1. Conference Schedule Overview

Day 1: September 11, 2019 (Wednesday)

9:00－10:50	President and Secretary General Meeting (1204, 2nd floor, 1st building, Niigata Seiryo University)

11:00－15:00 (10:30 － Registration)	Workshop (Conducted in Japanese) “Bayesian Statistics and its Application to Foreign Language Education Study” Kunihiro KUSANAGI (Hiroshima University) Niigata Seiryo University (4108PCL, 4th building)

15:30－18:00	Board Meeting (1204, 2nd floor, 1st building, Niigata Seiryo University)

Day 2: September 12, 2019 (Thursday)

8:20－	Registration	(Hallway, 2nd floor, 1st building)
9:00－9:15	Opening Ceremony	(Seiryo Hall, 1st floor, 1st building)
9:25－9:55	Presentation I	(1301, 1303, 1304)
10:00－10:30	Presentation II	(1301, 1302, 1304)
10:30－10:45	Break	(1206, 2nd floor, 1st building)
10:45－12:00	Keynote Speech	(Seiryo Hall, 1st floor, 1st building)
12:00－13:40	Lunch Break (JLTA Committee Meetings)	(Cafeteria, 1st floor, 3rd building) (1204, 2nd floor, 1st building)
13:40－14:10	Presentation III	(1301, 1302, 1303, 1304, 1305)
14:15－14:45	Presentation IV	(1301, 1302, 1303, 1304, 1305)
14:50－15:20	Presentation V (Institutional Member Presentations)	(1301, 1302, 1303, 1304)
15:20－15:40	Break	(1206, 2nd floor, 1st building)
15:40－17:10	Symposium	(Seiryo Hall, 1st floor, 1st building)
17:20－17:40	Closing Ceremony & JLTA Best Paper Award Ceremony	(Seiryo Hall, 1st floor, 1st building)
17:40－18:00	JLTA General Business Meeting	(Seiryo Hall, 1st floor, 1st building)
18:20－20:20	Banquet	(Cafeteria, 1st floor, 3rd building)

Commercial Exhibits:	1203, 2nd floor, 1st building
Lunch Room for Participants:	Cafeteria, 1st floor, 3rd building
Break Room (after 10:30):	1206, 2nd floor, 1st building
Family Waiting Room:	1205, 2nd floor, 1st building
Headquarters:	1204, 2nd floor, 1st building

✧ Complimentary refreshments are available in the 2nd floor hallway in the 1st building.

✧ Please see p. 11 for lunch information.

Program of the 22nd JLTA Conference

September 12, 2019 (Thursday)

8:20－ **Registration** (Hallway, 2nd floor, 1st building)
Conference Attendance Fee: JLTA Members & JALT TEVAL SIG Members: ¥1,000
Non-members: ¥3,000; Graduate students: ¥1,000
Undergraduate students (with a proper student ID): Free

8:30－ **Registration for Commercial Exhibits** (Hallway, 1st floor, 1st building)

9:00－9:15 **Opening Ceremony** (Seiryo Hall, 1st floor, 1st building)
Coordinator: Tomoko FUJITA (St. Andrew's University)
Greetings: Yoshinori WATANABE (JLTA President; Sophia University)
Yoshiko LEIBOWITZ (President, Niigata Seiryo University)

9:25－10:30 **Presentations I and II** (Presentation: 20 minutes; Discussion: 10 minutes)
(1301, 1302, 1303, 1304, 3rd floor, 1st building)

10:30－10:45 **Break** (1206, 2nd floor, 1st building)

10:45－12:00 **Keynote Speech** (Seiryo Hall, 1st floor, 1st building)
Coordinator: Yoshinori WATANABE (Sophia University)
Title: An Ethics-based Approach to the Evaluation of Language Assessments
Lecturer: Antony John KUNNAN (University of Macau)

12:00－13:40 **Lunch Break**
Lunch Room for Participants: Cafeteria, 1st floor, 3rd building
JLTA Committee Meetings: 1204, 2nd floor, 1st building

13:40－15:20 **Presentations III, IV, and Institutional Member Presentations (V)**
(Presentation: 20 minutes; Discussion: 10 minutes)
(1301, 1302, 1303, 1304, 1305, 3rd floor, 1st building)

15:20－15:40 **Break** (1206, 2nd floor, 1st building)

15:40－17:10 **Symposium** (Seiryo Hall, 1st floor, 1st building)
Theme: Evaluating Fairness and Justice of University Entrance English Examinations in Japan
Coordinator Hidetoshi SAITO (Ibaraki University)
Panelist 1 Hidetoshi SAITO (Ibaraki University)
An Introduction to the Symposium and External Testing Agencies Views on Effects on Teaching

Panelist 2 Kiwamu KASAHARA (Hokkaido University of Education)
External English Language Testing Agencies' Views on their Own Tests: Its Use and Washback on Teaching and Learning: Part 2

Panelist 3 Yasuyo SAWAKI (Waseda University)

Linking Commercial English Language Assessments to the CEFR and Using them for Admission Decision-Making: Challenges and Future Directions

Discussant Antony John KUNNAN (University of Macau)

17:20－17:40 **Closing Ceremony & JLTA Best Paper Award Ceremony** (Seiryo Hall, 1st floor, 1st building)

Coordinator: Youichi NAKAMURA (Seisen Jogakuin College)

