。作文はすべてデジタル化された上で、文法タグ (UCREL CLAWS7) が付けられ、課題ごとに作成された CEFR に準拠した評価基準をもとに、2 名の評価者により採点された。

分析対象としては、のべ 5,781 名、総語数約 825,000 語のデータが用いられ、JACET8000 の語彙リストとのマッチングがされた。分析方法としては、共起ネットワーク分析、コレスポネンス分析、N グラム分析が用いられ、A2、A2+、B1、B1+、B2 のそれぞれのレベルにおける課題タイプごとの文法や語彙の基準特性が示された。しかしながら、日本人学習者に特有の基準特性の分析には、EnglishProfile の語彙リスト等の基準リストとの比較や、高頻度語彙だけでなく、低頻度語彙の分析も必要となる。また、より詳細な文法パターンの分析も必要となるであろう。2011 年度以降も継続してデータを収集しており、現在人数の少ない B2 レベルのサンプルも増やしながら、さらなる分析を行なっていきたい。

第5室 (10209 教室) Part I
能力記述文による自己評価

木村 哲夫 (新潟青陵大学)
kimura@n-seiryu.ac.jp

STEPは2003～2006年に1級から5級の合格者のべ20,000人超に対してアンケート調査を実施し、英検 Can-do リストを作成した。これには各級・技能別に総括したまとめ表現もあるが、本研究ではこのうち英検4級～準1級の4技能別の能力記述文 (Can-do statement; CDS) を学習者の自己評価のツールとして用いることとした。

2 大学の1年生対象に開講された一般教養科目の英語7クラスの受講生合計295名の協力を得て、上記CDSsを使って、入学時の英語力を自己評価してもらった。

回答データは、潜在ランク理論 (Latent Rank Theory; LRT) により、分析ランク数を5として分析された。LRTによる各CDSsの困難度をIRT指標のひとつ β (基準値0.5にもっとも近いランク)とし、そのCDSsが所属する級の順序性が一貫しているかを調べた。英検4級を1に、3級、準2級、2級、準1級をそれぞれ2, 3, 4, 5と序数化して、 β との間の相関がどの程度あるか、スピアマンの順位相関を調べたところ、いずれの技能についても.93～.95と非常に高い相関を示した。このことは、英検 Can-do リストが自己評価のツールとして十分機能することを示唆している。

CDSsによる自己評価と別途実施された英語テストによる評価がどの程度一致するか、リーディングとリスニングについて分析したところ、いずれの技能についても約60%がほぼ一致したものの、自己評価において過小評価した者は、34人(29%)対24人(17%)でリーディングの方が多く、過大評価した者は、16人(14%)対30人(22%)でリスニングの方が多かった。この割合の差は統計的にも有意な差であった ($\chi^2=5.84, df=1, p=.016$)。具体的なCDSsを示しながら、自己評価とテスト評価のズレの問題について考察を加える。

第5室 (10209 教室) Part II

ビジネス分野の can-do statements による自己評価の分析 ―言語能力と業務経験の関係から―

伊東 田恵 (豊田工業大学)

tae@toyota-ti.ac.jp

川口 恵子 (芝浦工業大学)

太田 理津子 (慶應義塾大学)

佐々木 裕 (豊田工業大学)

Can-do statements による言語運用能力評価は、日本における外国語教育・日本語教育でも広まりつつある。近年の「英語力到達目標設定」の動きを受けて、CEFR を参照した言語ポートフォリオ、カリキュラム、教材などが開発されたり、言語能力評価が行われるようになってきている。しかしながら、can-do statements による自己評価値と実際の言語能力が一致するかについては依然として議論の余地が残されている。

先行研究 (伊東・川口・太田、2008) では、業務に関連した can-do statements を作成して調査を行い、評定者が類似した業務を英語で行ったことがあるかないかという「経験」が、自己評価値にどのように影響を及ぼすかについて調査をした。結果は「経験」は概ね影響を与えるが、言語に関する「不安感」や、評定者にとっての言語タスクの「難しさ」とも関連があることが示唆された。

