

日本語テスト学会 (JLTA)
第 17 回 (2013 年度) 全国研究大会発表要綱

Handbook
of
the 17th Annual Conference
of
the Japan Language Testing Association

Conference Theme:

Ensuring Appropriate Language Assessment Use in Globalizing Society

日時：2013年9月21日(土) 8:30-18:00

会場：早稲田大学早稲田キャンパス 22号館2・6・7階
〒169-8050 東京都新宿区西早稲田1-6-1 TEL: 03-3203-4141 (代表)

主催：日本語テスト学会 事務局
〒389-0813 長野県千曲市若宮758 TEL: 026-275-1964 FAX: 026-275-1970

共催：早稲田大学 教育・総合科学学術院
〒169-8050 東京都新宿区西早稲田1-6-1 TEL: 03-3203-4141 (代表)



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:
Waseda University, Waseda Campus
1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo 169-8050 Japan
Telephone: +81-3-3203-4141
<http://www.waseda.jp/eng/campus/map.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)
澤木 泰代	(早稲田大学)

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)
井上 千尋	(University of Bedfordshire)
金子 恵美子	(会津大学)
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)

全 国 研 究 大 会 実 行 委 員

澤木 泰代	(早稲田大学)
-------	---------

研 究 発 表 審 査 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. Conference Schedule Overview (大会日程表)	4
2. From the JLTA Office: Information for Conference Participants (学会事務局からのお知らせ)	12
3. Abstracts (発表要旨)	15
4. Workshop Information (ワークショップ情報)	50
5. Access to the Conference Venue (会場へのアクセス)	52
6. Special Conference Sponsors and Advertisement/Commercial Exhibit Sponsors (特別協賛企業、広告・展示協賛企業)	56

Conference Schedule Overview

September 20, 2013 (Friday)

16:30—18:30	Board Meeting	(TBA)
-------------	---------------	-------

September 21, 2013 (Saturday), 2nd, 6th, and 7th floors, Building 22, Waseda Campus, Waseda University

8:30—	Registration	(Room 203)
9:00—9:10	Opening Ceremony	(Room 202)
9:10—9:25	Report from the Web Publication Committee	(Room 202)
9:25—9:30	Report on the ILTA Code of Ethics Translation Project	(Room 202)
9:35—10:55	Keynote Speech	(Room 202)
11:05—11:35	Presentation I	
11:40—12:10	Presentation II	
12:15—12:45	Presentation III	
12:45—14:00	Lunch Break	(JLTA Committee Meetings: Room 203)
14:00—14:30	Presentation IV	
14:35—15:05	Presentation V	
15:10—15:40	Presentation VI	
15:55—17:25	Symposium	(Room 202)
17:30—17:50	JLTA General Business Meeting	(Room 202)
17:50—18:00	Closing Ceremony	(Room 202)
18:30—20:30	Banquet	(Restaurant & Café Takada Bokusha)

Commercial Exhibits: Room 617 (Free refreshments are available.)

Lunch Rooms for Participants & Participants' Lounges:

Rooms 601 & 616 (Please use only these rooms for lunch.)

Headquarters:

Room 203

September 22, 2013 (Sunday)

10:00—14:00	Post-Conference Workshop “Fundamentals of Item Response Theory” Jeffrey STEWART (Kyushu Sangyo University, Cardiff University) Aaron O. BATTY (Keio University, Lancaster University) (Conducted in English; Room 617, 6th floor, Building 22 Waseda Campus, Waseda University)
-------------	---

Program of the 17th JLTA Annual Conference

September 21, 2013 (Saturday)

8:30— Registration (Room 203, 2nd floor, Building 22)

Conference Attendance Fee: JLTA Members: ¥1,000

Non-members: ¥3,000 (Students: ¥1,000)

9:00—9:10 Opening Ceremony (Room 202, 2nd floor, Building 22)

Coordinator: Yo IN'NAMI (Shibaura Institute of Technology)

Greetings: Yuji NAKAMURA (JLTA President; Keio University)

9:10—9:25 Report from the Web Publication Committee (Room 202, 2nd floor, Building 22)

Coordinator: Yo IN'NAMI (Shibaura Institute of Technology)

Presenters: Yukie KOYAMA (Nagoya Institute of Technology), Yasuhiro IMAO (Osaka University), Rie KOIZUMI (Juntendo University), Randy THRASHER (Professor Emeritus, Okinawa Christian University & International Christian University), Youichi NAKAMURA (Seisen Jogakuin College), Minoru AKIYAMA

(e-Learning Service), Ken NORIZUKI (Shizuoka Sangyo University)
Theme: The Present Stage of Language Testing Web Tutorials

- 9:25—9:30 Report on the ILTA Code of Ethics Translation Project (Room 202, 2nd floor, Building 22)
Coordinator: Yo IN'NAMI (Shibaura Institute of Technology)
Presenter: Yasuyo SAWAKI (Waseda University), Rie KOIZUMI (Juntendo University), Yosuke YANASE (Hiroshima University)
Theme: Developing a Japanese Translation of the ILTA Code of Ethics
- 9:35—10:55 Keynote Speech (Room 202, 2nd floor, Building 22)
Coordinator: Yasuyo SAWAKI (Waseda University)
Introduction of the lecturer: Yuji NAKAMURA (JLTA President; Keio University)
Theme: Justifying the Uses of Language Assessments: Linking Test Performance to Consequences
Lecturer: Lyle F. BACHMAN (Professor Emeritus, University of California, Los Angeles)
- 11:05—12:45 Presentations I to III (Presentation: 20 minutes; Discussion: 10 minutes)
- 12:45—14:00 Lunch
JLTA Committee Meetings: Room 203, 2nd floor, Building 22
Lunch Room for Participants: Rooms 601 & 616, 6th floor, Building 22)
- 14:00—15:40 Presentations IV to VI (Presentation 20 minutes; Discussion 10 minutes)
- 15:55—17:25 Symposium (Room 202, 2nd floor, Building 22)
Theme: Building an Argument for Language Assessment Use in Japan
Coordinator: Yoshinori WATANABE (Sophia University)
Panelist: Yasuhiro IMAO (Osaka University)
Panelist: Keita NAKAMURA (EIKEN Foundation of Japan)
Panelist: Yoshinori WATANABE (Sophia University)
Discussant: Lyle F. BACHMAN (Professor Emeritus, University of California, Los Angeles)
- 17:30—17:50 JLTA General Business Meeting (Room 202, 2nd floor, Building 22)
Selection of the chair
Reporter: Youichi NAKAMURA (JLTA Secretary General; Seisen Jogakuin College)
- 17:50—18:00 Closing Ceremony (Room 202, 2nd floor, Building 22)
Coordinator: Yuko SHIMIZU (Ritsumeikan University)
- 18:30—20:30 Banquet (Restaurant & Café Takada Bokusha)
Coordinator: Tomoko FUJITA (Tokai University)

大会日程表

2013年9月20日(金)

16:30-18:30	理事会	(未定)
-------------	-----	------

2013年9月21日(土) 早稲田大学早稲田キャンパ22号館2・6・7階

8:30-	受付	(203教室)
9:00-9:10	開会行事	(202教室)
9:10-9:25	Web公開委員会報告	(202教室)
9:25-9:30	ILTA 倫理規範翻訳プロジェクト報告	(202教室)
9:35-10:55	基調講演	(202教室)
11:05-11:35	研究発表 I	
11:40-12:10	研究発表 II	
12:15-12:45	研究発表 III	
12:45-14:00	昼食	(役員会 : 203教室)
14:00-14:30	研究発表 IV	
14:35-15:05	研究発表 V	
15:10-15:40	研究発表 VI	
15:55-17:25	シンポジウム	(202教室)
17:30-17:50	JLTA 総会	(22号館2階202教室)
17:50-18:00	閉会行事	(22号館2階202教室)
18:30-20:30	懇親会	(レストラン&カフェ「高田牧舎」)

協賛企業展示 : 617教室 (無料の飲み物等がございます)
 一般参加者昼食・休憩室 : 601・616教室 (昼食は、この部屋のみでお願いいたします)
 大会本部 : 203教室

2013年9月22日(日)

10:00-14:00	ワークショップ「項目応答理論の基礎」(Fundamentals of Item Response Theory) 講師 : Jeffrey STEWART (九州産業大学、カーディフ大学) Aaron O. BATTY (慶應義塾大学、ランカスター大学) (早稲田大学早稲田キャンパス22号館6階617教室)
-------------	---

日本言語テスト学会第17回全国大会プログラム

2013年9月21日(土)

8:30- 受付 (22号館2階203教室)
 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 (22号館2階202教室)
 総合司会 印南洋 (芝浦工業大学)
 挨拶 中村優治 (JLTA 会長・慶應義塾大学)

- 9:10—9:25 **Web 公開委員会報告** (22 号館 2 階 202 教室)
 司会 印南 洋 (芝浦工業大学)
 発表者 小山 由紀江 (名古屋工業大学)、今尾 康裕 (大阪大学)、小泉 利恵 (順天堂大学)、Randy THRASHER (沖縄キリスト教学院大学・国際基督教大学名誉教授)、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)、法月 健 (静岡産業大学)
 題目 The Present Stage of Language Testing Web Tutorials
- 9:25—9:30 **国際言語テスト学会倫理規範翻訳プロジェクト報告** (22 号館 2 階 202 教室)
 司会 印南 洋 (芝浦工業大学)
 発表者 澤木 泰代 (早稲田大学)、小泉 利恵 (順天堂大学)、柳瀬 陽介 (広島大学)
 題目 Developing a Japanese Translation of the ILTA Code of Ethics
- 9:35—10:55 **基調講演** (22 号館 2 階 202 教室)
 司会 澤木 泰代 (早稲田大学)
 紹介 中村 優治 (JLTA 会長・慶應義塾大学)
 演題 Justifying the Uses of Language Assessments: Linking Test Performance to Consequences
 講師 Lyle F. BACHMAN (Professor Emeritus, University of California, Los Angeles)
- 11:05—12:45 **研究発表 I・II・III** (発表 20 分, 質疑応答 10 分)
- 12:45—14:00 **昼 食**
 (役員会 : 22 号館 2 階 203 教室 一般参加者昼食控室 : 22 号館 6 階 601・616 教室)
- 14:00—15:40 **研究発表 IV・V・VI** (発表 20 分, 質疑応答 10 分)
- 15:55—17:25 **シンポジウム** (22 号館 2 階 202 教室)
 テーマ Building an Argument for Language Assessment Use in Japan
 (使用言語 : 英語)
 コーディネーター 渡部 良典 (上智大学)
 パネリスト 今尾 康裕 (大阪大学)
 パネリスト 仲村 圭太 (日本英語検定協会)
 パネリスト 渡部 良典 (上智大学)
 討論者 Lyle F. BACHMAN (Professor Emeritus, University of California, Los Angeles)
- 17:30—17:50 **JLTA 総会** (22 号館 2 階 202 教室)
 議長選出
 報告 中村 洋一 (JLTA 事務局長・清泉女学院短期大学)
- 17:50—18:00 **閉会行事** (22 号館 2 階 202 教室)
 司会 清水 裕子 (立命館大学)
- 18:30—20:30 **懇親会** (レストラン&カフェ「高田牧者舎」)

司会 藤田 智子 (東海大学)

	Yasuyo SAWAKI (Waseda University), Rie KOIZUMI (Juntendo University), Yosuke YANASE (Hiroshima University)	Developing a Japanese Translation of the ILTA Code of Ethics (p. 21)
	Keynote speech Lyle F. BACHMAN (Professor Emeritus, University of California, Los Angeles)	Justifying the Uses of Language Assessments: Linking Test Performance to Consequences (p. 15)
I–VI	--	--
	Symposium Coordinator: Yoshinori WATANABE (Sophia University) Panelist: Yasuhiro IMAO (Osaka University) Panelist: Keita NAKAMURA (EIKEN Foundation of Japan) Panelist: Yoshinori WATANABE (Sophia University) Discussant: Lyle F. BACHMAN (Professor Emeritus, University of California, Los Angeles)	Building an Argument for Language Assessment Use in Japan (pp. 16-19)

Room 1 (Room 618)

Chair	Part I	Hidetoshi SAITO (Ibaraki University)
	Part II	Yo IN'NAMI (Shibaura Institute of Technology)
	Part III	Yo IN'NAMI (Shibaura Institute of Technology)
	Part IV	Yasuhiro IMAO (Osaka University)
	Part V	Emiko KANEKO (University of Aizu)
	Part VI	Takanori SATO (University of Melbourne)

Part	Presenter (Affiliation)	Title (Page)
I	Takanori SATO (Graduate School, University of Melbourne, Australia)	The Assessment Criteria for Oral Presentations Derived From Linguistic Laypersons (p. 22)
II	Veronika TIMPE (TU Dortmund University, Germany)	The Dependence of Socio-Pragmatic Competence on Learning Opportunities: A Fairness Issue for Foreign Language Learning? (p. 23)
III	Qin XIE (Hong Kong Institute of Education, China)	Does Test Design Affect Time Management and Approaches to Preparation? A Study on Washback Mechanism (p. 24)
IV	Hoky MIN (Korea Institute for Curriculum and Evaluation, South Korea)	Introducing the Writing Section of the NEAT (p. 25)
V	Daniel JOYCE (EIKEN Foundation of Japan)	Investigating the Importance Placed on Criterial Features by Raters of Productive Writing: The Case of Raters of the Writing Components of Two High Stakes English Tests in Japan, the EIKEN Grade 1 and Pre-1 Tests (p. 26)
VI	Jeffrey DURAND (Tokai University)	Rater Judging Plans and Sensitivity to Unexpected Scores (p. 27)

Room 2 (Room 619)

Chair	Part I	Takayuki NAKANISHI (Tokiwa University)
	Part II	Naoyuki NAGANUMA (Tokai University)
	Part III	Kahoko MATSUMOTO (Tokai University)
	Part IV	Trevor HOLSTER (Fukuoka Women's University)
	Part V	Hiroshi SHIMATANI (Kumamoto University)
	Part VI	Jeffrey K. HUBBELL (Hosei University)