2018 Best Paper Award Recipient: Akiyo HIRAI (University of Tsukuba)

17:40－18:00 **JLTA General Business Meeting** (Seiryo Hall, 1st floor, 1st building)

Selection of the chair

Reporter: Rie KOIZUMI (JLTA Secretary General; Juntendo University)
Kazuhiko KATAGIRI (JLTA Vice Secretary General; Senshu University)
Yuichiro YOKOUCHI (JLTA Vice Secretary General; Hirosaki University)
Makoto FUKAZAWA (JLTA Vice Secretary General; University of the Ryukyus)

18:20－20:20 **Banquet** (Cafeteria, 1st floor, 3rd building)

Coordinator: Tetsuo KIMURA (Niigata Seiryo University)

Presentation Details

Seiryo Hall, 1st floor, 1st building

Keynote speech chair　Yoshinori WATANABE (Sophia University)
Keynote speech summary　Hideaki OKA (University of Tsukuba)
Symposium summary　Hiroki MAEDA (University of Tsukuba)

Part	Presenter (Affiliation)	Title (Page)
I 9:25－9:55	------	------
II 10:00－10:30	------	------
10:45－12:00	**Keynote speech** Lecturer: Antony John KUNNAN (University of Macau)	An Ethics-based Approach to the Evaluation of Language Assessments (p. 14)
III 13:40－14:10	------	------
IV 14:15－14:45	------	------
V 14:50－15:20	------	------
15:40－17:10	**Symposium**: Evaluating Fairness and Justice of University Entrance English Examinations in Japan	
	Coordinator: Hidetoshi SAITO (Ibaraki University)	Introduction (p. 16)
	Panelist 1: Hidetoshi SAITO (Ibaraki University)	An Introduction to the Symposium and External Testing Agencies Views on Effects on Teaching (p. 16)
	Panelist 2: Kiwamu KASAHARA (Hokkaido University of Education)	External English Language Testing Agencies' Views on their Own Tests: Its Use and Washback on Teaching and Learning: Part 2 (p. 18)
	Panelist 3: Yasuyo SAWAKI (Waseda University)	Linking Commercial English Language Assessments to the CEFR and Using them for Admission Decision-Making: Challenges and Future Directions (p. 19)
	Discussant: Antony John KUNNAN (University of Macau)	

1301, 3rd floor, 1st building

Part	Presenter (Affiliation)	Title (Page)
I 9:25 － 9:55	白鳥 金吾 (北星学園大学短期大学部)	大学入学者選抜における民間試験の活用についてーケンブリッジ英検の検証を通した考察 (p. 20)
II 10:00 － 10:30	島田 勝正 (桃山学院大学)	結束性判断テストによるディスコース能力の発達の測定 (p. 21)
III 13:40 － 14:10	鈴木 健太郎 (共栄大学)	英単語の語源学習における学習者の注意―形態素の既知性，意味的透明性，テスト形式を焦点に― (p. 22)
IV 14:15 － 14:45	大木 俊英 (白鷗大学)	音声認識力を測る新しいタスクの可能性：Word Count Task の妥当性と信頼性について (p. 23)
V 14:50 － 15:20	平野 琢也 (株式会社アルク　アルク教育総合研究所) 木下 あおい（株式会社アルク　アルク教育総合研究所）	スピーキングテスト TSST を使用した「高校生の英語スピーキング力と学習実態に関する 3 年間追跡調査」及び「アルク語学教育研究支援制度」について (p. 36)

1302, 3rd floor, 1st building

Part	Presenter (Affiliation)	Title (Page)
I 9:25 － 9:55	------	------
II 10:00 － 10:30	Yujia ZHOU (Tokyo University of Foreign Studies) Jamie DUNLEA (British Council) Masashi NEGISHI (Tokyo University of Foreign Studies), Barry O'SULLIVAN (British Council) Asako YOSHITOMI (Tokyo University of Foreign Studies)	Gathering a Posteriori Validity Evidence of a Computer-based Speaking Test for Japanese University Admission (p. 24)
III 13:40 － 14:10	Gordon ALLAN (British Council)	Assessing Speaking: Current Practice and Perceptions of Japanese High School English Teachers (p. 25)
IV 14:15 － 14:45	島田 めぐみ (日本大学) 孫 媛 (国立情報学研究所) 保坂 敏子 (日本大学) 澁川 晶 (国際基督教大学) 谷部 弘子 (東京学芸大学)	日本語聴解認知診断テスト開発のための予備調査 (p. 26)
V 14:50 － 15:20	堀川 有美 (国際交流基金日本語試験センター) 板橋 貴子 (国際交流基金日本語試験センター) 田渕 咲子 (国際交流基金日本語試験センター)	日本語能力試験聴解「即時応答」のプロトコル調査の報告 (p. 37)

1303, 3rd floor, 1st building

Part	Presenter (Affiliation)	Title (Page)
I 9:25 – 9:55	Richard ARBER (Kwansei Gakuin University)	Filling in the Blanks: Rethinking EAP Assessment for CLIL Contexts† (p. 27)
II 10:00 – 10:30	------	------
III 13:40 – 14:10	Nick SAVILLE (IELTS Cambridge Assessment English)	Keeping IELTS Fit for Purpose: The Future of Learning and Assessment (p. 28)
IV 14:15 – 14:45	Myles GROGAN (Kwansei Gakuin University)	Grading the Grades: An Investigation into Classroom-based Compulsory University EFL Assessment (p. 29)
V 14:50 – 15:20	Nick SAVILLE (IELTS Cambridge Assessment English)	The Impact of IELTS in Japanese Higher Education (p. 38)