本研究では、先行研究で用いた英語業務の can-do statements を時代のニーズに合わせて改訂し、新たに調査を実施した。調査は客観テストの得点、can-do statements による自己評定、業務経験、必要性の有無、自由記述から構成されている。本発表では、客観テストの結果と業務の経験の有無を中心に、経験に関連した自由記述の分析を加えて、言語能力と業務経験が自己評定に与える影響を検証する。

本研究は、ビジネス分野の ESP カリキュラムを組む際のガイドラインを作成するためのプロジェクトの中間報告であり、データ分析の結果報告とともに、結果がどのようにカリキュラム作成に活用できるかも述べる。

第5室 (10209 教室) Part III

CEFR 読解尺度に基づいた日本語読解テスト開発の試み

谷 誠司 (常葉学園大学)

taniseiji@tokoha-u.ac.jp

近年、コミュニケーションや自律学習を重視する流れを受け、言語で何ができるか(Can-Do)という考えをもとに記述された言語能力記述の枠組みに関心が集まっている。その中で欧州評議会が開発した Common European Framework of Reference for Languages (以下、CEFR) は、外国語教育の参照枠としてスタートしたが、その影響力はヨーロッパの枠を超え、世界の外国語教育にインパクトを及ぼしている。日本語教育においても、CEFR に準拠した「JF 日本語教育スタンダード」(国際交流基金、2010)や「日本語口頭産出能力評価法」(OJAE: Oral Japanese Assessment Europe) が開発されている。

本研究では上記のような背景のもと、CEFR の読解尺度を土台にした成人韓国人日本語学習者用の日本語読解テストの開発を試みた。

テスト開発は予備テストと本テストの2段階に分けた。まず、予備テストでは日本語教育専攻の博士課程学生3名でテスト開発チームを組織し、CEFR への習熟化、テスト細目表の作成、項目作成を経て、予備テストを実施し、その結果から項目分析、妥当性の検証、CEFR の各レベルの難易度・困難度の範囲の設定を行った。予備テストでの結果を踏まえて、項目の差し替えを行い、本テストを実施し、項目分析、妥当性の検証、予備テストで設定した CEFR の各レベルの難易度・困難度の範囲との一致度の確認を行った。

第5室 (10209 教室) Part IV

英語教科書におけるリーディングの内容理解確認問題は何を測っているのか

高木 修一 (筑波大学大学院生・日本学術振興会特別研究員)

ka_to_n@yahoo.co.jp

英語教科書においては学習者の理解を確認するための問題が提示されているが、これらの内容理解確認問題は様々なバリエーションが存在している。例えば、同じ教科書においても、テキストに書かれている内容の真偽判断を求める問題が提示されている一方で、テキストに書かれていることに基づいて暗示的な内容の推論を求める問題が提示されている。

これらの内容理解確認問題のねらいは各教科書において定められてはいるが、それぞれの問題が学習者のどのような理解を測定しているのかについて統一的な枠組みで検討を行った研究はほとんどない (深澤, 2008, 2010)。

本研究では英語教科書におけるリーディングの内容理解問題に焦点を当てて分類を行った。その際、本研究では清水 (2005) の分類基準を採用した。清水 (2005) は読解問題を以下の 6 種類に分類している: (a) パラフレーズ質問、(b) 推論質問、(c) テーマ質問、(d) 指示質問、(e) 語彙質問及び (f) 文章構造質問。さらに、(b) 推論質問 については解答に必要な推論のタイプによって細かく下位分類 (e.g., 因果的先行詞) を行っている。テキストの深い読解にはテキストに明示されている情報に基づいた推論や背景知識の活性化が不可欠とされており、各質問に解答するために必要な学習者の理解を読解プロセスの観点から細かく分類している点において優れていると考えられる。深澤 (2008, 2010) では読解指導における望ましい発問の検討を目的として分類を行っているが、本研究では読解質問が学習者のどのような理解を測定しているのかという観点から分類を行う。そのため、暗示的な情報の推論に関してより詳細である清水 (2005) の分類基準は、本研究に適していると考えられる。

本研究では採用率の高さに基づいて教科書の選定を行い、それぞれの教科書における内容理解確認問題について上記の分類基準に従って分類した。そして、その結果に基づいて教科書間で比較を行い、各教科書の特徴について考察を行った。