Part	Presenter (Affiliation)	Title (Page)
I	Kiyomi YOSHIZAWA (Kansai University), Atsuko TAKASE (Kansai University), Kyoko OTSUKI (Hiroshima Shudo University)	A Comparison of the EPER Form A and Form E: Do They Work as Alternative Forms? (p. 28)
II	Trevor A. HOLSTER (Fukuoka Women's University), William R. PELLOWE (Kinki University), J W. LAKE (Fukuoka Women's University)	Monitoring Extensive Reading Using Mobile Phones (p. 29)
III	Jungok BAE (Kyungpook National University, South Korea)	The Relative Utility of Three Distinct Picture Prompts for Eliciting Language and Ideas (p. 30)
IV	Sébastien GEORGES (International Centre of Pedagogical Studies, France)	Cross-Actions between Linguists and Psychometricians Make Possible Reliable Tests for Multiple Examinee Profiles (p. 31)
V	Hideki IIMURA (Tokiwa University)	The Degree of Confidence in Distractors in Multiple-Choice Listening Tests (p. 32)
VI	Barry O'SULLIVAN (British Council, U.K.)	Breaking the Mould: Localisation and Validation in Test Development (p. 33)

Room 3 (Room 717)

Chair	Part I	Yukie KOYAMA (Nagoya Institute of Technology)
	Part II	Akiyo HIRAI (University of Tsukuba)
	Part III	Soo-im LEE (Ryukoku University)
	Part IV	Emiko KANEKO (University of Aizu)
	Part V	Kei MIYAZAKI (Keio Senior High School)

Part	Presenter (Affiliation)	Title (Page)
I	Atsushi MIZUMOTO (Kansai University)	Creating an In-House Computerized Adaptive Testing (CAT) Program with Concerto (p. 34)
II	Jinsong FAN (Fudan University, China)	The Factor Structure and Factorial Invariance of a University-Based EAP Test (p. 35)
III	Mehrasa ALIZADEH (Alzahra University, Iran)	Examining Local Dependence in the Iranian National University Entrance Examination (p. 36)
IV	Sachiyo TAKANAMI (Reitaku University)	Which is Better: Writing or Choosing in Spelling Tests? Comparing Recall with Recognition Test (p. 37)
V	Yuko HOSHINO (Tokyo Fuji University)	Development of a Test to Measure Knowledge Regarding Multiple Meanings of Basic Words (p. 38)
VI	--	--

Room 4 (Room 718)

Chair	Part II	藤田智子 (東海大学)
	Part III	藤田智子 (東海大学)
	Part IV	Keita NAKAMURA (EIKEN Foundation of Japan)
	Part V	Soo-im LEE (Ryukoku University)
	Part VI	Yoshinori WATANABE (Sophia University)

Part	Presenter (Affiliation)	Title (Page)
I	--	--
II	河住有希子 (日本工業大学)、藤田恵 (日本国際教育支援協会)、秋元美晴 (恵泉女学園大学)	言語テストにおける視覚障害者受験特別措置実施の目的と方法—日本語能力試験点字冊子試験を例に— (p. 39)
III	長沼君主 (東海大学)、高野正恵 (東京外国語大学)、ヘザー・ジョンソン (東京外国語大学)、工藤洋路 (駒沢女子大学)	CEFR に準拠したジャンル別ライティング及びスピーキング評価の開発と検討 (p. 40)
IV	Kaizhou LUO (Beijing Foreign Studies University, China), Baocheng HAN (Beijing Foreign Studies University, China)	Building an Assessment Use Argument for Matriculation English Test (Beijing Version): An Investigation into Refining the Reading Construct and Items (p. 41)
V	Parisa MEHRAN (Alzahra University, Iran)	Justifying the Use of an English Language Placement Test with an Assessment Use Argument (p. 42)
VI	Hossein KEYNEJAD (Islamic Azad University, Iran), Mehrasa ALIZADEH (Alzahra University, Iran)	Designing an Assessment Architecture: Evidence-Centered Design and Object-Oriented Methodology Intertwined (p. 43)

第5室 (719教室)

Chair	Part I	平野絹枝 (上越教育大学)
	Part II	島谷浩 (熊本大学)
	Part III	星野由子 (東京富士大学)
	Part IV	河住有希子 (日本工業大学)
	Part V	島田勝正 (桃山学院大学)
	Part VI	島田勝正 (桃山学院大学)

Part	Presenter (Affiliation)	Title (Page)
I	小林夏子 (教育測定研究所)、野上康子 (教育測定研究所)、吉川厚 (教育測定研究所)、林規生 (教育測定研究所)	ディクテーションテストにおける誤りの傾向 (p. 44)
II	船越貴美 (神戸学院大学附属高等学校)	高校生の語彙数はどのように伸びていくのか—英検取得級別に見た語彙サイズテストの成長曲線モデルによる分析 (p. 45)
III	宮本由美子 (長野県上田染谷丘高等学校)	高校生学習者の英語学習動機と自己評価の関係 (p. 46)
IV	横内裕一郎 (筑波大学大学院生)、青島健夫 (筑波大学大学院生)	異なる文章提示方法によってスピーキング再話課題の発話特徴に差は生じるのか (p. 47)
V	佐瀬文香 (筑波大学大学院生)	EFL 学習者における語順知識とスピーキング能力間の関係性 (p. 48)
VI	徳永美紀 (福岡大学)	日本人大学生の英語文法性判断力—試験的調査— (p. 49)

From the JLTA Office: Information for Conference Participants

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 in Room 203 on the 2nd floor of Building 22.
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 registration desk. The banquet will be held at Restaurant & Café Takada Bokusha (3 minutes' walk from the venue).
5. The conference handbook is available at the registration desk to JLTA non-members. We request JLTA members to bring with them the conference the handbook mailed in advance.

Lunch and Participants' Lounge

1. Please use Rooms 601 and 616 on the 6th floor of Building 22 for eating lunch as the Participants' Lounges. Lunch can be purchased at a nearby convenience store. There are a few convenience stores around the venue. The closest one is on the 1st floor of Building 11 on Waseda Campus (3 minutes' walk from the venue).
2. The following are locations on campus that are open for lunch on the conference day. See the Campus Map below for the locations.

Building 18 (2nd floor)	Center for Scholarly Information Tea Room
Building 23-5 (1st floor)	Uni. Shop & Café 125
Building 26 (15th floor)	Restaurant Seihoku no Kaze
3. There are a variety of restaurants and cafés off campus around the venue as well. For more information, please refer to the Waseda Town Guide. The PDF version of the Guide can be downloaded from: www.waseda.jp/rps/irp/handbook/ja/index.html
4. Complimentary refreshments are available in Room 617.
5. 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. 22号館2階の203教室で行います。
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。懇親会はレストラン&カフェ「高田牧舎」（大会会場から徒歩3分）にて開かれます。
5. 非会員の方には、『JLTA 第17回（2013年度）全国研究大会発表要綱』を受付で配布します。会員の方は、事前に郵送する『要綱』をご持参ください。

■昼食・休憩室

1. 昼食・休憩室として、22号館6階601・616教室をご利用ください。昼食は受付では販売いたしません。会場の周りにいくつかコンビニエンスストアがあります。一番近いものは早稲田キャンパス11号館1階です（会場から徒歩3分）。
2. 大会当日、大学施設では下記の食堂・喫茶、コーヒーショップが昼食時に営業予定です。場所はキャンパス・マップでご確認ください。

18号館（2階）	学術情報センター喫茶
23-5号館（1階）	Uni. Shop & Café 125
26号館（15階）	レストラン 西北の風
3. 大学キャンパス周辺には、その他にも昼食時にオープンしているレストランやカフェが多数あります。詳細は下記のURLから閲覧できる「早稲田・街の便利帖」(Waseda Town Guide) をご参照ください。(PDF版ダウンロード可) www.waseda.jp/rps/irp/handbook/ja/index.html
4. 無料の飲み物は22号館6階617教室にあります。
5. 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 202)

Justifying the Uses of Language Assessments: Linking Test Performance to Consequences

Lyle F. Bachman (Professor Emeritus, University of California, Los Angeles)
lfb@humnet.ucla.edu

We generally give a language assessment because we need to make some decisions. In language programs, we use assessments for selecting students for admission, placing them at appropriate levels for instructional purposes, assessing their progress and achievement in the program, or assigning grades. Language assessments are also used for making decisions about hiring or promoting individuals in companies, for certifying professionals, and for immigration and naturalization. *All* of these decisions will have consequences for stakeholders, and many of these decisions are high-stakes, entailing major consequences for stakeholders. We therefore need to be able to justify the decisions we make on the basis of test scores, so that we can be accountable to the stakeholders—the various individuals who will be affected in one way or another by the assessment and by the way we use it.

In order to justify using the results of a language assessment for making decisions, we need to provide a rationale for linking students' performance on the assessment with the intended uses—the decisions that are made and the consequences of these. This rationale is provided by an **assessment use argument (AUA)**. By *demonstrating*, through argumentation and the collection of supporting evidence, that our assessment is useful for its intended purpose, we provide the justification we need to be accountable to the individuals who are affected by the assessment and the way it is used.

Bio

Lyle F. Bachman is Professor Emeritus of Applied Linguistics at the University of California, Los Angeles. He is a Past President of the American Association for Applied Linguistics and of the International Language Testing Association. He has received numerous awards for his research and service to the profession, including the TESOL/Newbury House Award for Outstanding Research, the Modern Language Association of America's Kenneth Mildenerger Award for outstanding research publication, the Sage/International Language Testing Association award for the best book published in language testing, the Lifetime Achievement Award from the International Language Testing Association, and the Distinguished Scholarship and Service Award from the American Association for Applied Linguistics. He has published numerous articles and books in language testing and other areas of Applied Linguistics. His current research interests include validation theory, classroom assessment, and epistemological issues in Applied Linguistics research.

Symposium (Room 202)

Building an Argument for Language Assessment Use in Japan

Coordinator Yoshinori WATANABE (Sophia University)

Panelists Yasuhiro IMAO (Osaka University)
Keita NAKAMURA (EIKEN Foundation of Japan)
Yoshinori WATANABE (Sophia University)

Discussant Lyle F. BACHMAN (Professor Emeritus, University of California, Los Angeles)

Introduction

Coordinator: Yoshinori WATANABE (Sophia University)

yjwatana@gmail.co.jp

The symposium is intended to demonstrate how an argument-based approach to validity in general and an Assessment Use Argument in particular can be implemented to justify the use of assessments in real world settings. The topics vary among the three presenters, but are in common in that they all deal with the validation of assessments which are used to make high-stakes decisions in the context of Japanese educational system.

Imao argues that the current practice of university admission examinations has to fulfill double-functions, one for assessing test-takers' achievement and the other for assessing their proficiency levels. This ambiguous status of the examination is likely to be counterproductive. He illustrates the use of AUA to resolve the dilemma by offering practical suggestions. Nakamura and Green report on the result of the questionnaire they administered to high-school students and teachers about the possible consequences of a new test of Academic English for university admission. By doing so, the authors demonstrate how the survey helps examine if the intended impact of the assessment has successfully been engineered to pre-college level education. Watanabe focuses on washback effects of assessments in Japan. He reformulates a number of claims and assertions that have been made concerning the university admissions in the framework of an argument-based approach to validation. He gives examples of how to carry out empirical research to explore washback to gather evidence (i.e. backing) by observations. Based on the analysis, he suggests the way of generating positive impact on pre-college level education in Japan.

The symposium concludes by offering suggestions as to the way of implementing the argument-based validation to the context of Japan in an attempt to facilitate appropriate use of high-stakes assessments for EFL learners in the country.

Paper 1: Examining the issues in the English portion of university entrance exams in Japan using the Assessment Use Argument (AUA) framework

Yasuhiro Imao (Osaka University)

imao@lang.osaka-u.ac.jp

In the last several months, university entrance examinations have once again been a hot issue in Japan. There are, however, some fundamental problems in the debates of the university entrance examination reform in Japan, such as overreliance on unwarranted claim of beneficial consequences, or positive washback. This paper reviews one of the most fundamental issues using the AUA as a framework (Bachman & Palmer, 2010) to introduce a conceptually new approach to test validation.

One of the fundamental issues in the English portion of the university entrance examinations is a lack of consensus among test developers and test users on the construct to be measured for university entrance decision making. There are at least two possible approaches to defining the construct. One approach would be to base the construct on the National Curriculum Guidelines, which essentially makes the entrance examination an achievement test. Another approach would be to define the construct based on what students need to be able to do in English at universities.

These two approaches require totally different warrants and backings in interpretation of test scores. In the National Curriculum Guidelines approach, the interpretation of test scores needs to be meaningful with respect to the Guidelines and the tasks should be generalizable to the language instructional tasks although those tasks are hard to identify because of the diversity of the tasks used in secondary schools all over Japan. The National Center Test for University is designed in this approach and should be evaluated accordingly.

When defining the construct based on needs analyses, the target language use tasks would be language instructional tasks or real life tasks in university contexts. In most universities in Japan, English is taught as a general education subject, which makes the detailed analysis of English language classes more appropriate as a reference point to evaluate the score interpretation. If English is used as an instructional language in content classes and/or in out-of-class communication, meaningfulness and generalizability should be evaluated as such. Therefore, unless a consensus on the construct is reached, the interpretation of scores would not be warranted.

The present paper concludes by suggesting that this high-stakes test in Japan will be improved by creating a new test or selecting an existing English test whose construct reflects or matches English language instruction/use in university contexts in Japan. This approach warrants the claim that score interpretations are meaningful and generalizable, which will provide useful information for university entrance decision making.

Bio

Yasuhiro Imao is an associate professor in the Graduate School of Language and Culture, Osaka University. He received his Ph.D from UCLA in 2010 under the supervision of Prof. Lyle Bachman. His research interests include writing assessment, academic writing, and corpus analysis. He also develops various computer programs for Mac, including corpus analysis tools, a transcription aiding tool, and a GUI frontend for a statistical computing environment, R.

Paper 2: An application of AUA to examining possible consequences of a new test of English for university entrance

Keita Nakamura (EIKEN Foundation of Japan)

ke-nakamura@eiken.or.jp

Anthony Green (CRELLA, University of Bedfordshire)

Taking the perspective of an Assessment Use Argument (Bachman & Palmer, 2010), this paper reports on the possible consequences of introducing an innovative four-skills test of English for Academic Purposes for university entrance in Japan. It presents and discusses an anticipatory baseline questionnaire targeting potential test consequences and looks forward to the planned implementation of the test from the 2015 school year.