†Assessment practice presentation

1304, 3rd floor, 1st building

Part	Presenter (Affiliation)	Title (Page)
I 9:25 – 9:55	Nan-Young JI (Korea Polytechnic University) (*KELTA delegate presentation)	Investigation into Validity of Paraphrasing Task as a Writing Performance Test Item for EFL Learners (p. 30)
II 10:00 – 10:30	Yunjeong CHOI (Korea University) (*KELTA delegate presentation)	The Effectiveness of Peer Feedback and Writing Conference as Learning-Oriented Writing Assessment (p. 31)
III 13:40 – 14:10	Ryoko FUJITA (Juntendo University)	The Effects of Listening Instructions Using Materials with Background Noise on EFL Learners' Listening Abilities (p. 32)
IV 14:15 – 14:45	Keita NAKAMURA (Eiken Foundation of Japan)	A Validation Study of New Business Speaking Test (p. 33)
V 14:50 – 15:20	青山 智恵 (ケンブリッジ大学英語検定機構)	新 4 技能 CBT Linguaskill リンガスキルの活用報告 (p. 39)

1305, 3rd floor, 1st building

Part	Presenter (Affiliation)	Title (Page)
I 9:25 – 9:55	------	------
II 10:00 – 10:30	------	------
III 13:40 – 14:10	Paul WICKING (Meijo University)	How Japanese Students Conceptualize and Experience University Assessment (p. 34)
IV 14:15 – 14:45	David ALLEN (Ochanomizu University)	The Impact of Japanese-L1 Loanword Knowledge in Tests of English-L2 Lexical Knowledge (p. 35)
V 14:50 – 15:20	------	------

2. From the JLTA Office: Information for Conference Participants

Registration

1. There is no need to register in advance.
2. The conference registration site is Hallway on the 2nd floor of the 1st building.
3. The conference attendance fee is ¥1,000 for members (including institutional members) and ¥3,000 for non-members (¥1,000 for non-member graduate students and ¥0 for non-member undergraduate students).
4. If non-members apply for membership at the registration desk, the conference attendance fee will be ¥1,000. The JLTA annual membership 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.
5. Please wear your conference name card strap throughout the conference.
6. The banquet fee is ¥3,000. The banquet registration is conducted at the registration desk. There is no pre-conference registration. The banquet will be held at the Cafeteria in the Niigata Seiryo University. (See the map on p. 44).
7. The conference handbook is available at the registration desk on the day of the conference and is not sent by post in advance.

Family Waiting Room

1. A family waiting room is available for family members (junior high school age and above) who are not attending the JLTA events but are accompanying an adult(s) attending the events.
2. Desks and chairs are available.
3. As no JLTA or care staff is present in the room, its use is limited to people from junior high school age and above and at their own risk.
4. Members of a participant's family who do not attend presentations or lectures and only use the family waiting room are exempt from the conference attendance fee. Please ask for a "participant's family" tag at the registration desk when your family member registers for the JLTA events.
5. The family waiting room is at 1205 on the 2nd floor in the 1st building. Complimentary refreshments are available in the 1st floor hallway in the same building. Feel free to enjoy them.

Lunch and Participants' Lounge, Etc.

1. Please use the Cafeteria on the 1st floor in the 3rd building for lunch.
2. Complimentary refreshments are available in the 2nd floor hallway.
3. There are a convenience store (FamilyMart) and several restaurants within a ten-minute walk. The on-campus Cafeteria does not operate, but University CO-OP in the 5th building is available for participants to buy food and beverages.

Accommodation

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

Smoking

Smoking is prohibited on campus.

***Emergency Contact E-Mail Address*:** u16yoko@gmail.com (Yuichiro YOKOUCHI, JLTA Vice Secretary General)

To Presenters

1. Presenters will have 20 minutes to present their paper, followed by 10 minutes for discussion.
2. There will be no chair person in the presentation room. A time keeper will show you the time left.
3. Please register at the registration desk first. Please go to the designated room 5 minutes prior to the starting time of the presentation.
4. Presenters are expected to bring a PC. There will be an audio terminal connector (for PC connection through a stereo mini plug) and a D-sub 15-pin cable in the presentation room. If necessary, please prepare an HDMI to VGA adaptor. Mac users should bring their own Mini DisplayPort to VGA Adapter. Third-party adapters do not work properly sometimes.
5. Eduroam or other Wi-Fi Internet access is not available.
6. Please bring handouts in case your PC or the projector does not work.
7. If you need a letter of invitation, contact Rie KOIZUMI (JLTA Secretary General) at rie-koizumi@mwa.biglobe.ne.jp