第5室 (10209 教室) Part V

VELC テストによる TOEIC スコアの予測: リスニングとリーディングについて示唆されるもの

静 哲人 (大東文化大学・VELC 研究会)

shizuka320@gmail.com

VELC テストとは我々の研究会が新たに開発したリスニング 60 問、リーディング 60 問、合計 120 問、所要時間約 70 分の日本人大学生向けの熟達度テストである。リスニング、リーディングともに 3 つのパート (各 20 問) に分かれており、リスニング語彙力スコア(L1)、リスニング音声分析力スコア(L2)、リスニング内容把握力スコア(L3)、リーディング語彙力スコア(R1)、リーディング文法構文力スコア(R2)、リーディング内容把握力スコア(R3)の 6 つのパートスコア、L1 と L2 と L3 を合わせたリスニングセクションスコア、R1, R2, R3 を合わせたリーディングスコア、全てを合わせたトータルスコアを算出する。ラッシュモデルによる等化を経た複数フォームのどれを受けても同一尺度上での結果比較が可能である。

VELC テストはそれ自体として日本人大学生の英語力を可能な限り正確に測定し、それによってクラス分けや授業効果の測定に役立つことを目的とする熟達度テストだが、そのような目的ですでに広く使用されている TOEIC の結果との関連について質問が多い。そこで、本発表では TOEIC の結果との相関を検証する。VELC テストの開発段階で試行テストを受けた大学生 375 名が申告した TOEIC リスニングおよびリーディングを別々の目標変数として、VELC の L1, L2, L3, R1, R2, R3 を予測変数とした重回帰分析を行った。ステップワイズ法によって予測変数を絞り込んだところ TOEIC リスニングは L1, L2, L3 によって分散の 58% を、TOEIC リーディングは L1, L2, L3, R1, R2, R3 の全てをモデルに使用して分散の 64% を予測した。またこれらの予測スコアの合計によって TOEIC トータルを予測したところ分散の 68% を予測した。また 120 問の VELC テストを仮に 90 問に短縮した場合にも TOEIC トータルの 66% を予測できることが半明した。60 分で実施できるテストで 2 時間の TOEIC の 65% 以上を予測できるのは十分な妥当性を有すると言えよう。また TOEIC リスニングの予測式には R1, R2, R3 が含まれないが、TOEIC リーディングの予測式には L1, L2, L3 が含まれるのは、リスニングとリーディングの関係に関して示唆的である。

Workshop Information (ワークショップ情報)

題目：確認的因子分析 (Confirmatory Factor Analysis)

講師 印南洋 (芝浦工業大学)
司会 平井 明代 (筑波大学)

日時：2012年10月26日(金) 14:00-17:00 (受付 13:30-)

場所：専修大学生田キャンパス 10号館1階 101B (ゼミ室)

参加費：500円

定員：20名(申し込み順)

参加条件：SPSS・エクセルなどで相関分析を行ったことがある方が望ましい。確認的因子分析が初めての方から参加可能。

形式：

コンピュータのない教室で、講師のコンピュータを映したスクリーンを見ながら、理論と手順を学びます。ご自身のコンピュータを使って行う実習はありません。

概要：

1. 確認的因子分析の理論
2. Amos を用いた確認的因子分析方法
3. 質疑応答

申し込み：

参加希望の方は、以下の情報とともに、10月20日(土)までに小泉宛にご連絡ください。定員に達していない場合には当日参加も可能です。

早めにご質問をいただきますと、ワークショップの内容に反映しやすくなります。たくさんのお申し込みをお待ちしております。

■ワークショップの申し込み時にお教えてください。

- (1) 氏名とご所属
- (2) 以下の分析を使ったことはありますか。あるものを挙げてください。
 - (a) 相関分析
 - (b) 重回帰分析
 - (c) 探索的因子分析 (exploratory factor analysis)
 - (d) 確認的因子分析 (confirmatory factor analysis)
 - (e) その他
- (3) 何かご質問はありますか。
- (4) その他、ワークショップに対して何かご要望がありましたらお書きください。