In order to investigate the prospects for the overall claim that the introduction of the new test will bring positive consequences for stakeholders, the authors conceptualized ways in which characteristics of the test tasks, procedures and use might impact on teaching and learning of English in Japan. We incorporated these into a detailed statement of intended impact indicating how features of the test design, procedures and use are intended to influence educational practice at high schools. As one element of the validation plan, the team gathered data on the perceptions of the new test among relevant stakeholder. A survey was conducted of 3,868 high school students and 423 high school teachers. The questionnaires covered five main categories, but this paper will focus on two: stakeholder perceptions of 1) the washback effects of the current university entrance examinations, and 2) likely effects of changes to the content of university entrance examinations.

The paper will outline how questionnaire data can be used to evaluate how far the test developers' intentions are shared by stakeholders, to anticipate reactions to the introduction of the new test, and to plan strategies to promote positive impact.

Bio

Keita Nakamura is researcher at Eiken Foundation of Japan. He is currently studying at the Ed.D program at Temple University. His main area of the work is on the test validation using both qualitative and quantitative methods. As a member of the test development section at Eiken, he has worked with various research projects such as the development of the measurement tool of young learners' English proficiency, the investigation of brain activities during English interview test.

Anthony Green is Professor of Applied Linguistics in the Centre for English Language Learning and Assessment at the University of Bedfordshire, UK. As a member of one of Europe's leading research centres for language testing and assessment, he works on a wide range of test development and validation programmes around the world. He has extensive experience as a language teacher, a lecturer in language assessment, test developer and examiner and has worked for national and international testing organisations. He has a particular research interest in the relationships between testing and teaching: He has published a number of journal papers on this topic as well as the book, *IELTS Washback in Context* (2007, Cambridge University Press).

**Paper 3: Washback of high-stakes assessments in Japan:
Retrospect and prospect evaluated in the framework of
Argument-based approaches to test validation**

Yoshinori Watanabe (Sophia University)
yjwatana@gmail.co.jp

An increasing number of empirical studies have been reported on washback effects of language assessments in the past two decades. The body of the research to date indicates that washback does exist, but its nature is extremely complex being observed on multiple dimensions. Washback is a socio-psychological phenomenon in the sense that it is always mediated by test user and contextual factors rather than being inherent in the assessment itself. However, there are still too many claims that are based on a naïve view as if innovating in assessments would automatically induce innovation in education in a corresponding manner. In order to generate positive washback, we need more evidence indicating the presence or absence of washback. And only after that, can we identify the conditions under which washback operates.

In this presentation, I will summarize the claims and assertions that have been made concerning the role of high-stakes assessments in Japanese education, including the criterion-referenced grading system at secondary schools, the National Center Test and the university entrance examination, amongst others. By so doing, an attempt will be made to show the usefulness of argument-based approaches to test validation to deepen our understanding of washback effects of language assessment. Based on the analysis, I will demonstrate the process of validating empirically through observations. To conclude, several recommendations will be made so they may help generate positive washback to education, based on the findings that have been made in the field exploring the issue of test specifications, assessment literacy, diffusion of innovations, and learning-oriented use of language assessment.

Bio

Yoshinori Watanabe, PhD (Lancaster, UK), is professor of Faculty of Foreign Studies at Sophia University. His dissertation topic was the washback effect of Japanese university entrance examinations and he has edited and authored a number of publications in this area (e.g. *Washback in Language Testing* edited with Liying Cheng, 2004, Routledge). His recent contribution includes *CLIL (Content and Language Integrated Learning) Volume I* (2012) and *Volume II* (with Ikeda and Izumi, 2013). He serves on the editorial advisory board of *Language Assessment Quarterly*.

Report from the Web Publication Committee (Room 202)

The Present Stage of Language Testing Web Tutorials

Yukie KOYAMA (Nagoya Institute of Technology)

koyama@nitech.ac.jp

Yasuhiro IMAO (Osaka University)

Rie KOIZUMI (Juntendo University)

**Randy THRASHER (Professor Emeritus, Okinawa Christian
University & International Christian University)**

Youichi NAKAMURA (Seisen Jogakuin College)

Minoru AKIYAMA (e Learning Service)

Ken NORIZUKI (Shizuoka Sangyo University)

As was already reported at the 2012 JLTA conference, the Web Publication Committee has been working on the following two areas: (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.

(A) Registration of JLTA Journal on CiNii

The committee already registered all the past JLTA Journal articles on CiNii.

(B) Publishing useful materials on the JLTA web site

This includes the following three types of publications; 1) digitalized materials useful for language testing such as booklets, 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.

Based on the results from the questionnaire conducted in the previous year, the Web Publication Committee has focused on the creation of JLTA original on-line tutorials in answer to the needs of junior- and senior-high school teachers. Of the on-line tutorials, “Practical considerations in developing language tests” in the whole frame-work was chosen to launch now. The first part of the on-line tutorials was completed in July, 2013, with contributions from many JLTA members, teachers at high schools and universities.

This presentation explains the procedure of the tutorial creation, and shows the products in addition to the progress of the committee’s other activities.

Report on the ILTA Code of Ethics Translation Project (Room 202)

Developing a Japanese Translation of the ILTA Code of Ethics

Yasuyo SAWAKI (Waseda University)

ysawaki@waseda.jp

Rie KOIZUMI (Juntendo University)

Yosuke YANASE (Hiroshima University)

The Code of Ethics of the International Language Testing Association (ILTA) was adopted at its annual meeting held in Vancouver in 2000 to promote ethical conduct of all ILTA members as professionals engaged in language testing and assessment services. The document comprises nine fundamental principles and their annotations that specify what ILTA members ought to do and should not do as language testing professionals as well as challenges and exceptions that ILTA members may face in putting the principles into practice. In 2011 the ILTA Executive Board launched a project to translate the document into various languages to facilitate dissemination of the principles laid out in the document across different countries and regions of the world. So far, provisional translations of the ILTA Code of Ethics into four languages, including Japanese, are complete, while translation work into five other languages are currently under way. In this presentation we will provide an overview of the translation project, including the procedure employed to develop the Japanese version of the document, and future plans for releasing the provisional Japanese translation on the ILTA website for use by Japanese-speaking language testing professionals.

Paper Presentations

Room 1 (Room 618) Part I

The Assessment Criteria for Oral Presentations Derived From Linguistic Laypersons

Takanori SATO (Graduate School, University of Melbourne)
t.sato@student.unimelb.edu.au

Most English proficiency tests are *weak performance tests* that focus exclusively on the test-takers' linguistic quality of performance. Accordingly, the assessment criteria used by most speaking tests consist of linguistic features such as grammatical accuracy, vocabulary, and pronunciation. These linguistically oriented criteria are usually the only features that language professionals (e.g., applied linguists and language teachers) are interested in. Nevertheless, research has shown that various linguistic lay domain experts (e.g., physicists) do not judge second language (L2) communicative performance based on linguistic quality per se. This indicates the possibility that test scores on weak performance tests may not reflect the subjective judgment of stakeholders in real-life domains. This study aims to explore the assessment criteria employed by linguistic laypersons in assessing L2 speakers' oral presentations.

Twenty-three graduate students of disciplines other than applied linguistics and TESOL participated in this study. First, the participants watched videos of short oral presentations given by seven test-takers of the national College English Test-Spoken English Test of China (CET-SET). Second, they indicated their impression of each test-taker's *communication ability* on a scale of 1 (Poor) to 7 (Excellent), without being given any pre-determined criteria. Third, the participants were asked to verbalize the reasons for their rating. Fourth, they reviewed the same performance and participated in stimulated recall sessions, verbalizing features of the performances that influenced their impressions. Post-session interviews were also conducted to solicit their perceptions of salient features and behaviors of the test-takers.

The participants' ratings were analyzed using FACETS. The results showed that their ratings were aligned with the test-takers' CET-SET scores; overall, the participants rated high score achievers highly and low score achievers lowly. This finding seems to indicate that linguistic laypersons' impression of communication ability is dependent on English proficiency. However, the participants' verbal protocols showed that they focused not only on linguistic quality but also on a variety of features that language professionals would consider construct-irrelevant: demeanor, non-verbal behaviors, content, topical knowledge, and global comprehensibility. In addition, the post-session interviews revealed that linguistic resources (i.e., grammar and vocabulary) were not deemed influential and salient in the judgments of communication ability.

This study suggests that linguistically oriented assessment criteria do not represent all the features attended by general listeners in judging oral presentations outside the testing milieu. Language professionals should pay closer attention to non-linguistic features that may strongly affect listeners' judgment of a speaker's communicative performance.

Room 1 (Room 618) Part II

The Dependence of Socio-pragmatic Competence on Learning Opportunities: A Fairness Issue for Foreign Language Learning?

Veronika TIMPE (TU Dortmund University)
veronika.timpe@udo.edu

Several language frameworks identify receptive and productive pragmatic abilities as constitutive components of (intercultural) communicative competence (e.g., Bachman, 1990; Bachman & Palmer, 2010; Byram, 1997; Canale, 1983; Canale & Swain 1980; Hymes, 1972). However, even though pragmatics has long been anchored in language competence frameworks, it is still an underrepresented and challenging domain in language teaching and assessment.

This study investigated the development of socio-pragmatic comprehension in relation to learning opportunities as experienced by German university learners of English and thus, explored the following questions: (1) Do German university-level EFL/ESL learners vary in receptive socio-pragmatic competence depending on different types and amounts of English learning opportunities and contexts? (2) Which types of target language input contribute to higher levels of receptive socio-pragmatic competence?

The study was subdivided into two phases. First, a test of socio-pragmatic comprehension was developed, which operationalized and measured students' pragmatic knowledge of U.S.-American English. Once the test was found to provide reliable results, the second and primary part of the study was conducted with 105 L1 German university-level learners of English ($N = 105$). All candidates took the socio-pragmatic comprehension test as well as an online questionnaire which assessed their experience in the TL environment (United States) and amount of exposure to target language input in the home country (Germany). In an initial analysis, students were grouped based on their prior residence in the U.S. and amount of TL input. A one-way MANOVA was used to investigate between-group differences. Subsequently, multiple regression analyses were employed to examine the effect of learning contexts and different types of target language input on the L2 learners' sociopragmatic comprehension. Findings revealed that frequent exposure to audiovisual media – more than a prolonged sojourn in the United States – was a major influential factor for the development of receptive pragmatic competence in U.S.-American English.

The findings have a number of practical implications. For example, films, sitcoms, and soap operas are highly contextualized, authentic forms of input that may provide a useful means of promoting L2 pragmatic competence in the FL learning context. Moreover, the possibility that learners can gain pragmatic competence within the home context would support the inclusion of pragmatic knowledge in tests of English as a foreign language. That is, if all FL students have the opportunity to gain pragmatic knowledge, not just those who study abroad, then there is little reason from a fairness standpoint to avoid testing this domain.

References

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford Univ. Press.
- Bachman, L. F., & Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence. Multilingual matters*. Clevedon: Multilingual Matters.
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 2–27). New York: Longman.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, (1), 1–47.
- Hymes, D. (1972). On communicative competence. In J. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Middlesex: Penguin Books Ltd.

Room 1 (Room 618) Part III

Does test design affect time management and approaches to preparation? A study on washback mechanism

**Qin XIE (Hong Kong Institute of Education)
qxie@ied.edu.hk**

Existing studies on the washback of high-stakes language proficiency tests are mostly qualitative; there is a lack of quantitative studies exploring its mechanism. This study utilized structural equation modeling to investigate washback mechanism, focusing on two aspects of test design: component weighting and subjective test methods, and their washback on test preparation. Test-taker perceptions of test design and their test preparation activities were surveyed ($N = 1,003$); their test scores were also collected. Data was analyzed to estimate the washback effects on test-taker time management, approaches to preparation, and test performance. The study found that test takers spent more time on the components with higher weight and much less on those with lower weight. Reporting component scores separately did not seem to adjust this tendency. Meanwhile, favorable perception of test validity was associated with a higher level of engagement in both language learning activities and focused test preparation. This suggests that favorable perception may not be able to reduce negative washback effects, but may be able to promote positive ones.

Key words: Component weighting, washback, subjective testing methods, test preparation, Structural equation modeling

Room 1 (Room 618) Part IV
Introducing the Writing Section of the NEAT

Hoky MIN (the Korea Institute for Curriculum and Evaluation)
hoky@kice.re.kr

The writing section of the NEAT is composed of six items for level 2 and level 3. The items for level 2 include “Writing about daily lives,” where the test takers describe their personal experience about a given event or an object and “Expressing one’s own opinion,” where the test takers discuss pros or cons about a given particular issue. The items for level 3, on the other hand, include “Selective picture description with given words.” In this item, the test takers choose one among the three given situations and describe it using given words. In “One-picture description,” they describe the actions or behaviors of the people in a picture. In “Letter writing,” the test takers write an email or a letter based on given an advertisement. Lastly, in “Two-picture description and inference” the test takers describe two given pictures and make an inference about the following situation. The rating domains of the writing sections are basically content, organization, language use, and task completion. The definition of each of the rating domains is carefully constructed. In the content domain, the main idea in student’s response must be clearly and deeply discussed, providing adequate supporting details. The organization domain tests whether the writing is logical and consistent in cohesion and coherence, thereby increasing the efficiency of the delivery of the information. Language use is the domain which evaluates whether the structure of the sentences, and the usage of grammar and the spelling are accurate. The usage of various expressions and vocabulary must be in accordance with the situation. The flow of the writing should be natural in its various forms, genre, and situations. Finally, the task completion examines whether the given conditions are completed with proper and appropriate sentences, thereby increasing the probability of general comprehension. Here, the rater must eliminate the evaluation of values or truth. The test takers’ written responses are scored by the group of English teachers who have been trained for the online KICE rater training program. The chief raters grade the samples which will be used during the online training in order to compare the trainees’ scores for the first and second pilot ratings with their own scores. The trainees are certified by passing the final rating test at the end of the online program. In the 2012-2013 NEAT administrations, the certified raters showed significantly reliable results.