3. Abstracts (発表要旨)

Keynote Speech (Seiryo Hall, 1st floor, 1st building) 10:45－12:00

An Ethics-based Approach to the Evaluation of Language Assessments

Antony John KUNNAN (University of Macau)
akunnan@umac.mo, akunnan@gmail.com

The dominant 20th century approach to the evaluation of language assessments was the Standards-based approach. The *Standards* most evaluators referred to were the APA, AERA, NCME *Standards* (1999; 2014) or the derivative ILTA, ALTE or EALTA *Standards*. These *Standards* considered the center-piece of their evaluations to be evidence from studies of validation, reliability and consequences. In the early part of the 21st century, the Argument-based approach (proposed by Bachman, 2005; Bachman and Palmer, 2014) emerged as a new approach that used the Toulmin way of structuring arguments (with claims, warrants, backing and rebuttal). Their emphasis of this approach was to include consequences and to clarify evaluation procedures using Toulmin's framework. The reviewers who used the *Standards*-based and Argument-based approaches published their evaluations in the *Mental Measurements Yearbook*, *Language Testing* and *Language Assessment Quarterly*. These published evaluations have many deficiencies in critical ways: they were mainly descriptive (not evaluative), they were insider evaluations (not independent as the authors are often the testing agencies' collaborators), they did not have test performance data (for secondary analyses), they accepted the test agencies' claims rather than evaluate the assessment against principles (for example, of fairness, justice, etc.), and they were lacking in any intellectual foundation (as test agencies did not explicitly state their ethical beliefs).

To remedy this situation, I am proposing an ethics-based approach to assessment evaluation. In this approach, a principled basis for fairness of assessments and justice in institutions is used as a framework that in turn is used to develop the *Principle of Fairness* and *Principle of Justice*. Procedurally, Toulmin's structuring of arguments is used: Principles, claims, warrants, backing, qualifier, and rebuttals or counter-claims. I will examine three claims from Principle of Fairness (Opportunity to Learn, Meaningfulness, and Absence of bias) and one claim from Principle of Justice (Consequences). I will provide evidence of support for the various claims and also offer rebuttals of claims. The claims examined are (1) **opportunity-to-learn** in the classroom of two automated essay evaluation software (Vantage Learning's *MY Access!* and Pearson's *WriteToLearn*),

(2) **meaningfulness** in terms of consistency and dependability and the internal structure of a placement test (UCLA's *New ESL Placement Examination*), (3) **absence of bias** in terms of differential item functioning based on age (in the Cambridge English Language Assessments' *Certificate in Advanced English*), and **consequences** (of the *U.S. Naturalization Test*). Analyses used were correlations, exploratory and confirmatory factor analyses. These studies showed that some test agencies' claims could not be supported and rebuttals or counter-claims could be entertained. In other cases, more independent research studies are needed to find evidence to support or rebut claims. I will conclude with some remarks regarding rights and responsibilities of test takers and test users.

Bio

Antony John KUNNAN is a language assessment specialist. His research interests are fairness of tests and testing practice, assessment literacy, research methods and statistics, ethics and standards, and language assessment policy. After completing his Ph.D. from UCLA in 1991, he was awarded a post-doctoral fellowship at the University of Michigan, Ann Arbor for a year. From 1992 until 2013, he was assistant, associate and full professor at California State University, Los Angeles. In 2006, he received a Fulbright scholarship to Tunghai University, Taiwan where he was a visiting professor and scholar. He also was professor (and now Honorary Professor) at the University of Hong Kong and a professor at Nanyang Technological University, Singapore. From 2016, he has been Professor of English and Associate Dean of the Faculty of Arts and Humanities at the University of Macau.

He has served in many capacities at the international level: as secretary-treasurer and president of the International Language Testing Association. He was the founding president of the Asian Association for Language Assessment, and the founding editor of *Language Assessment Quarterly* (2003-2013). He was a member of the TOEFL Committee of Examiners and the New TOEFL (now iBT) at Educational Testing Service, Princeton, and a research consultant at the University of Cambridge English Language Assessment where he conducted research workshops and projects.

His latest publications include: edited volumes *The Companion to Language Assessment* (in 4 volumes, Wiley, 2014), *Language Testing and Assessment* (in 4 volumes, Routledge, 2014) and *Talking about Language Assessment* (Routledge, 2015) and an authored book *Evaluating Language Assessments* (Routledge, 2018).

Note: The first and second paragraphs of this bio statement were adapted from the University of Macau's website: https://fah.umac.mo/staff/staff-do/antony-kunnan/

Symposium (Seiryo Hall, 1st floor, 1st building)　15:40－17:10

Evaluating Fairness and Justice of University Entrance English Examinations in Japan
(日本の大学入試英語科目を公平性と公正性の観点から評価する試み)

Coordinator	**Hidetoshi SAITO (Ibaraki University)**
Panelists	**Hidetoshi SAITO (Ibaraki University)**
	Kiwamu KASAHARA (Hokkaido University of Education)
	Yasuyo SAWAKI (Waseda University)
Discussant	**Antony John Kunnan (University of Macau)**

Introduction and Symposium Paper 1: An Introduction to the Symposium and External Testing Agencies Views on Effects on Teaching

Coordinator and panelist: Hidetoshi SAITO (Ibaraki University)
hidetoshi.saito.cldwtr@vc.ibaraki.ac.jp

This symposium is a three-part report of a survey of six testing agencies whose English language tests have just been endorsed for use as part of the revised Japanese university entrance exam system starting in 2021. This will be followed by Dr. Antony Kunnan's comments on the issues we will raise.

According to this reform plan, both commercial proficiency tests and the new Common Test for English language will be used until 2023, and then the education ministry plans to withdraw the new Common Test and replace it entirely with external proficiency tests. This change will most likely generate unexpected consequences, positive or negative, which require serious consideration.