申し込み：小泉 利恵 e-mail: rie-koizumi@mwa.biglobe.ne.jp

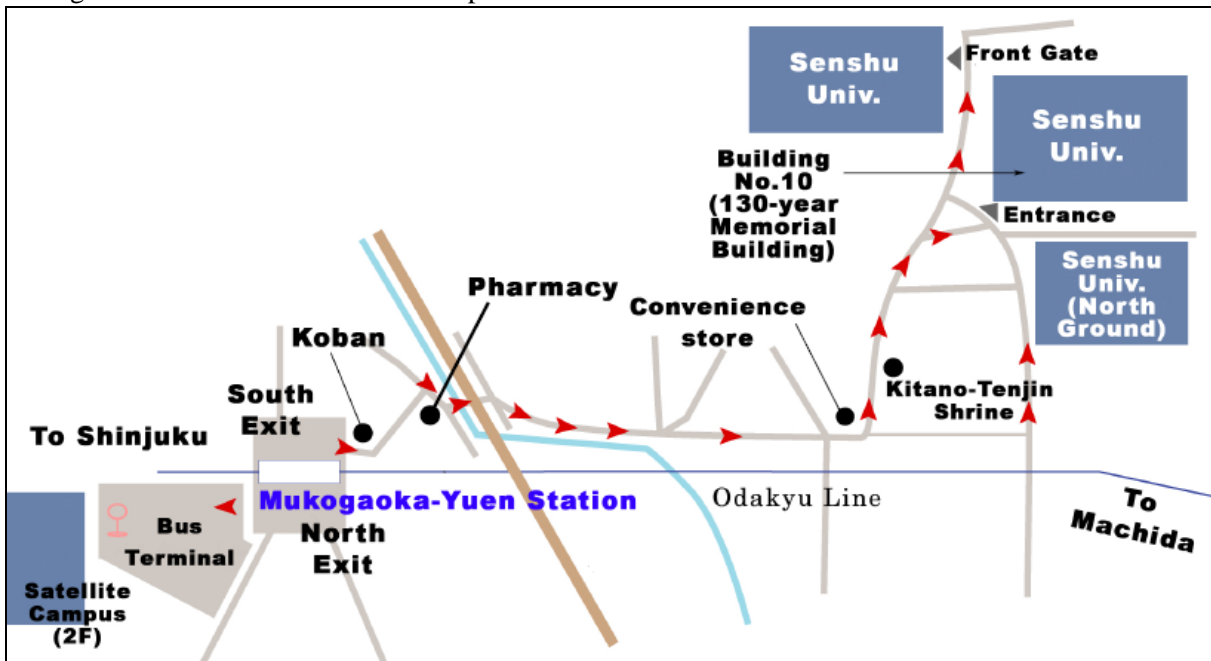
Access to the Conference Venue (Senshu University)