Room 1 (Room 618) Part V

Investigating the importance placed on criterial features by raters of productive writing: The case of raters of the writing components of two high stakes English tests in Japan, the EIKEN Grade 1 and Pre-1 tests

Daniel JOYCE (Eiken Foundation of Japan)

d-joyce@eiken.or.jp

This paper reports on one part of a comprehensive study to validate the rating scales for the writing components of two high stakes English proficiency tests in Japan, the EIKEN Grade 1 and Grade Pre-1 tests. The rating scales were introduced in 2004 when the writing components were revised to provide an opportunity to test productive English writing skills in the form of a communicative task. The revision was in response to a trend within EFL in Japan at the time toward greater emphasis on communicative competence (Sasaki, 2008), and in particular the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) *Action Plan to Cultivate “Japanese with English Abilities”* (MEXT, 2002; MEXT, 2003). The scales were thus designed to guide raters toward placing greatest importance on the extent to which examinees can communicate their message, rather than on mechanical accuracy. One goal of this study was to evaluate the effectiveness over time of rater training and support materials in maintaining an interpretation of the scales by raters that is consistent with the intentions of the test developers and the original revised writing test design. Building on the methodology described in Eckes (2005), a questionnaire study was carried out to investigate the salient features raters attend to when rating scripts and to what degree raters attend to those features. The questionnaire was administered to raters following live rating, and the data obtained was analyzed using multi-facet Rasch analysis (FACETS 3.67, Linacre, 2010). The results showed that raters are attending to the features intended by the scale developers and to the degree intended. Furthermore, it was found that the criterial features to which raters are attending fall into three distinct groupings based on the level of importance raters place on them, and that these groupings correspond to three broad categories of writing performance: communicative effectiveness, linguistic resources, and mechanics. These groupings are also consistent with the intentions of the test developers, and raters for both tests broadly allocated the same criterial features in the same way to these three groupings. Some differences between the two tests were identified, however, and it is posited these can be linked to differences in the tasks. This presentation will introduce the study background and explain the procedures followed, from questionnaire design to data analysis methods. The presenter will also give the results of the study and discuss how similar studies may be useful in the development and revision of rating scales.

Room 1 (Room 618) Part VI

Rater Judging Plans and Sensitivity to Unexpected Scores

Jeffrey DURAND (Tokai University)

kandajeffd@gmail.com

In many tests, performance on tasks like holding a conversation or writing an essay is important. In general, these tasks require raters to judge the performance. With a large number of test takers, it may not be feasible for all raters to judge each performance. The result is that a couple, or even just one, rater may be the judge of a performance. This can be problematic in that raters generally differ in how strict they are. The results of test takers performances may depend as much on the strictness and consistency of their raters as on their abilities.

Multi-Facet Rasch Measurement (MFRM) (Linacre, 1994) was developed, in part, to deal with this issue. This measurement technique allows ratings of ability to be adjusted for rater strictness. It also allows the identification of unusual or unexpected scores. The analysis and detection of problems depends, however, on being able to compare all raters in a “network.” Two raters can have judged the same students, providing a direct comparison of their strictness. They could also both have worked with a third rater, though not with each other, providing an indirect comparison of their strictness. However, if one group of raters works together to judge one set of students while another group of raters works independently of them with a different set of students, it is not possible to adjust for strictness across these two groups.

While Multi-Facet Rasch Measurement requires a network of raters, it does not specify the exact nature of that network. The way that raters are assigned to judge the same performances, the judging plan, is flexible as long as all raters are part of the network. If all raters judge every performance in a consistent manner, the judging plan is not so important. In practice, raters are not perfectly consistent, and it is not always clear how the inconsistencies affect performance scores. In this situation, the ability to detect problems and provide fair scores may depend on the judging plan.

This simulation study investigates different judging plans and the effects that unexpected ratings have on all performance scores. It also investigates the ease of identifying unexpected ratings under different judging plans. The goal of this research is to improve judging plans and predict when certain judging plans could lead to errors in measurement.

Reference

Linacre J.M. (1994). *Many-Facet Rasch Measurement*, 2nd Ed. Chicago: MESA Press.

Room 2 (Room 619) Part I

A comparison of the EPER Form A and Form E: do they work as alternative forms?

Kiyomi YOSHIKAWA (Kansai University)

yoshizaw@kansai-u.ac.jp

Atsuko TAKASE (Kansai University)

Kyoko Otsuki (Hiroshima Shudo University)

Extensive reading (henceforth ER) has been recognized as one of the effective methods for improving language abilities of ESL and EFL learners. It is essential for a successful ER program that learners read materials appropriate for their reading levels. In order to decide the appropriate reading levels for learners, the Edinburgh Project on Extensive Reading (EPER) test (cloze test) has been administered in many ER classes. Although three forms of the test are available, i.e. Forms A, B and E, Form A is most widely used. These three forms are supposed to serve as alternative forms of the EPER test. However, the instructors who use these forms perceive that Form E is more difficult than Form 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.

The present study investigates whether EPER Form A and Form E can be treated as alternative forms. Two types of analysis were conducted: qualitative and quantitative. In the qualitative analysis, a content analysis of Form A and Form E was conducted. Along with the length and readability of the passages in each form, deleted items were examined in terms of the levels of knowledge learners need to fill in the blanks and the grammatical features. In the quantitative analysis, the difficulty levels of the two forms were examined. To this end, two forms were linked using a common-item design. Using the results of the content analyses of the two forms, 35 common items were created based on the three new passages. Those items were merged into the two forms. A total of 537 students at four four-year universities in the western Japan participated in the study. There were two groups of students: one group took Form A with the common items and the other group took Form E with the common items. Rasch analysis was conducted to analyze the data files of the two forms. Then, they were linked and the difficulty levels of the two forms were analyzed. The presentation includes the results of the content analysis, those of equating, and educational implication.

Room 2 (Room 619) Part II

Monitoring Extensive Reading using Mobile Phones

Trevor A. HOLSTER (Fukuoka Women's University)

trevholster@gmail.com

William R. PELLOWE (Kinki University)

J W. LAKE (Fukuoka Women's University)

Extensive reading programs aim to increase students' reading automaticity through processing large quantities of text. To achieve this, students should choose books that are both interesting and of the appropriate difficulty. Common ways of monitoring student reading include quizzes, book reports, and word counts. However, these may result in intensive (rather than extensive) reading, which in turn reduces long-term motivation. Instead, the presenters developed an extensive reading module for an open-source audience response system. Students report which books they have read, and rate the books' difficulty and interest level. Classroom teachers can access summaries of the number of books each student has read, as well as popularity ratings for the books. Those doing research can easily access data files formatted for many-faceted Rasch analysis, providing measures of the reading ability of individual students and difficulty of book titles. Piloting of the system in 2012 provided data from 122 students and 197 books, with reliability coefficients of .91 for students and .84 for books, sufficient to inform text recommendations and future purchasing decisions. The Japanese Ministry of Education, Culture, Sports, Science and Technology provided a research grant to develop and operationalize the system. Results from operational use of the system in the first semester of 2013 will be presented, comparing the difficulty of books as rated by students with publishers' claimed difficulty levels. Predicted gains in reading speed will be investigated using pre-test and post-test measures of reading speed based on Quinn and Nation (1974).

Room 2 (Room 619) Part III

The Relative Utility of Three Distinct Picture Prompts for Eliciting Language and Ideas

Jungok BAE (Kyungpook National University)

jungokbae@knu.ac.kr

Pictures are the most spontaneous and cross-cultural source of meaning creation, narration, and language use. The use of pictures for eliciting language and ideas has always been attractive and appropriate in a globalizing society. This study presents three distinct types of picture prompts developed and used one after another over the years to evaluate the language skills of youths and to select those gifted with verbal creativity. The study aims to compare their utility and difficulty levels.

Three Prompts. The three types of the picture prompts manipulated for the purpose of the study were as follows: Version 1 ('Series') had sequenced events with five scenes; students were asked to write a story based on the sequence. Version 2 ('Predict') had the first scene taken from the 'Series' version, and students were asked to write a story, imagining what was happening now, and what would next. Version 3 ('Connect') had seven isolated objects taken from version 1 ('Series'), and these objects were presented randomly on one page; students were asked to create a story using at least five of these objects. 'Series' and 'Predict' are commonly used formats, and 'Connect' is a new format developed recently by the author. Across these types, writers were told to use their imaginations, and they were given 30 minutes for composing.

Participants, Test Administrations, and Analysis. The participants were EFL students in grades 3 to 6 ($N =$ about 180) enrolled in an elementary school that implements immersion-based language education. The three versions were distributed simultaneously to the students by random sequential assignment; this procedure generated three different prompt groups, each of which received a different version. The stories were evaluated by both human readers and Coh-Metrix software. Data analysis was performed with MANCOVA and also qualitatively. The means for writing qualities were compared across the prompt groups, first for the entire data set and then by grade level, classified into lower (grades 3 and 4) and upper levels (grades 5 and 6).

Results. The findings include the following. First, word count, readability, grammatical complexity had no statistical differences across the three prompt groups, and this equality held true for both grade levels. Second, the mean scores for originality from highest to lowest ranked the three prompt groups as follows: Connect, Predict, and Series for the entire data. Third, the coherence mean was the highest for the Connect group for the lower grade level. For the upper grade level and the entire data set, however, coherence means did not vary across the prompt groups. Fourth, the Connect prompt generated the largest number of storylines followed by Predict and then Series with the smallest number of storylines. Finally, all of the stories based on one prompt were put together, forming three corpora of stories, each created from a different prompt. Vocabulary diversity was the same across the three corpora for lower graders. However, for upper graders, greater vocabulary diversity was detected in the Predict corpora.

Implications. The implications of these findings to be drawn will help test developers and book authors who utilize pictures to assess, teach, and inspire language and creative thinking skills. Because pictures are a universal tool, the findings will be applicable globally across assessments in different languages.

Room 2 (Room 619) Part IV

Cross-actions between Linguists and Psychometricians make possible reliable Tests for multiple Examinee profiles

Sébastien GEORGES (International Centre of Pedagogical Studies, France)
georges@ciep.fr

The Centre international d'études pédagogiques (CIEP) is a national public operator for the Ministry of Education and Higher Education, and the Ministry of Foreign Affairs. It aims mainly at promoting French language. For this purpose, the CIEP manages and designs the administrative and pedagogical features of the DELF-DALF exams and of the TCF -French as a foreign language certificates-. These French language tests measure a person's level of French for different purposes. All the tests are aligned with the six levels of the Common European Framework of Reference for Languages (CEFRL). The TCF enables candidates to see what level they have reached in reading, listening, and/or writing, and/or speaking. It is intended for people who are not native French speakers and who, for professional, personal, immigration (e.g., to Quebec), citizenship (e.g., in France) or academic purposes (e.g., in 8 Japanese universities), wish to have their abilities in French assessed in a reliable, simple and quick way. The CIEP has designed as many assessment tools as requested by policy and decision-makers for specific age groups and goals. These computer-based and/or paper and pencil tests –administered roughly to 500 000 people in more than 175 countries worldwide each year– include multiple choice questions, matching tasks and/or open-ended questions. A linguistic profile of the examinees in the 4 skills can be inferred from the test results.

The aim of this communication is to show:

- I) how the CIEP designs these tests to deliver a relevant answer from the initial request to the final users;
- II) the common and the specific procedures linked to the different tests;
- III) how the pre-, post- and seeding-tests are used to maximize the item bank size as a function of item exposure and the number of test sessions, and finally
- IV) how the results of psychometrical analyses -across Classical Test Theory, one parameter Rasch model for dichotomous items, or Partial Credit Model for polytomous items- a) are jointly used by psychometricians and French as a foreign language experts to make valid, reliable, sensitive, and fair assessments and b) enable to design parallel tests versions which confer independency between examinees' results and contents' items and tests.

We will also illustrate this last point by showing how we designed a standard-setting to get a single test aligned both with the CEFRL and the Canadian Language Benchmark to comply with a specific request from Canada.

Room 2 (Room 619) Part V

The Degree of Confidence in Distractors in Multiple-Choice Listening Tests

Hideki IIMURA (Tokiwa University)

iimurahideki@gmail.com

This study investigates how distractors function in multiple-choice listening tests. Usually each distractor has been evaluated by its attractiveness. Namely, distractors that can succeed in attracting many test-takers are considered to perform well on the test. On the other hand, distractors which end up attracting only a few test-takers are recognized as performing poorly. Thus, those distractors judged to be unattractive might be a target for restructuring.

Given that test-takers have to choose only one response among several options, it is possible to assume that some distractors could have functioned well even if they had not been selected. In other words, we should evaluate each distractor's level of attractiveness before test-takers select one option.

This study presents data from the questionnaire in which a self-rating scale of confidence in both choosing a correct option and not choosing incorrect options (i.e., distractors) has been developed and tested. The researcher tries to explain the test-taking process of eliminating distractors and reevaluates the distractors that are not chosen by test-takers.

Room 2 (Room 619) Part VI

Breaking the Mould: Localisation and Validation in Test Development

Barry O'SULLIVAN (British Council)

barry.o.sullivan@britishcouncil.org

The concept of localisation in the area of technology has been with us for some time, its discussion in language test development only really began in earnest with O'Sullivan (2011). In this paper I will focus on the theoretical link between localisation and a model of test validation – establishing an initial broad framework for further study and understanding of the concept. Broadly speaking, localisation is defined here as attending to the requirements of a specific test taker or test taking population in an appropriate way. Within the validation model, this essentially means consideration of not just the test taker (representing the cognitive dimension of language ability) but also of the test itself (representing the social dimension of language use) and the scoring system (which should 'fit' theoretically and philosophically with the other dimensions).

Following on from this more theory-driven section of the talk, I would then like to demonstrate how the concept of localisation has been built into a new language testing service called Aptis. Aptis was developed by the British Council over a two year period and launched in 2012. Since then, it has been used in over 40 countries, often with significant local input. The service is built around a number of key concepts, including validity, flexibility, accessibility and localizability. While this talk will focus on the latter concept, the others will also be touched upon in order to present a more complete picture of the approach taken in this new test.

Room 3 (Room 717) Part I

Creating an in-house computerized adaptive testing (CAT) program with Concerto

Atsushi MIZUMOTO (Kansai University)

atsushi@mizumot.com

Many researchers and practitioners who engage in language testing understand the importance of Item Response Theory (IRT) and Rasch Modelling for better linguistic competence measurement. Computerized Adaptive Testing (CAT) utilizes these theories and is an ideal way to administer a test, measure the ability of test takers, and give them feedback. With almost 40 years of use in research, CAT has a well established framework among testing specialists (Thomson & Weiss, 2011). This is why many large-scale tests have adopted IRT-based CAT.

However, even for researchers and practitioners who have some knowledge of IRT and CAT, it is almost impossible for them to independently develop and administer CAT because it requires substantial technical abilities. Except for a few innovative Moodle plug-ins developed by Kimura and Akiyama (2009), the majorities of us have had to depend on commercial test publishers for CAT.