We (Saito, Kasahara, & Sawaki) have decided to 'preliminarily evaluate' the English language tests developed and administered by the six agencies using Kunnan's principles of fairness and justice (2018). Using his principles along with Toulumin's argument approach, one can evaluate the feasibility of test and argue for and against its use for intended purposes and consequences.

The six agencies responded to the questions concerning issues that the test-takers and their teachers might wonder about: the tests' potential use and consequences. The questions were targeted at revealing communicability of the test agencies' current positions and practice about the future test use. Our 'evaluation' necessarily contains speculative arguments, because the new exam system has yet to be implemented. Nevertheless, we are confident of the significance of our preliminary evaluation because of the magnitude of impact the reforms will have on more than 500,000 high school students and their parents alike.

The first question in the survey was "What do the agencies think about minimizing the potential risk their test may pose of narrowing the curriculum and teaching to the test?" I will discuss their responses (claims) based on Kunnan's principle of justice: subprinciple 1, which states that "an assessment institution ought to foster beneficial consequences to the test-taking community" (2018, p. 80).

Bio

Hidetoshi SAITO (PhD) currently teaches pre- and in-service English teachers and graduate students at Ibaraki University. His papers have appeared in journals including *Language Assessment Quarterly, Language Testing,* and *Language Teaching Research.* He has just completed a chapter for a book on assessment for learning in CLIL. His research interest includes formative assessment, CLIL, and discussion instruction. Saito currently

serves as a Vice President of the Japan Language Testing Association.

Symposium Paper 2: External English Language Testing Agencies' Views on their Own Tests: Its Use and Washback on Teaching and Learning: Part 2

Kiwamu KASAHARA (Hokkaido University of Education)
kasahara.kiwamu@a.hokkyodai.ac.jp

Based on Kunnan's principles of fairness and justice (2018), this presentation reports on a preliminary evaluation of the English language tests developed and administered by the six agencies, and their use as part of the nationwide university entrance examination system. In particular, of the five questions we posed in the survey administered to the six agencies, this presentation reports on responses to Questions 2 through 4.

Question 2 asked: "Research has shown that when such tests are used for entrance examination purposes, students could mainly focus on test preparation, consequently narrowing the content they learn. Do you have any advice for examinees regarding this point?" Question 3 asked: "Additionally, what initiative do you think should be taken to have beneficial effects on high school teachers and examinees? Please describe the current plan and direction for the future plan." Question 4 asked: "How would you respond to examinees when they point out that your test(s) include(s) vocabulary that exceeds the 5,000 words that high school graduates should know as specified in the New Course of Study beginning from elementary school starting in 2020?" Questions 2 and 3 concerned Kunnan's Principle of Justice: subprinciple 1, which states that "An assessment institution ought to foster beneficial consequences to the test-taking community" (p. 80). Question 4 concerned Kunnan's Principle of Fairness: subprinciple 1, which states that "an assessment institution ought to provide adequate opportunity to acquire the knowledge, abilities or skills to be assessed for all test takers" (p. 80).

After analyzing responses to Questions 2 through 4, the presentation will indicate the extent to which each test could serve as part of the nationwide university entrance examination system, along with areas of concern or interest that have emerged in the process of analyzing the responses.

Bio

Kiwamu KASAHARA currently teaches pre- and in-service English teachers and graduate students at Hokkaido University of Education. He has a testing and assessment course with his recently published book on testing. His research interest includes vocabulary acquisition, classroom SLA, and teaching testing for pre-service teachers. He serves as the editorial chief of *Annual Review of English Language Education in Japan.* He is also an editorial advisor of a nationally approved junior high textbook series, *New Horizon*.

Symposium Paper 3: Linking Commercial English Language Assessments to the CEFR and Using them for Admission Decision-Making: Challenges and Future Directions

Yasuyo SAWAKI (Waseda University)
ysawaki@y.waseda.jp

This presentation focuses on the final question included in the survey conducted with the six testing agencies participating in the score reporting system for the new university entrance examination administered by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT): "Test takers' scores on your test will be converted to the levels of the CEFR (Common European Framework of Reference for Languages; Council of Europe, 2001). How would you respond if a test taker asks how accurate the score boundaries between the different CEFR levels are for your test?" Currently the information relevant to this question is provided by MEXT in the form of a concordance table specifying the score mapping results for these agencies' assessments to the CEFR levels, which has generated controversies over the feasibility and fairness of the score-to-CEFR level conversion based on this information. This question mainly concerns two aspects of Kunnan's (2018) Principle of Fairness (p. 80). One is Sub-principle 2, the degree to which test score interpretation is meaningful for and consistent across all test takers. A fundamental issue is the variation across the external English language assessments in assessment purpose, target population, construct representation, and design. This makes it difficult to map these tests to a common scale such as the CEFR, which is context-free in nature (Deygerts, van Gorp, Demeester, 2018; Papageorgiou, Tannenbaum, Bridgeman, Cho, 2015). The other relevant principle is Sub-principle 4, the degree to which the standard-setting procedures employed to link those assessments to the CEFR are appropriate for equitable decision-making. The standard-setting methods that those agencies employ determine the quality of the information presented in MEXT's score concordance table. Moreover, guidance that these agencies provide to stakeholders for score interpretation would also be critical in minimizing test score misuse.