<http://www.senshu-u.ac.jp/univguide/profile/access.html>

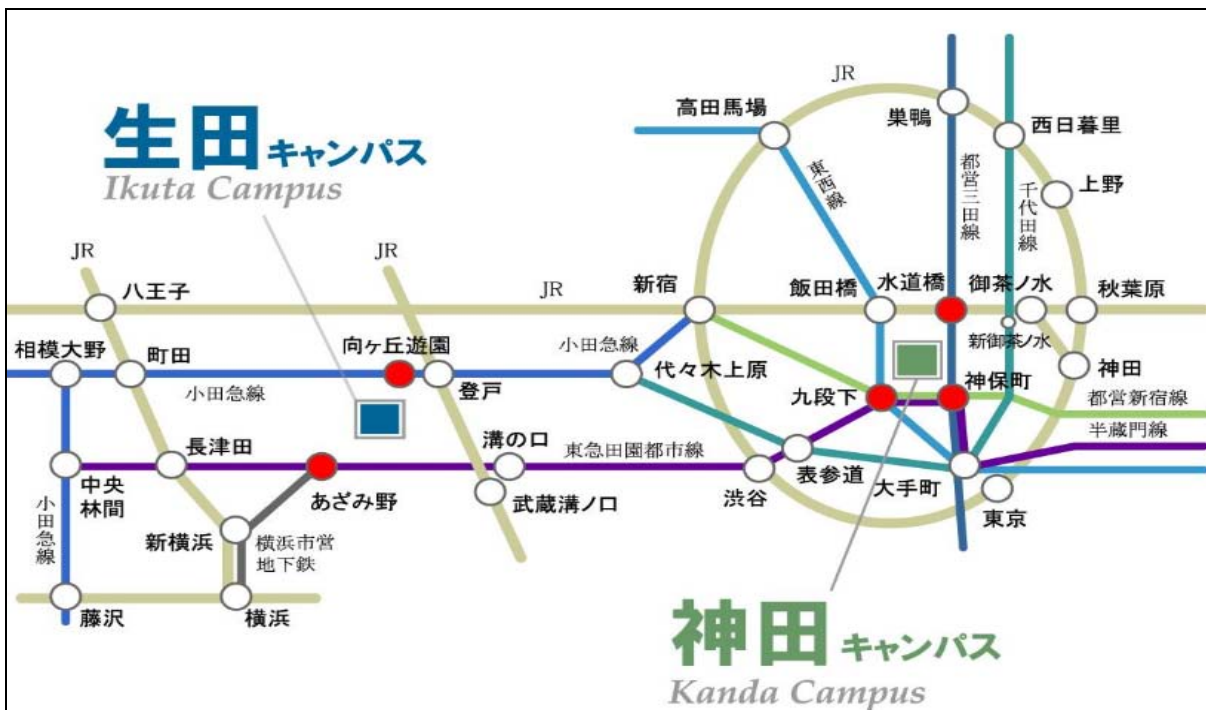
<http://www.senshu-u.ac.jp/english/campus/ikuta/campus.html> 等より引用

2-chome-1-1, Higashi-Mita, Tama-ku, Kawasaki-shi, Kanagawa

- 14 minutes' walk from the south exit of the Mukogaoka-Yuen Station on the Odakyu Line.
- 10 minutes by bus (for Senshu-daigaku-mae) from the north exit of the Mukogaoka-Yuen Station.
- 35 minutes by bus (for Mukogaoka-Yuen) from Azamino Station on the Tokyu Den-en-toshi Line or the Yokohama Municipal Subway line to Senshu Daigaku / Kawasaki Golf-jo Iriguchi Bus Stop or Senshu Daigaku 120-nen Kinenkan-mae Bus Stop.



神奈川県川崎市多摩区東三田 2-1-1



- 向ヶ丘遊園駅（小田急線）南口より徒歩 14 分（10 号館入口）。途中、坂を登ります。
- 向ヶ丘遊園駅（小田急線）北口より「専修大学前」行きバスで約 10 分
→専修大学 120 年記念館前下車（下車後、会場の 10 号館へは、トンネルを歩いて徒歩 5 分。）
- 向ヶ丘遊園駅（小田急線）北口より「聖マリアンナ医科大学」「あざみ野」行きバスで約 10 分
→専修大学 120 年記念館前下車（下車後、会場の 10 号館へは、トンネルを歩いて徒歩 5 分。）
- 向ヶ丘遊園駅（小田急線）北口より タクシー 1,000～1,500 円程度
南口より タクシー 1,300～1,800 円程度
→ 専修大学 120 年記念館（9 号館）1 階正面玄関 降車
- あざみ野駅（東急田園都市線・横浜市営地下鉄）より「向ヶ丘遊園駅」行きバスで約 35 分
→ 専修大学 120 年記念館前下車（下車後、会場の 10 号館へは、トンネルを歩いて徒歩 5 分）

※ バスの運賃はいずれも 200 円です。

※ 向ヶ丘遊園駅からのバス本数は、1 時間当たり 6～9 本です。

あざみ野駅からのバス本数は、1 時間当たり 3 本です。

※ 向ヶ丘遊園駅とあざみ野駅それぞれのバス乗り場の地図やバス時刻に関しては、
専修大学ホームページの生田キャンパスへのアクセスをご参照ください。
専大 HP トップ画面（右上）「交通案内」をクリック → 生田キャンパスへのアクセス
http://www.senshu-u.ac.jp/univguide/campus_info/campus_info_index.html