To suggest ways for researchers and practitioners to develop their own CAT programs, I will introduce Concerto, an open-source online R-based adaptive testing platform. Concerto was developed by The Psychometrics Centre, University of Cambridge (<http://www.psychometrics.cam.ac.uk/page/338/concerto-testing-platform.htm>). With some knowledge of the HTML language and R (statistical programming software), researchers and practitioners can create their own CAT programs. It can be installed on to a server, or a free Concerto account can be created on the developer's website (limited to 150 respondents monthly).

In my presentation, I will report how I developed a CAT version of a vocabulary size test and describe the results of pilot testing. Practical issues and implications associated with the CAT development will also be discussed.

References

- Kimura, T., & Akiyama, M. (2009). Reinforcing development of a Moodle-based English test with multiple choice maker, TDAP block, and CAT module. *Language Education & Technology*, 46, 233–245. Retrieved from <http://ci.nii.ac.jp/naid/110008441797>
- Thompson, N., & Weiss, D. (2011). A framework for the development of computerized adaptive tests. *Practical Assessment, Research & Evaluation*, 16, 1–9. Retrieved from <http://www.pareonline.net/pdf/v16n1.pdf>

Room 3 (Room 717) Part II

The factor structure and factorial invariance of a university-based EAP test

Jinsong FAN (Fudan University)

jasonfann@126.com

Though language testing researchers have generally come to the consensus that language ability is a multi-componential trait, mixed findings have been reported by researchers as to what these components represent and the relationships between these components. More empirical studies are therefore warranted to further investigate the construct structure of language ability by analyzing the data of language tests developed and used in different contexts. This study investigated the factor structure and factorial invariance of the Fudan English Test (FET) with structural equation modeling, a powerful data analysis method which integrates multiple regression, path analysis, and confirmatory factor analysis. The FET is a high-stakes English for Academic Purposes (EAP) test developed and used within Fudan University in Shanghai, China. It is expected that this study can present some empirical evidence as to the construct validity of the FET, and at the same time, help language testers and educators better understand the nature of language ability.

The data in this study were the test scores of 800 students on the different sections of the 2012 FET administration with 361 males and 439 females. To investigate the FET factor structure, five theoretical models were posited *a priori*, including a bi-factor model, a correlated first-order four-factor model, a first-order one-factor model, a higher-order factor model, and a first-order two-factor model. These five hypothesized models were subsequently tested for model fit through examining their model fit indices, including, for example, χ^2/df , GFI, AGFI, CFI, and RMSEA, as well as checking their parameter estimates and model parsimony. In addition, Chi-square difference tests were performed of the nested models. The results indicated that among the five hypothesized models, the higher-order factor model best fit the FET data. To investigate whether this model could maintain factorial invariance, multi-group analyses were performed on male (N=361) and female (N=439) test candidates. Measurement weights, structural weights, structural covariances, structural residuals, and measurement residuals were posited to be equal between the models of male and female groups, and Chi-square difference tests as well as CFI difference tests were performed to investigate whether factorial invariance could be maintained across the two groups. The test results indicated that the model was group invariant at all five levels, suggesting that the moderator variable, sex, had no effect on the hypothesize model. The findings of this study lend support to the current practice of the FET reporting both the composite score and the four profile scores. Meanwhile, this study also presents further empirical evidence as to the multi-componential nature of language ability.

Room 3 (Room 717) Part III

Examining Local Dependence in the Iranian National University Entrance Examination

Mehrasa ALIZADEH (Alzahra University)

mehrasa.alizadeh@yahoo.com

Local independence, as a primary assumption of item response theory (IRT) models, tends to be violated in reading comprehension tests which comprise several short passages, each one followed by a number of questions. This study examined local dependence (LD) in the Reading Comprehension section of the Iranian National University Entrance Examination. Test performance data of 212 test takers were selected from the responses of 133,832 Iranian test takers who sat for the language-major Entrance Examination in July 2010. The data were subjected to the Rasch model which provided evidence of local dependence. It was also found that collapsing reading comprehension items into a super-item for each passage and the subsequent use of the partial credit model for data analysis rectified the problem of local dependence. The findings of the study have implications for passage-related LD and test development.

Room 3 (Room 717) Part IV

Which is Better: Writing or Choosing in Spelling Tests? Comparing Recall with Recognition Test

Sachiyo TAKANAMI (Reitaku University)

sachiyotakanami@gmail.com

The purposes of this study are (a) to compare performances on recall and recognition tasks in English spelling tests and (b) to consider the difficulties of spelling tests for Japanese EFL learners.

Some have asserted that there are three symbolic codes in English words: sounds, letters, and meanings. These codes are essential to the tasks of describing an idea through spoken language (i.e., sound), writing language (i.e., letter), and understanding spoken and written language (i.e., meaning). That is to say, acquisition of the knowledge of these essential codes is the first important step for language learners.

As described in his book, Nation (1990, 2001) presents three major aspects of vocabulary knowledge: form, meaning, and use. In addition, each aspect has three different subdivisions, respectively, and two more subdivisions according to knowledge type (i.e., receptive and productive knowledge). In short, there are 18 aspects of vocabulary knowledge. Combining Nation's definition of vocabulary knowledge and the three language codes described above, 12 tests were conducted for this research.

The first half of the six tests consisted of recall tests, and the other half involved recognition tests. Two sets of the tests were basically designed to measure the same aspects of vocabulary knowledge. The six aspects are as follows: (a) sound to letter, (b) sound to meaning, (c) letter to sound, (d) letter to meaning, (e) meaning to letter, and (f) meaning to sound.

The results showed that, as expected, learners' performance on the recall tests were inferior to the recognition tests. Also noteworthy were the difficulty levels on the recall tests. Most of the learners had difficulty writing the correct spellings (i.e., meaning to letter and sound to letter) or producing correct pronunciation (i.e., meaning to sound). Another noteworthy result was gathered from the recognition test. Learners had difficulty recognizing (or choosing) the correct spellings. Producing or recognizing English spellings, more specifically alphabet letter strings, were considerably more challenging tasks for the Japanese EFL learners, even though the tested words were already learned words. Thus, EFL learners' lack of spelling knowledge is considered to be a problem that needs to be addressed.

Room 3 (Room 717) Part V

Development of a Test to Measure Knowledge Regarding Multiple Meanings of Basic Words

Yuko HOSHINO (Tokyo Fuji University)
yukohoshino@live.jp

The volume of research conducted concerning vocabulary acquisition is increasing. However, a limited number of studies have focused on the manner in which knowledge regarding basic words develops. The fact that the most frequent 2000 word families cover about 80% of running words (Nation, 2001) suggests that these words play an extremely important role. Because they appear rather frequently, they often have multiple meanings, and therefore, it is difficult to acquire complete knowledge of word meanings. Wolter (2009) called this phenomenon meaning-last acquisition, which indicates that the meanings of words appearing with various collocated words are understood later than other aspects of vocabulary knowledge. Furthermore, L1 and L2 meanings of words seldom match exactly, which increases the difficulty of learning meanings. However, past research concerning basic words has often dealt with other aspects (e.g., association and collocation) and neglected their meanings. This study focuses on meaning and investigates the manner in which various meanings of words can be measured effectively. Seven basic verbs with multiple meanings are presented in two types of contexts (collocation and sentence). The participants answered either one of the tests in context, and they were asked (a) to decide whether each expression is in correct English, (b) to translate each expression, and (c) to identify which of the expressions have the same meaning out of six expressions for each verb. The results suggested that the correct rates did not differ between the two types of contexts and their correlation coefficients with estimated vocabulary size and Oxford Quick Placement Test (OQPT) were very low or negative. Hence, the concept of knowledge regarding multiple meanings is different from knowing one core meaning (vocabulary size) and finding suitable words to contexts (OQPT). Therefore, it is necessary that tests to measure knowledge regarding multiple meanings for basic words are developed.

Room 4 (Room 718) Part II

言語テストにおける視覚障害者受験特別措置実施の目的と方法－日本語能力試験点字冊子試験を例に－

河住有希子（日本工業大学）

y_kawasu@nit.ac.jp

藤田恵（日本国際教育支援協会）

秋元美晴（恵泉女学園大学）

本研究は日本語能力試験（以下 JLPT）視覚障害者受験特別措置である点字冊子試験を題材に、「言語テストの点字冊子試験」として共有可能な情報を集約し、言語テストの更なる充実に役立てることを目的とする。

現在、「障害者の権利に関する条約」（07年に外務省が署名）批准に向けての準備が進められている。同条約第二十四条教育5には「締約国は、障害者が、差別なしに、かつ他の者と平等に高等教育一般、職業訓練、成人教育及び生涯学習の機会を与えられることを確保する。このため、締約国は、合理的配慮が障害者に提供されることを確保する」とある。学習の状況・成果を測定し、自己理解、学習意欲喚起、教養の涵養などに寄与する言語テストにおいても、あらゆる受験者が合理的配慮のもとで受験できるよう、試験問題の設計及び受験環境の整備に取り組む必要があると言えよう。文部科学省が2011年に取りまとめた「検定試験の自己評価シート」に「障がい者について、検定試験の目的や内容、規模等に応じた一定の配慮が考慮されている」という項目が立てられたことから、特別措置への要請が高いことが伺える。

本発表ではまず、日本国内で実施されている各種言語テストの点字冊子試験実施状況を概観する。次に JLPT 点字冊子試験受験経験者への聞き取り調査の結果から、点字冊子試験実施の意義を述べ、最後に JLPT 点字冊子試験の実施方法と改善に向けての検討の枠組みを「言語テストの点字冊子試験」という観点から示す。

点字冊子試験実施に関連する議論は藤芳（1996）に始まる一連の研究などごく限られている。しかし、グローバル化が進む社会において、言語テストは、信頼性・妥当性の保証のみならず、すべての受験者に、平等に受験の機会を提供することが、世界と同水準で求められるようになるであろう。本研究は、受験特別措置の充実に向けて、一定の役割を果たすものとする。

Room 4 (Room 718) Part III

CEFR に準拠したジャンル別ライティング及びスピーキング評価の開発と検討

長沼君主 (東海大学)

n.naganuma@tokai-u.jp

高野正恵 (東京外国語大学)

ヘザー・ジョンソン (東京外国語大学)

工藤洋路 (駒沢女子大学)

東京外国語大学英語学習支援センター (TUFS English Learning Center) では、ヨーロッパ言語共通参照枠(CEFR)に準拠した言語パスポートの発行を行っている。そのうち、ライティング評価では、ケンブリッジ大学 ESOL との提携による English Profile Programme の評価タスクをベースとし、Eメール文、物語文、論説文の評価開発を行ってきた。CEFR の Written Assessment Criteria Grid などの関連資料を参考に、それぞれのジャンルごとに独立した内容面の採点基準を設け、正確さや一貫性といった共通した言語面の評価と合わせてレベル判断を行った。

一方、スピーキング評価では、ACTFL-OPI に基づいた SST(Standard Speaking Test)の評価タスクを参照し、一部、ケンブリッジ英検の面接評価タスクを加えるなどしながら、課題解決場面を設定したロール・プレイ課題、ひと続きの絵や写真に基づいた物語描写課題、対となるイラストを対比させながら意見を述べる意見陳述課題を設けた。それぞれのタスクは複数の難易度が設定されており、インタラクティブに受験者のレベルを判断しながら出題した。採点基準としては、ACTFL Proficiency Guidelines を参照しながらも、CEFR の Oral Assessment Criteria Grid などの関連資料を参考に、独自の評価開発を行った。レベル判断にあたっては、包括的評価に基づきつつ、タスクごとの分析的評価も合わせて行った。

ライティング評価はこれまで主副専攻英語授業内で年 1 回実施してきた。スピーキング評価は実施上の負荷の点から任意受験形式で年 2 回実施してきた。本研究では 2009 年度から 2012 年度にかけてのそれぞれのジャンル別評価データをもとに、それぞれのジャンル間のテキストタイプに応じた能力の関連性ととも、ライティング評価タスクとスピーキング評価タスクで測定している課題ごとの能力が、概ね対応しているとの仮説のもと、その相互の関連性を探る。また、ジャンル別の評価開発にあたっての問題点を議論しつつ、タスクに合わせた評価基準開発のあり方やライティングとスピーキング採点基準の相互の関連についても考えたい。

Room 4 (Room 718) Part IV

Building an Assessment Use Argument for Matriculation English Test (Beijing Version): An investigation into refining the reading construct and items

Kaizhou LUO (Beijing Foreign Studies University)

kevinlkz@hotmail.com

Baocheng HAN (Beijing Foreign Studies University)

The only way for most Chinese high school students to go to college is to pass college entrance examination. English is one of the mandatory subjects no matter whether a student sits for the national matriculation examination or the one designed by local testing authority. However, a large number of studies show that, on the one hand, the matriculation English test with its high-stakes nature has been definitely exerting unintended influences (negative washback) on normal high school teaching and learning, on the other, there really exist a huge number of students whose English performance does not meet the college requirements for academic communications. Therefore, more and more stake holders urgently call for the reform of matriculation English tests.

A matriculation English test reform project sponsored by Chinese Ministry of Education is underway. Adopting Bachman and Palmer's Assessment Use Argument (AUA) as theoretical framework for guiding test development and justification, this reform project stands poised to change the exam-oriented *status quo* of high school English teaching and learning through refining the construct, redesigning the test format and justifying the intended use in readiness for college academic study.

This paper deals with a preliminary investigation into the gap between the construct as measured by the present reading module in matriculation English test (Beijing version) and the real reading activities and experiences of the first/second-year college students'. Besides, it also reports the results of need analysis on the academic reading performance and specific tasks expected by college teaching staff from different disciplines. Researchers hope that results can help to make the initial claims for the "interpretations" within AUA. Detailed information and practical implications are now being analyzed and will be reported in the presentation. All suggestions on this research are welcome.