In this session, the presenter will first summarize survey results on this question and issues of consideration that emerge from them. This is followed by a proposal of potential future directions for building and supporting a fairness argument (1) by facilitating stakeholders' understanding of similarities and differences among the assessments and how MEXT's concordance table should be interpreted and used in specifying admission requirements; (2) by enhancing the standard-setting practice on which the table is based; and (3) by planning and conducting empirical validation studies of this new score reporting system by the collaboration of MEXT, the testing agencies, and universities.

Bio

Yasuyo SAWAKI (Professor, Faculty of Education and Integrated Arts and Sciences, Waseda University) currently teaches various undergraduate- and graduate-level courses in language assessment and academic English at Waseda University. Her current research interests include diagnostic English language assessment and feedback as well as integrated writing assessment. Sawaki is a board member of the Japan Language Testing Association, Secretary/Treasurer of the Asian Association for Language Assessment, and member of the editorial advisory boards of the *Language Testing* and *Language Assessment Quarterly* journals.

Paper Session

1301, 3rd floor, 1st building　Part I (9:25-9:55)

大学入学者選抜における民間試験の活用についてーケンブリッジ英検の検証を通した考察

白鳥 金吾 (北星学園大学短期大学部)

2018 年、大学入試センターは、TOEFL、ケンブリッジ英語検定といった試験を「大学入試英語成績提供システム参加要件を満たしていることが確認された資格・検定試験」（以下、認定試験）として公表した。しかし、これらの認定試験は、言うまでもなく、日本の大学入試を目的に開発されたものではないことから、それぞれの試験が高等学校の教育内容を反映しているか、大学教育を受けるに相応しい英語力を測定する試験として妥当であるかを検証する必要がある。本発表では、認定試験の一つであるケンブリッジ英語検定 B1 Preliminary が高校の教育内容と整合しているか、さらに、CEFR B1 が入学者選抜の基準となり得るか否かについて考察する。はじめに、B1 Preliminary と平成 21 年改訂高等学校学習指導要領外国語編それぞれの目標、トピック、タスク、文法・語彙、音声等の項目を照合した。その結果、本テストは日常生活で求められるコミュニケーション能力を評価する出題内容や形式となっており、多くの項目において、学習指導要領との対応関係が確認された。次に、B1 Preliminary の結果をもとに、受検者の総合的な英語力及び技能別の状況を分析した。国内の短期大学の学生を対象に実施した本テストの結果から、CEFR B1 は高校卒業時の達成可能な目標であり、受検者の英語力を識別する指標となり得ることが示唆された。こうした結果を踏まえ、良質な英語 4 技能テストが有する優れた機能に着目し、「英語学習者」の視点を大切にした検証を行い、民間英語試験を大学入試として活用する際の検討事項について議論を行う。

1301, 3rd floor, 1st building　Part II (10:00-10:30)

結束性判断テストによるディスコース能力の発達の測定

島田 勝正 (桃山学院大学)

英語の結束性のあるディスコースにおいては、情報は文末焦点の原則 (Maxim of End-focus) に基づいて、旧情報から新情報へと提示される (Leech, 1983)。

この研究の目的は、この文末焦点の原則を発見させるディスコーストレーニングの処遇効果として、コミュニケーション能力の下位要素で、文脈に応じて適切な言語形式を選択できるディスコース(結束性)能力がいかに向上するかを測定することである。そのために、結束性判断テストを開発し、処遇の事前、直後、および、処遇の 1 週間後と 3 回実施した。

結束性判断テストは複文、与格交替、態の 3 つの文法範疇を取り扱っている。テスト問題は複文の主節の位置が 12 項目（前半・後半、それぞれ 6 項目）、与格交替が 12 項目（SVO・SVOO、それぞれ 6 項目）、態が 12 項目（能動態・受動態、それぞれ 6 項目）で、合計 36 項目から構成されている。

テスト形式は、問題文を読んでそれに続くより適切な発話を A, B のいずれかから選ぶ二者択一方式である。問題文は Globalvoice English (Ver. 2) で録音し、参加者が問題文を聞いて 3 秒以内に解答するように、パワーポイントを使って時間制限を課した。採点は、文末焦点の原則にしたがったより適切な選択肢を選んだ場合に 1 点を与えた。

結果として、処遇効果として複文において、事後、遅延事後テストの得点に上昇がみられた。特記すべきは、転移効果として与格交替と態にも事後、遅延事後テストにおいて得点の上昇に有意差が見られた。さらに、特記すべきことは、複文、与格交替、態の文法範疇間の相関が事後、遅延事後とより強くなっている点である。なお、信頼性係数（クロンバックα）は、事前テストが 0.643、事後テストが 0.764、遅延事後テストが 0.854 と上昇している。これは学習により文末焦点の原則が 3 つの文法範疇間で一次元化したことを示唆する。

1301, 3rd floor, 1st building　Part III (13:40-14:10)

英単語の語源学習における学習者の注意―形態素の既知性，意味的透明性，テスト形式を焦点に―

鈴木 健太郎 (共栄大学)

本研究の目的は，単語に含まれる形態素を利用した語源学習の効果を，学習者要因として形態素の既知性 (形態素の意味を知っているか)，語彙要因として形態素の意味的透明性 (形態素の意味が語の意味にどの程度関連しているか) の 2 つの観点から明らかにすることである。その際，複数の語彙テスト実施することで，学習者がどのような情報に注意を向けながら学習しているかを検証した。