<帰りのタクシーを呼ぶ場合: Telephone number for taxi >

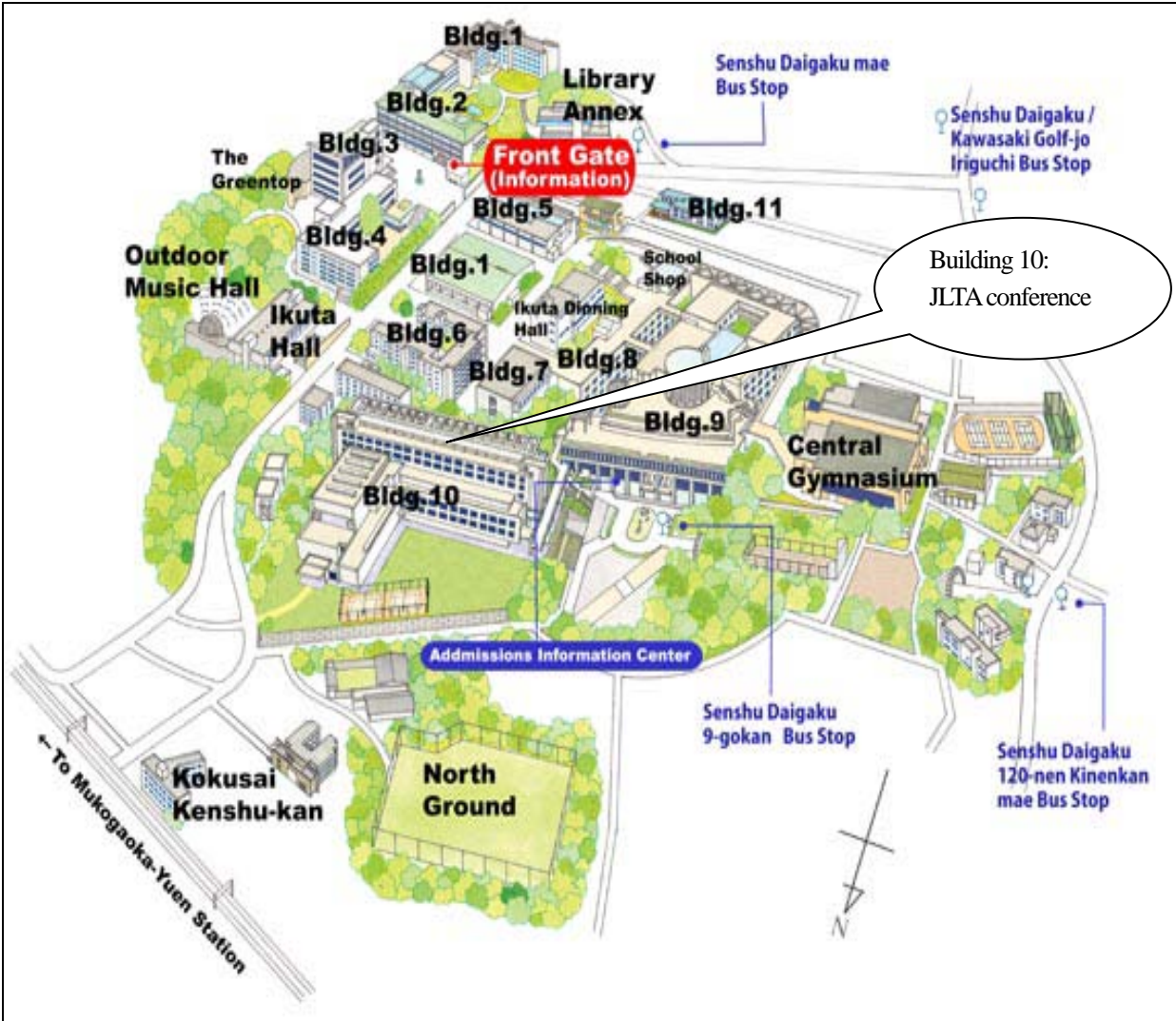
生田交通 (Ikuta Kotsu) 044-955-5388

川崎交通 (Kawasaki Kotsu) 044-933-1118

専修大学 9 号館 1 階正面玄関から乗車。

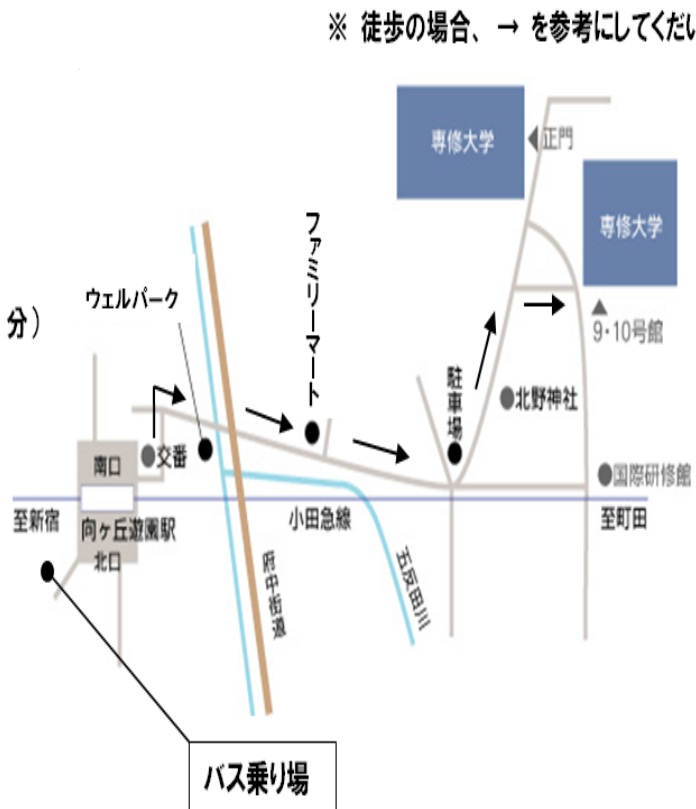
(You can get on the taxi at the Front Gate, 1st floor, Building 9, Senshu University.)