Room 4 (Room 718) Part V

Justifying the Use of an English Language Placement Test with an Assessment Use Argument

**Parisa MEHRAN (Alzahra University)
parisa_mehran@yahoo.com**

This study was an attempt to justify the use of an English language placement test, which is composed of the Oxford Quick Placement Test (the OQPT) and the follow-up oral examination, based on Bachman and Palmer's (2010) assessment use argument (AUA) framework. To do so, different kinds of relevant evidence, both quantitative (classical item analysis, estimations of reliability, and correlational analyses) and qualitative (interview and observation), which supported the claims and weakened the potential rebuttals, were collected and analyzed. The study was conducted at one of the English language institutes in Tehran, Iran. Three hundred and thirty two newcomers to the institute who took the placement test participated in this study, and 15 of them were interviewed. The head of the institute, three examiners of the placement test, ten teachers, and four experts also attended the current study. The results contributed both positive and negative attributes to the validity argument for the placement test. Based on the gathered evidence, this study found that the assessment records of the OQPT and the oral examination are consistent across different assessment tasks, different aspects of the assessment procedure, and across different groups of test takers. However, the oral examination requires a set of criteria. The findings also indicated that the OQPT scores and the oral examination results can be interpreted somewhat as test takers' level of English proficiency and place them in their appropriate levels. Such interpretations are meaningful, impartial, relevant, and sufficient, although lack of a listening section in the OQPT and lack of a rubric for the oral examination can be threatening, and generalizability of the results is to some extent under question. In addition, the placement decisions that are made on the basis of the OQPT scores and the oral examination results are not sensitive to local values and equitable to all stakeholders due to the subjectivity of the oral examination and the economic considerations of the institute. Lastly, by and large, the consequences of the placement decisions based on the OQPT scores and the oral examination results are beneficial to all stakeholders that use the test, including the test takers, the institution, the teachers, and the supervisor. The findings of this study have local implications for the institute which administers the placement test. Furthermore, it serves as an illustration of the merits/demerits of using an AUA.

Room 4 (Room 718) Part VI

Designing an assessment architecture: Evidence-centered design and object-oriented methodology intertwined

Hossein KEYNEJAD (Islamic Azad University)

Mehrasa ALIZADEH (Alzahra University)

mehrasa.alizadeh@yahoo.com

The evidence-centered design (ECD), as proposed by Mislevy and his colleagues, defines any assessment system in terms of a four-process architecture. This article aims at employing the ECD as a framework to set up an online testing center, which is capable of supporting various assessment types. Once combined with ECD, the object-oriented (OO) methodology can be made use of in designing assessment components. Furthermore, reusability, as a basic feature of OO methodologies, guarantees the most appropriate relationships among assessment components. Assessment components themselves are defined in terms of classes which own a set of attributes and methods. Inheritance and polymorphism also establish well-defined relationships among classes. Thanks to UML, assessment components are modeled, and roles, activities, artifacts, and workflow are designed. This is the first step in designing and delivering any assessment system. In summary, the OO methodology along with its distinctive features can act as an appropriate model for assessment components, such as the student model, evidence model, and task model, in the conceptual assessment framework (CAF).

Room 5 (Room 719) Part I

ディクテーションテストにおける誤りの傾向

小林夏子 (株式会社教育測定研究所)

kobayashi@jiem.co.jp

野上康子 (株式会社教育測定研究所)

吉川厚 (株式会社教育測定研究所)

林規生 (株式会社教育測定研究所)

ディクテーションは、言語習得の学習方法として、また、学習者の習熟度を測定する手法として利用されてきたが、その評価については確立された方法が未だない。多くの場合、手採点における複雑さを回避するという作業上の理由により、採点対象とする語彙ごとの完全一致で評価されている。しかし、誤りと判断された解答の内容を単語単位で吟味すると、単語の音としては聞き取れているが解答の際にスペルを間違えた場合、音自体が聞き取れていないがその箇所に何らかの語が存在したことは認識していた場合、語の存在自体を認識できなかった場合など、様々なケースが考えられる。また、連続する複数の単語で誤りが生じている解答の中には、語順が入れ替わってしまった場合や2語がつながって聞こえてしまった場合などが考えられる。このように誤りの内容は多種多様であり、誤りに関してより詳細な分析に基づいた評価が求められている。

本研究では、コンピュータで採取した解答データを分析し誤りの傾向を調べた結果を報告する。また、その誤り傾向と受験者の英語力との関係について調査した結果を報告する。

Room 5 (Room 719) Part II

高校生の語彙数はどのように伸びていくのか—英検取得級別に見た語彙サイズテストの成長曲線モデルによる分析

船越貴美 (神戸学院大学附属高等学校)

funakoshiy@kcc.zaq.ne.jp

旧学習指導要領で定められていた中学校の指導語数は900語、高等学校では1300語～1800語となっており、中学・高等学校を通じて2200語～2700語が必修単語数として学習することとなっていた。しかし、大学入試に必要な単語数は、大学入試センター試験で約3000語、難関大学で4000語～5000語であり、教科書だけでは語彙数を増やすことはできない。

本研究では、高校生の語彙数が3年間でどのように伸びていくかを調査するために、実験に参加してくれた高校生に年1回、望月(1998)による語彙サイズテストを受験してもらい、3回の結果を英検取得級別に成長曲線モデルで分析した。結果は、高校3年間で生徒の語彙数は時間の経過とともに増えているが、英検取得級が上級になるほど、語彙数の増加が著しかった。特に、高校卒業時までに英検2級を取得した生徒が最も語彙数を伸ばしており、英検取得のための学習が結果的に語彙数を増やしたと思われる。一方で、英検を積極的に受験しなかった生徒でも、大学受験のために語彙数を増やす努力をした生徒は、2年次から3年次にかけて急速に語彙数を増やしていたことがわかった。

望月正道. (1998). 「日本人学習者のための英語語彙サイズテスト」『語学教育研究所紀要』 第12巻, 27-53.

Room 5 (Room 719) Part III

高校生学習者の英語学習動機と自己評価の関係

宮本由美子 (長野県上田染谷丘高等学校)

myumiko@po6.ueda.ne.jp

近年、教師の評価や外部からの視点からの評価に加えて、学習者自身が自己の「振り返り」によって成長できるという、自己評価の役割が注目されている。しかしながら、自己評価は教育的意義が認められながらも「教員が行う生徒の評価資料としないこと」(文科省)など補完的なものにとどまり、現場で広く活用されているとは言い難い。自己評価と実際の能力との関係については、過小評価や過大評価の傾向が経験上指摘されるが、現場では当然あることと、問題視されないことが多い。英語学習動機については、英語運用能力(GTEC)とどのような関係があるか、高校生を対象にした研究は非常に少ない。以上の点から、高校3年生278名に対し、同時期にGTEC for STUDENTS (以下GTEC)と質問紙調査(英語学習動機、英語各技能ごとの自己評価)を実施し、以下の点について考察する。

- (1) 自己評価は英語運用能力(GTEC)と、どの程度一致しているか。一致していない場合は、どのような特徴が見られるか。(例) A:自己評価・GTECのスコアともに高い、B:自己評価は高いが、GTECのスコアが低い、C:自己評価は低いがGTECのスコアが高い、D:自己評価、GTECのスコアともに低い。(結果は当日発表する)
- (2) 英語学習動機 質問紙:「英語の勉強が大切だと思う理由」に関しては9項目中3項目を複数回答させた。9項目のうち類似の項目を結合し、「就職」、「海外渡航」、「国内外のグローバル社会への意識」、「高校の成績や大学受験」とまとめ、割合を計算し、クラスター分析を行った。クラスター数は3と解釈し、クラスター毎にGTECのスコアについて記述統計を行い、クラスター間平均点の分散分析をGTECの技能別に行った。動機については4つの動機ともバランスの良いクラスターにおいて、全てのスキルと総合得点において平均点の差が有意である結果が得られた。

Room 5 (Room 719) Part IV

異なる文章提示方法によってスピーキング再話課題の発話特徴に差は生じるのか

横内裕一郎 (筑波大学大学院)

u16yoko@gmail.com

青島健夫 (筑波大学大学院)

学習指導要領では4技能の総合的な向上を、特にコミュニケーション能力の向上が叫ばれる中、教室内でスピーキング能力を評価する機会は少ない。このような環境でスピーキング能力を評価する際に想定されるタスクとして再話課題があるが、この課題を行う際に、情報をどのように提示すべきなのだろうか。情報提示方法は、リスニングによる場合とリーディングによる場合と考えられるが、リスニングによるインプットは、「語彙、文法、音韻などの知識を同時に活用して、瞬時に言語処理を行う極めて能動的な力(高梨、2009)」であるため、読み返しが可能なリーディングよりも困難度が高くなると考えられる。本研究では、リスニングあるいはリーディングによる情報提示の後に、スピーキングによる再話課題を行わせた場合の発話の特徴について調査を行った。

日本人大学生 56 名を対象にリーディングインプットの再話課題とリスニングインプットの再話課題を課した。これらの発話を(1)有効語数(Token 数)、(2)Types 数、(3)Guiraud index、(4)WPM の4観点と、客観評価による結果を対応のある t 検定で分析を行った。その結果、(1)、(2)、(4)では有意にリーディングインプットの場合に数値が高くなったが、(3)については有意な差が見られなかった。また、発話を分析的に客観評価し、それらの結果を比較したところ、伝達能力、文法と語彙、内容、発音の各項目でリーディングの方が有意に高い評価が得られた。

この結果から、リーディングベースのインプットの方が、学習者の発話をより多く引き出すことがわかり、発話の質(使用語彙数、流暢さ、客観評価)も高かったことがわかった。今回の実験では、リーディングとリスニングマテリアルの難易度に差があったため、この結果を一般化することはできないが、リーディングをインプットとした再話課題を用いることで、より学習者のスピーキング能力を評価するのに必要な情報を引き出すために有効であると考えられる。この結果に基づいて、再話課題の利用について考察する。

Room 5 (Room 719) Part V

EFL 学習者における語順知識とスピーキング能力間の関係性

佐瀬文香 (筑波大学大学院)

s1220041@u.tsukuba.ac.jp

現行の学習指導要領では、「話すこと」及び「書くこと」に関する技能を中心とした「英語会話」という科目が新設され、これまで以上にアウトプットの能力の育成や測定が求められている。田地野 (1999) は語句の配列順序の重要性を述べ、語順によって文の意味が左右されると主張した。語順整序問題は大学入試等でも使用されていることから、こうした知識の重要性がわかる。そこで、本研究では、整序問題で問われる語順知識と、スピーキング能力の関連性について調査を行った。

本研究では大学生 69 名を対象に、再話課題と、課題文中で使用されている文法項目を中心とした整序問題の 2 タスクを作成した。再話課題を行った後、その音声を録音し Hirai and Koizumi (2008) の EBB 尺度を使用して伝達能力・文法と語彙・発音の 3 つの観点から評価を行った。その後、全 15 問の整序問題を実施した。そしてスピーキング評価と整序問題の解答とを相関分析にかけて検証した。

結果として、評価の 3 観点それぞれにおいて相関が見られなかったため、スピーキングパフォーマンスと整序問題で問われる語順知識との間に関連がほぼ見られないということがわかった。その後、整序問題正答者及びスピーキング評価の各上位 30% の相関を産出したところ、全体よりもこちらの相関の方が高かったが、有意な結果は見られなかった。しかし、本実験では協力者の熟達度がほぼ等質であり、句間語順や句内語順という語順知識の 2 分類(山岡, 2000) を考慮に入れていない。また、整序問題を解答する際には時間が与えられる一方で、スピーキング産出の際にはそのような時間はほとんど与えられないという即時性の問題も挙げられる。そこで、これらを考慮に入れてさらなる調査を行った。再話課題及び、語順知識を問う問題を課し、それぞれの結果の比較を行った。当日の発表では詳しい分析結果及び考察を紹介する。

Room 5 (Room 719) Part VI

日本人大学生の英語文法性判断力～試験的調査～

徳永美紀 (福岡大学)

tokunagamiki@fukuoka-u.ac.jp

本研究は、日本人大学生が英文の文法性をどのように判断しているのかを検証しようとするものである。第二言語習得において、目標言語に関する明示的知識と暗示的知識を測定する方法として文法性判断テストが使われてきた。日本における研究では、共通して日本人英語学習者は明示的知識への依存が大きく、暗示的知識に基づく訂正が少ないという結果がでている。暗示的知識が乏しいということは、英文の文法性を判断する際に「なんとなく違う」というフィーリングによる判断が難しいということになる。しかし、与えられた英文の文法性を判断するという作業は、その結果が文法能力を表すのか、読解力や単語力を表すのか明確ではない。そこで、本研究では意味処理能力などの影響を抑える為に英文に日本語訳をつけ、文法性判断に与える影響を検証する。

この試験的調査では、Ellis 他(2009)で使用された文法性判断テスト 68 項目の固有名詞を日本人に理解しやすい名前に変更して使用した。対象者は私立大学で文系学部所属する 23 名の学生である。まず、英文のみを提示し、それらの英文が「正しい気がする」か「間違っている気がする」かを問うテストを行った。次に、その中で難易度が高かった 33 項目に日本語訳を付け、再度「正しい気がする」か「間違っている気がする」かの判断を求めた。最後に、33 項目の中で間違っている英文 21 項目を対象に、英文の訂正を行った。

予測に反し、日本語訳が英文の文法性判断の大きな助けにはならず、日本語訳が悪影響を与えた項目も存在した。間違った英文の訂正では、文法性判断では「間違っている」と判断できたものの訂正が困難である項目と、文法性判断は困難であったが、「間違っている」と言われれば間違っている部分を見つけて訂正出来る項目が存在した。当日は詳しい結果を報告する。

Workshop Information

Title: Fundamentals of Item Response Theory (Conducted in English)

Lecturer: Jeffrey STEWART (Kyushu Sangyo University, Cardiff University)

Aaron O. BATTY (Keio University, Lancaster University)

Chair: Akiyo HIRAI (University of Tsukuba)

- Date: September 22, 2013 (Sunday), 10:00–14:00 (Lunch break included)
- Venue: Waseda University, Waseda Campus, Building 22, Room 617 (6th floor)
- Attendance Fee: 1,000 yen
- Max Capacity: 30 (first-come, first-served basis)
- Prerequisite: Basic knowledge of statistics is recommended.

● Aims

1. To describe the underlying principles of item response theory, detail and compare the properties of common item response models such as the Rasch, 2-parameter, and 3-parameter models, and explain advantages item response theory offers over classical test theory.
2. To gain experience analysing example data sets (provided) under various item response models using the ltm package for R. (Note: attendees should bring their own notebook computers. Both Windows and Mac operating systems will be supported.)

● Procedure

1. Lecture and Group Work
2. Workshop using the ltm package for R.
3. Q&A

● How to register

1. The deadline of the registration is Tuesday, September 10th. (Note: If the workshop does not reach the maximum capacity, the registration on the day of the workshop conducted is allowed.)
2. When you register, provide the information below and email it to Rie Koizumi (Juntendo University) at rie-koizumi@mwa.biglobe.ne.jp [Note: If you write your questions in (3) below, the lecturers may be able to answer them during the workshop.]