実験では，120 名の日本人大学生が，意味的透明性の異なる接頭辞と語根からなる 16 の派生語の目標語 (例. 透明性高: conspire「団結する」= con-「共に」+ -spire「息づく」; 透明性低: supersede「後任となる」= super-「超えて」+ -sede「座る」) を，複数の学習条件に分かれて 16 分間で学習した。語源学習を行う 2 条件では，テスト時の形態素とその訳の提示の事前告知の有無 (形態素既知，形態素未知) により形態素の既知性を実験的に操作した。また，一般的な学習と比較するために，形態素を提示しない条件を用意した (通常学習)。学習成果を測定するために，学習の直後と 1 週間後 (遅延) にそれぞれ語形のみを提示するテスト (手がかり無) と形態素とその意味を提示したも (手がかり有) 2 種類の語彙テストを実施した。

語彙テスト得点を条件間で比較した結果，手がかり無テストでは，直後遅延ともに形態素未知条件の方が既知条件よりも成績が良く，未知条件では形態素だけでなく語形と意味のつながりにもより注意を向けていたことが示された。透明性の効果に関しては，手がかり有テスト (遅延) においてのみ，透明性の高い語に対する得点が透明性の低いものよりも高かった。しかしながら，ほとんどのテストにおいて通常学習条件の成績は語源学習の 2 条件を上回っていた。発表では，これらの結果に加え，学習に要した時間やアンケート結果などを踏まえて，語源学習の効率・効果について総合的に議論する。

1301, 3rd floor, 1st building　Part IV (14:15-14:45)

音声認識力を測る新しいタスクの可能性：Word Count Task の妥当性と信頼性について

大木 俊英 (白鷗大学)

入力された音声を瞬時に処理しなければいけないリスニングでは、素早く正確な音声認識が不可欠だが、外国語学習者はその能力が十分でないという指摘がある（e.g., Field, 2008; Tsui & Fullilove, 1998）。学習者にこの能力向上のための学習を促す方法の 1 つは、ハイステークスなテストにおいて音声認識力を測る問題を出題することである。しかし、センターリスニングをはじめとした大規模テストでは、文章を聞いてその内容と合致する絵や文を選ばせることで内容理解度を測る問題が多く、基本的な音声認識力の測定をねらいとした問題は出題されていない。

連続発話の認識度を測るタスクとしてはディクテーションが候補として挙がるが、書き取りに時間がかかるため試験時間が長くなる、採点に時間がかかる、綴りの知識が成績に影響する、といった欠点がある。しかし、Shi and Farooq（2012）などの研究で用いられている音声復唱タスクに比べれば、スピーキングを苦手とする日本人学習者には適したタスクだと言えよう。このような背景のもと、筆者はディクテーションに代わる簡便なタスクとして「Word Count Task（以下 WCT）」を考案した。これは、文を聞いて空所に入る単語の数を答えるだけの簡易なタスクだが、文章の意味を必ずしも理解できなくても答えられるという欠点があり妥当性が怪しまれる。

そこで本研究では、私立大学の 1〜2 年生 123 名に TOEIC の英文を使ってディクテーションと WCT を行わせ、これらの得点を用いて WCT の妥当性と信頼性の検証を試みた。まず、2 つの得点の相関が有意だったため（$r = .46$）、併存的妥当性の観点から、2 つのタスクは同じような能力を測っていることが示唆された。しかし信頼性分析の結果、ディクテーションのα係数（.70）に比べて、WCT のα係数（.45）が低いことがわかり、項目を増やすなどの改善が必要であることが明らかとなった。現在、項目応答理論（2PLM）に基づいた分析を行っており、学会当日はその結果も含めて発表する予定である。

1302, 3rd floor, 1st building　Part II (10:00-10:30)

Gathering a Posteriori Validity Evidence of a Computer-based Speaking Test for Japanese University Admission

Yujia ZHOU (Tokyo University of Foreign Studies)
Jamie DUNLEA (British Council)
Masashi NEGISHI (Tokyo University of Foreign Studies)
Barry O'SULLIVAN (British Council)
Asako YOSHITOMI (Tokyo University of Foreign Studies)

To foster positive washback on English education, the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) has announced a new policy of encouraging universities to use four-skill English tests for entrance purposes (MEXT, 2017). In response to this call, Tokyo University of Foreign Studies (TUFS), in collaboration with the British Council, developed the British Council TUFS-Speaking Test for Japanese Universities (BCT-S). The joint development project, drawing on the British Council's expertise in developing and delivering the online speaking component of Aptis, offers one potential solution to the challenges faced by Japanese universities in introducing an independent speaking component for university-specific exams.

According to the socio-cognitive model for validating speaking tests (O'Sullivan, 2011; O'Sullivan & Weir, 2011), test developers should attend to three basic components during the test development and validation cycle: The test taker, the test system and the scoring system. A priori validity evidence on these three aspects has been reported in Zhou et al., 2018. To inform further test development, the present study focused on providing a posteriori validity evidence related to the operational test.

TUFS students who took the operational test in February 2019 completed a questionnaire designed to elicit their perceptions of BCT-S regarding test validity, testing procedure, and test prompts, and some of them participated in follow-up interviews as well. Their responses to the questionnaire were compared to those from TUFS students and high school students elicited in the a priori validation studies. A multiple-rater study was also carried out using the operational test responses from the February 2019 administration.