Campus map

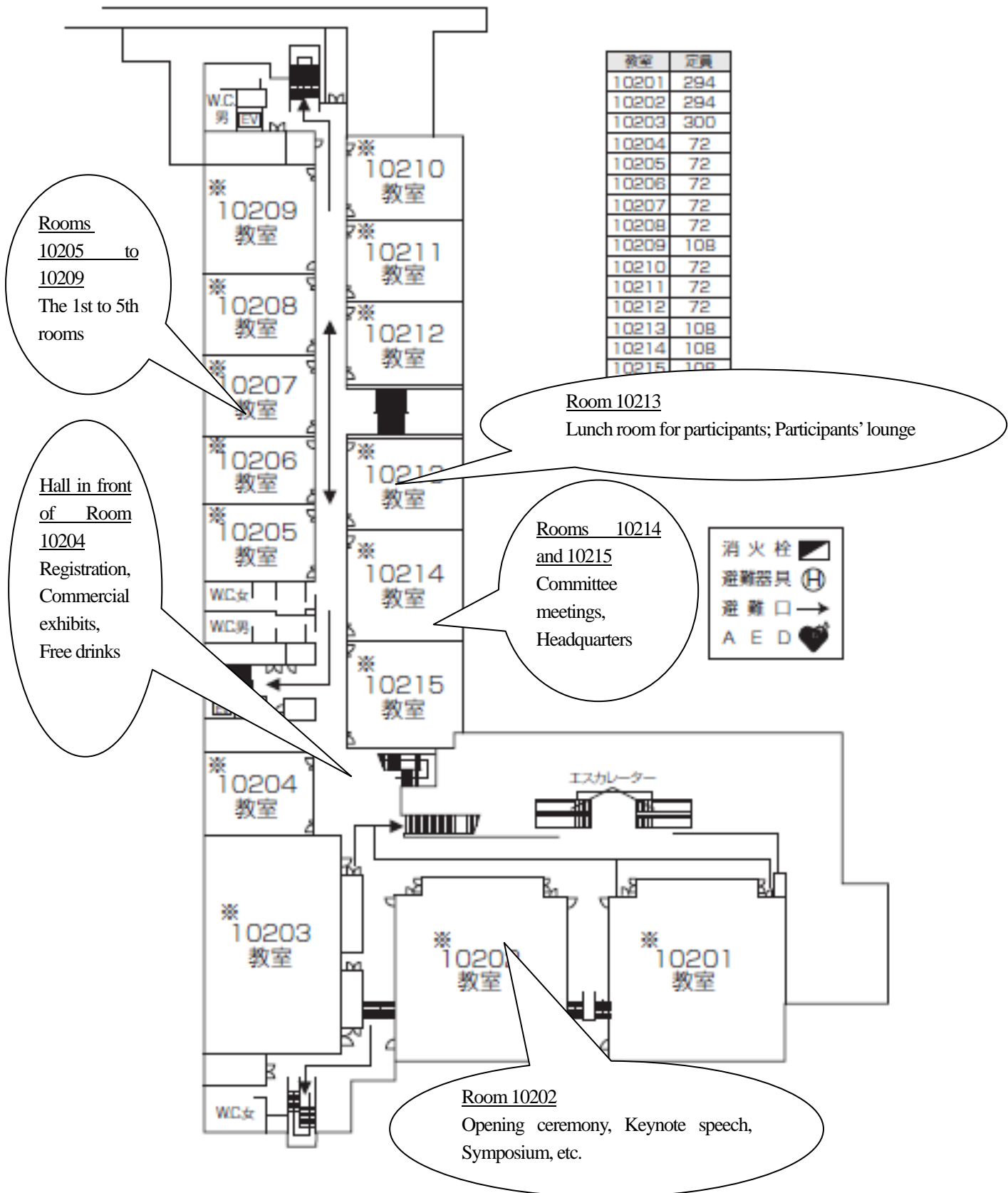




- 小田急線向ヶ丘遊園駅下車(新宿より急行で25分)
- 小田急バス向ヶ丘遊園北口より専修大学前行(約10分)
- 徒歩、南口より(約14分:130年記念館入口)

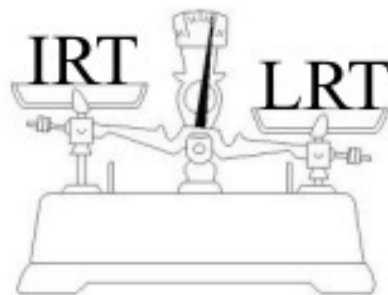


Building 10 (2nd floor)



項目応答理論？潜在ランク理論？

あなたのテストを分析できますか？



Latent Rank Theory

潜在ランク理論に基づくアダプティブテストを
試しませんか？

株式会社eラーニングサービス
askme@e-learning-service.co.jp

Commercial Exhibits (展示協賛企業)

株式会社 eラーニングサービス
e Learning Service

<http://www.e-learning-service.co.jp/index.html>

英語能力測定・評価研究会 (VELC 研究会)
Research Group for Visualizing English Language Competency

<http://www.velctest.org/group/>

Commercial exhibits are located at the hall in front of Room 10204. Please ask staff at the registration desk for details.

10204 教室前ホールに展示があります。詳しくは受付でお尋ねください。

We would like to acknowledge the generous financial support through the use of the Campus from Senshu University to make the annual conference of the Japan Language Testing Association possible. 第16回日本言語テスト学会全国研究大会の実現にあたりましては、専修大学より、会場の無償提供を始めご支援・ご協力を賜りました。誠にありがとうございました。

The next year's annual conference will be held in fall 2013, at the Waseda Campus of Waseda University in Shinjuku, Tokyo. The conference schedule will be announced via the JLTA website as soon as the details become available. We look forward to seeing you there.

2013年度の日本言語テスト学会全国研究大会は、2013(平成25)年秋に早稲田大学早稲田キャンパスで行われます。詳細が決まり次第、JLTAのホームページでお知らせいたします。どうぞご参加のほどよろしく願いいたします。

日本語テスト学会 (JLTA) 第16回 (2012年度) 全国研究大会発表要綱
Handbook of the 16th Annual Conference of the Japan Language Testing Association

発行日：2012年9月1日

発行：日本語テスト学会 (JLTA)

会長：中村 優治 (慶應義塾大学)

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

編集：全国研究大会実行委員会

印刷：(有) ユニオン・プレス 〒380-0815 長野県長野市鶴賀田町2152

TEL: 026-235-6236 FAX: 026-235-6258