Let us know the following information when you register the workshop.

- (1) Your name and affiliation.
- (2) Your experience of using the following statistics. Please list the names.
 - (a) Classical Test Theory
 - (b) Rasch Analysis (Name of the software used)
 - (c) 2 or 3 parameter Item Response Theory (Type and name of the software used)
 - (d) Others ()
- (3) Questions to lectures, if you have. (Optional)
- (4) Request to this workshop, or JLTA workshops in general (Optional)

Workshop Information (ワークショップ情報)

題目：「項目応答理論の基礎 (Fundamentals of Item Response Theory)」(英語で実施)

講師 Jeffrey STEWART (九州産業大学、カーディフ大学)
Aaron O. BATTY (慶應義塾大学、ランカスター大学)
司会 平井 明代 (筑波大学)

日時：2013年9月22日(日) 10:00-14:00 (途中、昼食休憩あり)

場所：早稲田大学早稲田キャンパス 22号館 617教室 (6階)

参加費：1,000円

定員：30名 (申し込み順)

参加条件：基礎的な統計の知識があることが望ましい。

目的

1. 項目応答理論の基本的原理を述べ、ラッシュ、2・3パラメータモデルのような一般的な項目応答モデルを詳述し、比較する。また、古典的テスト理論と比較して、項目応答理論の利点を説明する。
2. Rのltmパッケージを用いて、様々な項目応答モデルでサンプルデータ(配布予定)の分析実習を行う。(注：参加者は、ノート型PCを持参してください。WindowsとMacのPCで実習可能です。)

手順

1. レクチャー及びグループワーク
2. Rのltmパッケージを用いた実習
3. 質疑応答

申し込み方法

1. 申し込みの締切日は、9月10日(火)となります。(定員に達していない場合には当日参加も可能です。)
2. 申し込み時に、下記の情報を小泉 利恵(順天堂大学) rie-koizumi@mwa.biglobe.ne.jp まで e-mail でご連絡ください。(注：(3)で質問を書かれると、ワークショップ中に回答できるようにします。)

■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) 古典的テスト理論)
(b) ラッシュモデル (使用 software 名)
(c) 2、3パラメータの項目応答理論 (種類、使用 software 名)
(d) その他 ()

(3) 何かご質問はありますか。

(4) その他、ワークショップまたはJLTAワークショップ全体に対して何かご要望がありましたらお書きください。(希望者のみ)

Access to the Conference Venue (Waseda University, Waseda Campus)

(会場へのアクセス)

www.waseda.jp/jp/campus/waseda.html

www.waseda.jp/eng/campus/map.html 等参照

1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo (東京都新宿区西早稲田 1-6-1)

- 8 minutes' walk from Exit 3b of Tokyo Metro Waseda Station on the Tozai Line.
- 5 minutes' walk from Waseda Station on the Toden Arakawa Line.
- 20 minutes' walk from Exit 1 of Tokyo Metro Nishi-waseda Station on the Fukutoshin Line.
- 23 minutes' walk from Exit 2 of Takadanobaba Station on the JR Yamanote Line and the Seibu Shinjuku Line.
- 13 minutes by Tokyo City Bus from Takadanobaba Station and Sodai Seimon.

(Take Exit 2 for the bus stop. The buses numbered 「(学02) 早大正門行き」 OR 「(早81 出入) 早大正門行き」 will go to the Main Gate of Waseda University, Waseda Campus.)

Bus fare 200 yen for regular buses

 170 yen for the (学02) buses (Gaku Bus)

Bus schedule 8-10 buses per hour on Saturdays

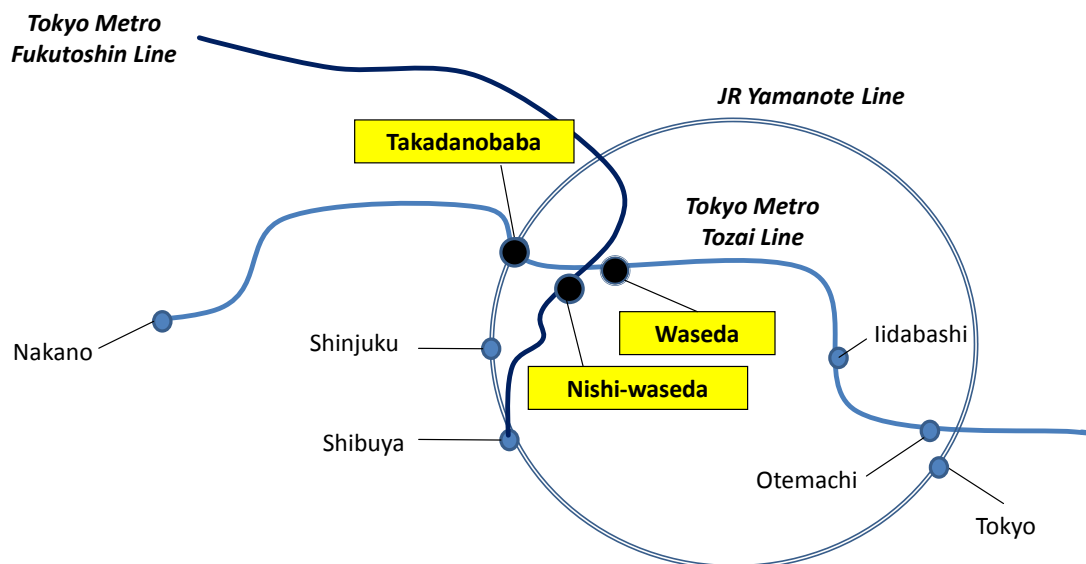
Refer to the websites below for further information about transportation to the venue.

Tokyo Metro Subway Map: www.tokyometro.jp/en/subwaymap/

East Japan Railway Company (JR East) Maps & Guides: <https://www.jreast.co.jp/e/downloads/index.html>

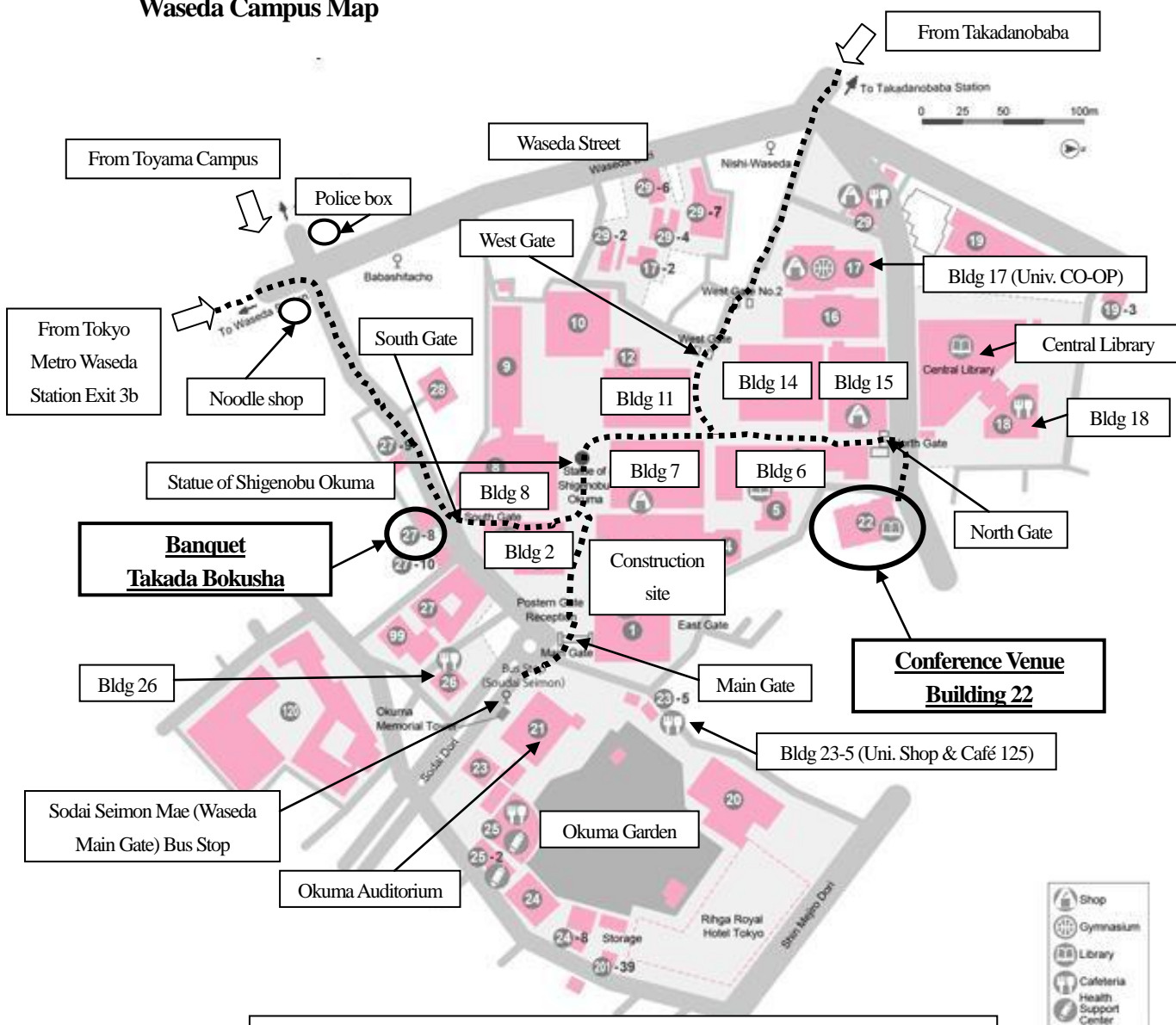
Toei Streetcar (Toden) Arakawa Line: www.kotsu.metro.tokyo.jp/eng/services/streetcar.html

Toei Bus website: tobus.jp/blsys/navi?LCD=e



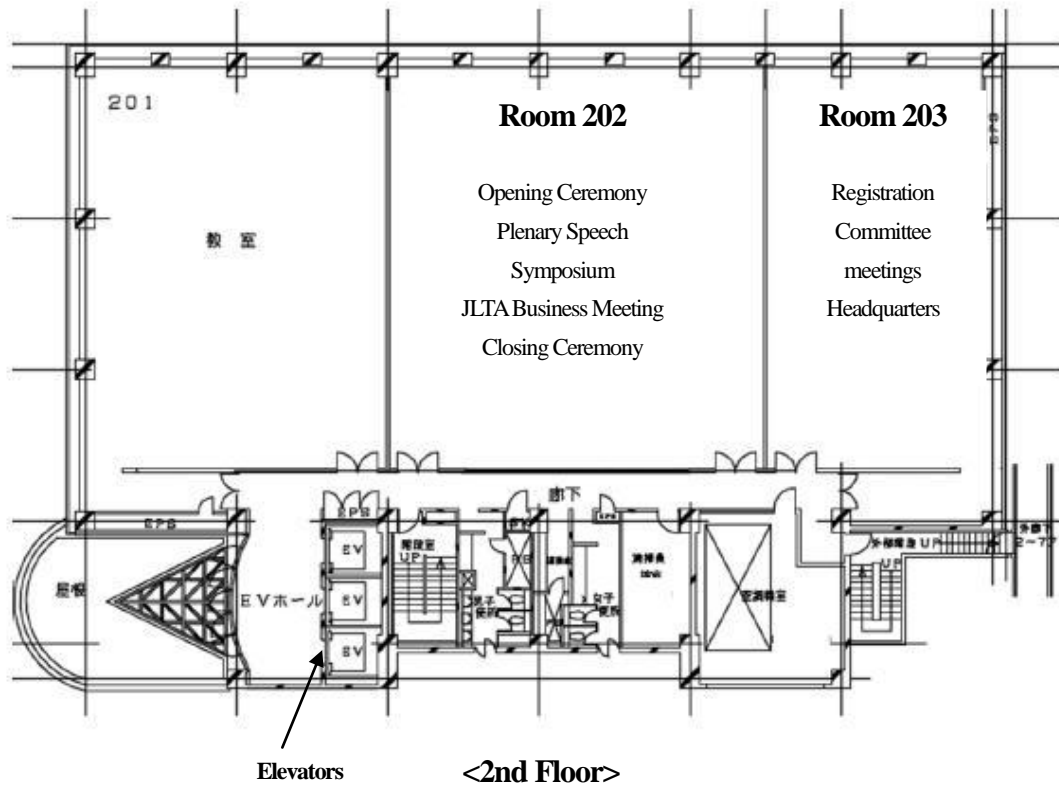
- 東京メトロ早稲田駅（東西線）3b 出口より徒歩 8 分。
 - 東京メトロ西早稲田駅（副都心線）1 番出口より徒歩 20 分。
 - 都電荒川線早稲田駅より徒歩 5 分。
 - 高田馬場駅（JR 山手線、西武新宿線）2 番出口より徒歩 23 分。
 - 高田馬場駅（JR 山手線、西武新宿線）2 番出口より都バスで 13 分（「(学02) 早大正門行き」又は「(早 81 出入) 早大正門行き」）。
- ※ バスの運賃は通常 200 円、学バス（学02）は 170 円です。
- ※ 高田馬場駅からのバス本数は、1 時間当たり 8~10 本です。
- ※ 高田馬場駅のバス乗り場の地図やバス時刻に関しては下記の東京都交通局ウェブサイトをご覧ください。 www.kotsu.metro.tokyo.jp/bus/

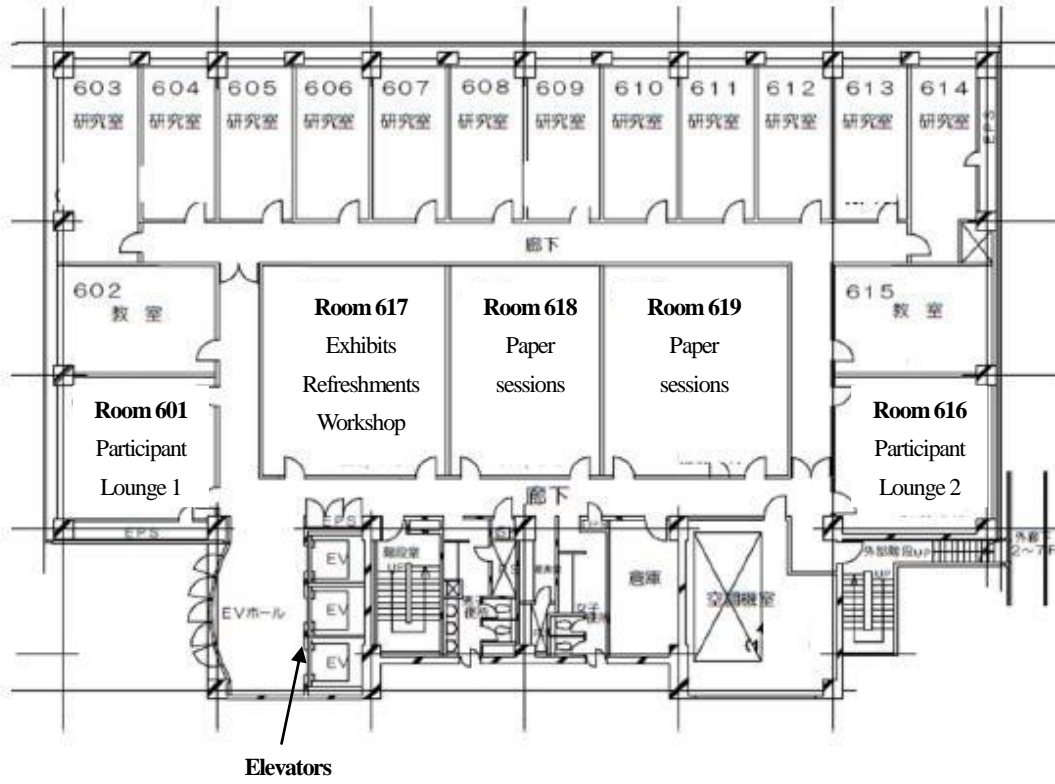
Waseda Campus Map



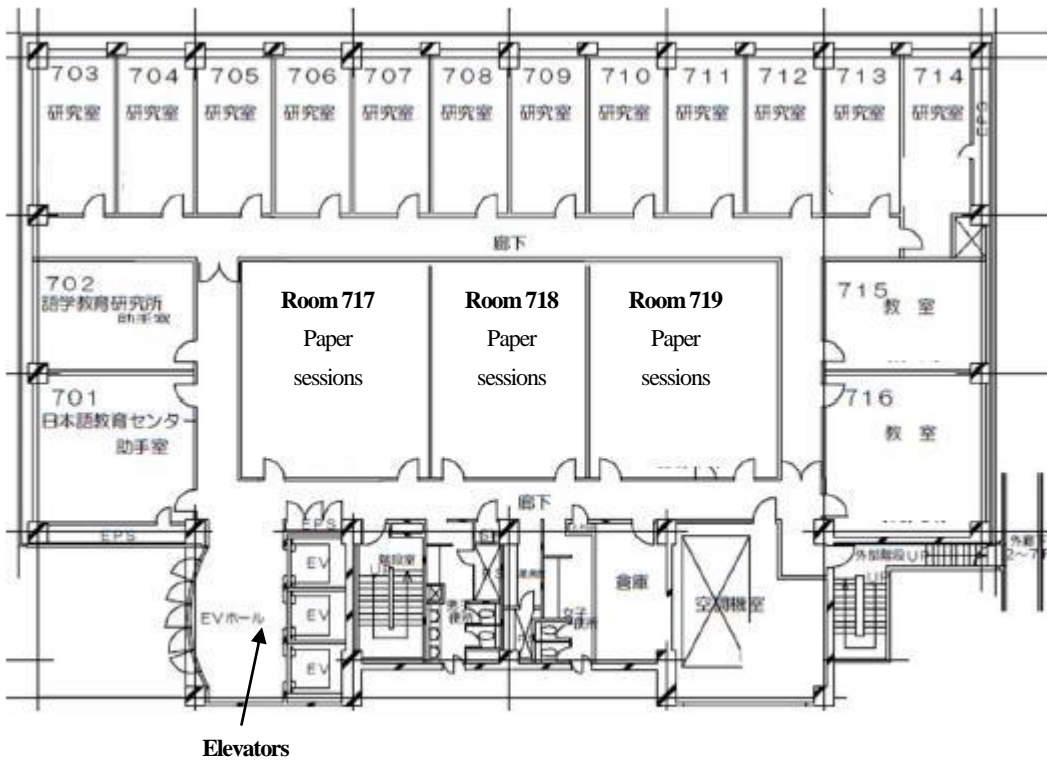
NOTE: The entrance to Building 22 is facing the street, across the Central Library (There is no access to Building 22 from inside Waseda Campus).

Map of Conference Venue (Building 22)





<6th Floor>



<7th Floor>

Special Conference Sponsors and Advertisement/Commercial Exhibit Sponsors

(特別協賛企業、広告・展示協賛企業)

Special Conference Sponsors (特別協賛企業)

第17回日本語テスト学会全国研究大会の実現にあたりましては、下記の特別協賛企業より格別の経済的支援を賜りました。誠にありがとうございました。

We are indebted to the special conference sponsors below for their general financial support, which has made the 17th Annual Conference of the Japan Language Testing Association possible.

株式会社アルク

ALC Press Inc.

<http://www.alc.co.jp/press/>

教育測定研究所

The Japan Institute for Educational Measurement, Inc.

<http://www.jiem.co.jp/>

国際教育交換協議会

Council on International Educational Exchange

<http://www.cieej.or.jp/ciee/>

公益財団法人日本英語検定協会

EIKEN Foundation of Japan

<http://www.eiken.or.jp/>

(50音順)

The New ETS Criterion.



国内 70 以上の教育機関で利用されている
ライティング指導ツールが 2013 年 8 月、リニューアルしました

<p>インターフェース改訂</p> <p>インターフェースが大幅に改訂され、明るく、より使いやすくなります。 Revise 画面は左右配置になり、より見やすくなります。 ついに iPad に対応します。</p>	<p>提出エッセイ全保存</p> <p>これまでは 1 つの課題につき 2 つ (最初と最後) の提出しか保存されませんでしたが、新バージョンでは提出したものがすべて保存されるようになります。</p>	<p>ピアグループ機能</p> <p>2 人以上の Student グループを作成できるようになります。グループ内でお互いのエッセイを見るのももちろん、ダイアログ機能も追加されます。</p>
<p>課題への URL 添付機能</p> <p>課題 (Assignment) に URL を添付できるようになります。画像や音声のファイル、ニュースのページなどをリンクさせることで、これまでできなかった課題を作成可能になります。</p>	<p>新・コメント機能</p> <p>これまでの Pop-up Note 機能では箇々の単語にしかコメントをつけることできませんでした。新しくなったコメント機能ではコメントをつける範囲を自由に選択できるようになります。</p>	<p>その他</p> <ul style="list-style-type: none"> • Student 画面のモニター機能 • Thesaurus 機能の追加

So, What is Criterion?

Criterion は、TOEFL® テストなどで知られる米国 ETS が開発したライティング指導のためのオンラインツールです。最大の特長は、エッセイ評価エンジン「e-rater®」により提出後わずか数秒で出るスコアとフィードバックです。WEB アプリケーションのため、余計なコストや管理の手間もかからず、ブラウザ上で全てのことを行えます。「ライティング指導を効率的に行える」ツールとして、多くの教育機関で利用されています。

CIEE ブース
at JLTA 2013

Criterion や TOEFL テストに関する最新情報を提供しております。是非お立ち寄りください。

お問い合わせ Criterion ご案内ページ www.cieej.or.jp/toefl/criterion

英語コミュニケーション能力判定テスト



Computerized Assessment System for English Communication



累計 118 万人が受けた

CASECで英語力チェック!

CASEC (キャセック)とは…

公益財団法人日本英語検定協会が基礎試験を行い、株式会社 教育測定研究所が開発・運営を行っている、日本で初めて「(項目別者理編)に基づいたCATIコンピュータ適応型テスト」として開発された、英語能力判定テストです。回答の正解・不正解に合わせて次の段階の難易度を変化させていくことにより、個人の英語能力値を従来よりも短いテスト時間で正確かつ客観的に測定することができます。また、インターネット上で時間と場所を選ばずに受験ができ、スコアもその場ですぐに判ります。

CASECのオフィシャルホームページ

<http://casec.evidus.com/>

CASEC

検索

納得の3つのポイント

POINT 1

40分ですぐ判定!

- ・平均試験時間は40分~50分
- ・終了後すぐに結果を表示

POINT 2

オンラインでいつでもどこでも!

- ・試験会場に行かなくてもインターネット環境さえあれば24時間165日いつでもオンラインで受験可能

POINT 3

TOEIC®スコア目安がわかる!

- ・TOEIC®のスコア目安・英検級の目安が表示され、しかも高い相関を誇ります

テスト概要

- ・計 60 問 / 1000 点満点
- ・セクションは 4 つ
- ・受験は平均 40 分 ~ 50 分
- ・受験後即時に結果を確度

SEC1 語彙の知識



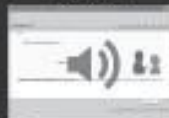
16 問 4 肢択一
(配点: 250 点満点)

SEC2 表現の知識



16 問 4 肢択一
(配点: 250 点満点)

SEC3 リスニングでの 大意把握



17 問 4 肢択一
(配点: 250 点満点)

SEC4 具体情報の 書き取り能力



11 問 キーボードでの入力
(配点: 250 点満点)



テストのスコアだけでなく、今後の学習をモチベートする情報が満載!
CASECであなただけの英語学習の強化ポイントが分かり、次の目標設定がしやすい!

オフィシャルスコアレポート



より詳しく習熟度が把握できるフィードバックシートで弱点を明確に!

TOEIC®スコア目安、全検定級の中での自己位置をマークアップ、難易度に応じた学習アドバイス、語彙力、読解力、聴解力、それぞれに対して記載されているので学習の手がかりが得られます。

NEW CASEC CAP-00 (新機能)



取得したスコアで、英語を用いてできることを具体的に記述!

スコアの意味する英語力をより明確にイメージできるようになりました。ビジネス旅行、受験シーンでの目標が定まってきました。

導入実績 (50校以上)

【企業】アセビコー株式会社/アプライマテリアル株式会社/株式会社イーオン/株式会社インテリジェンス/宇都宮県立国際情報専門学校/大塚製薬株式会社/大塚製薬株式会社/株式会社大林組/精工株式会社/株式会社ASAHI/レスコスシステム株式会社/住友スリーエム株式会社/ソノタ株式会社/地産地消推進委員会/経済企画庁/日本政策投資銀行/バリエル株式会社/ファイザー株式会社/富士ゼロックス株式会社/富士通株式会社/ブッポウ工業株式会社/三井物産株式会社/サンエー株式会社/フコク生命株式会社/株式会社リコー

【大学】東北大学/福岡大学/早稲田大学/福井県立大学/徳島大学/富山大学/立正大学/愛知大学/宇都宮大学/宇都宮女子大学/宇都宮短期大学/宇都宮女子短期大学/宇都宮女子短期大学/宇都宮女子短期大学/宇都宮女子短期大学/宇都宮女子短期大学/宇都宮女子短期大学

お問い合わせ先: (株)教育測定研究所 CASEC サポートセンター

TEL: 03-5775-9071 (平日 10:00 ~ 17:00) E-mail: casec-qa@evidus.com

JTEM 株式会社教育測定研究所

The Japan Institute for Educational Measurement, Inc.



英検

後援：文部科学省

おもてなしの心も英語で伝えたい。

2013年度試験日程

■ 第2回

一次試験 **10月13日** 日 二次試験 **11月10日** 日

申込受付 ▶ 8/1～9/19 ㊦ 協会必着 9/17 ㊦ 書店締切

■ 第3回

一次試験 ^{2014年}**1月26日** 日 二次試験 ^{2014年}**2月23日** 日

申込受付 ▶ 12/2～12/24 ㊦ 協会必着 12/20 ㊦ 書店締切

英検 公益財団法人
日本英語検定協会

詳細は英検ウェブサイトをご覧ください

英検

検索

公正と信頼—— プロメトリックの CBT配信サービス

The **Right** Expertise for the
Right Computerized Test

私たちプロメトリックは、実施する全ての試験が、受験者一人一人の人生と、
将来のキャリア形成に対して影響を与えるものであると考えています。
私たちは誇りを持って、この重要な使命に取り組んでいます。

試験作成ソリューション

私たちは、試験作成のプロセス（集約分析、試験問題作成、試験仕様設計、フォー
ム・プール構築、合否基準設定および項目分析など）において、経験・知識を
長年積み上げてきました。その専門性をベースに、受験者のスキルや知識、能
力を測定する妥当性・信頼性の高い試験プログラムの作成をサポートいたします。

試験配信ソリューション

私たちは、すべての都道府県に合計200カ所以上の試験会場を設け、また世界で
は8,000カ所以上の試験会場を通じて、毎年1,000万件の試験を配信しています。ま
た、試験認定団体それぞれのニーズに合わせ、CBTによる最適な試験配信ソリュー
ションをご提供いたします。

私たちは、全ての受験者が、
公正な試験を適切な環境で
受験し、正確かつ迅速に結
果を確認できるようサポー
トいたします。

Trusted Provider of Market Leading Test Development Solutions
www.prometric.jp



Commercial Exhibits (展示協賛企業 50 音順)

株式会社アルク

ALC Press Inc.

<http://www.alc.co.jp/press/>

NECラーニング株式会社

NEC Learning, Ltd.

<http://www.neclearning.jp/>

オックスフォード大学出版局

Oxford University Press

<http://www.oupjapan.co.jp/>

教育測定研究所

The Japan Institute for Educational Measurement, Inc.

<http://www.jiem.co.jp/>

国際教育交換協議会

Council on International Educational Exchange

<http://www.ciej.or.jp/ciee/>

Commercial exhibits are located in Room 617. Please ask staff at the registration desk for details.
展示は 617 教室で実施いたしております。詳しくは受付でお尋ねください。

Our special gratitude goes to Waseda University for making Waseda Campus available as the venue for the 17th Annual Conference of the Japan Language Testing Association.
第 17 回日本語テスト学会全国研究大会の実現にあたりましては、早稲田大学より、会場の提供を始めご支援・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 September 2014, at Ritsumeikan University. The conference schedule will be announced via the JLTA website as soon as the details become available.
We look forward to seeing you there.
2014 年度の日本語テスト学会全国研究大会は、2014 (平成 26) 年 9 月に立命館大学で行われます。
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 17 回 (2013 年度) 全国研究大会発表要綱
Handbook of the 17th Annual Conference of the Japan Language Testing Association

発行日 : 2013 年 8 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