Regarding perceptions of BCT-S, test takers who took the operational test showed a similar pattern to university and high school students in the a priori validation studies: They were in general satisfied with the test validity and testing procedure. While it seems that test prompts used in the operational test caused less confusion among test takers compared to those used in the pretests, interview data indicated that some areas of the test needed further improvement. These results, along with those from Many Facet Rasch Measurement (MFRM) analysis are reported in detail in the presentation, and the implications of the findings for building validity argument of BCT-S and future test development are discussed.

1302, 3rd floor, 1st building　Part III (13:40-14:10)

Assessing Speaking: Current Practice and Perceptions of Japanese High School English Teachers

Gordon ALLAN (British Council)

With a view to the introduction of speaking into university entrance exams in 2020, this study sought to investigate the construct of speaking currently taught and tested in Japanese high schools, and compare it with high school English teachers' perceptions of the construct of speaking tested in a computer-mediated standardised international test.

The introduction of performance testing to university entrance exams (UEEs) in Japan is at least partly aimed at encouraging the teaching of speaking and writing skills (MEXT, 2014). Teachers provide the link between the content of new tests and the proficiency outcomes sought by policy makers, so their perceptions of what is being tested may influence these outcomes. This study focused on speaking because, given the history of English teaching in Japan, it is likely to be the least familiar element of new UEEs for many teachers.

Data were collected in 2017 from high school teachers participating in the Leaders of English Education Project (LEEP), a national training project created and delivered by the British Council for MEXT. As part of LEEP, all had taken the Aptis test, including a computer-mediated speaking component. Quantitative data were collected from 40 teachers via an online survey, giving an indication of current practice and perceptions. These were supplemented by qualitative data from seven interviewees to explore the issues and underlying thinking in greater depth.

The reported practical challenges of testing students' speaking included test security, and the time and resources required to test a large number of students. This caused teachers to either avoid the need for security by using presentations or group conversations as tests, or to use technology to test large numbers of students simultaneously. The most frequently used criteria tended to be non-linguistic. In contrast teachers perceived Aptis to use more directly language-related criteria. (Since speaking will almost certainly be computer-mediated in new UEEs, Aptis was used as an example of the kind of test teachers might have to prepare students for).

These teachers perceived a difference between their current practice and the demands of the Aptis test. This paper explores the reasons for their choices about speaking assessment, and their perceptions of what they need in order to adequately prepare students for the forthcoming UEE English speaking tests. The aim is to enhance understanding of the educational context into which these new tests will be introduced, and identify support that may help to facilitate positive washback.

1302, 3rd floor, 1st building　Part IV (14:15-14:45)

日本語聴解認知診断テスト開発のための予備調査

島田 めぐみ（日本大学）
孫 媛（国立情報学研究所）
保坂 敏子（日本大学）
澁川 晶（国際基督教大学）
谷部 弘子（東京学芸大学）

本発表では，日本語聴解認知診断テストを開発するために実施した予備調査の結果を報告する。認知診断テスト（Cognitive Diagnostic Assessment）とは，測定対象となっている領域・分野の学習や理解に必要なアトリビュート（attribute）の習得状況を推定するテストであり，各学習者の強みや弱みを，学習支援のための情報として提供することができる。日本語聴解認知診断テストの測定対象は，CEFR（Common European Framework of Reference for Languages）の B1 レベルを目指す日本語学習者であり，B1 レベル到達のためにどのような学習が必要かという情報を提供することを目的としている。そのため，CEFR の枠組みを参考に，12 のアトリビュートを設定した。テストは，4 枝選択形式の合計 27 項目からなる。

2018 年 6 月から 9 月にかけて，日本国内外 5 都市でテストを実施し，298 人からデータを得た。認知診断モデルによる分析を行い，各学習者の 12 のアトリビュートの習得確率を計算した。12 のアトリビュート習得確率の平均とテスト得点の相関は 0.846 であり，アトリビュート習得確率の平均が高い者はテスト得点も高いという結果であった。しかし，同じ得点を得た受験者であっても，各アトリビュートの習得確率は異なっており，このことは認知診断テストが一人ひとりの学習者に適したフィードバックを行うことができることを示唆している。

各項目の分析の結果，必要なアトリビュートを保持しているにも関わらず誤答してしまう確率（slip）と偶然にその項目に正答してしまう確率（guessing）の値が高い項目が複数観察された。このことにより，該当項目に対するアトリビュートの特定に誤りがある，12 のアトリビュート以外に関連するものが存在するという可能性が考えられ，アトリビュートの再検討が課題と言える。

1303, 3rd floor, 1st building　Part I (9:25-9:55)

Filling in the Blanks: Rethinking EAP Assessment for CLIL Contexts

Richard ARBER (Kwansei Gakuin University)

†Assessment practice presentation

While numerous forms of testing remain highly prevalent through language proficiency exams such as IELTS and TOEFL, it is widely acknowledged that assessment in the EAP/ESP field is neglected both in the production of meaningful testing frameworks and in the 'post-proficiency exam' stage (Schmitt & Hamp-Lyons, 2015). Organisations such as BALEAP and EALTA have tentatively produced guidelines for test development and assessment but these remain largely tethered to the CEFR, placing little emphasis on the research process. By extension, the role of the EAP teacher remains highly multidisciplinary and fluid in its nature, often requiring practitioners to develop a knowledge of various research genres, deliver curriculums driven to accommodate students entering a wide range of academic fields, perform examination roles with overly generalized criteria and supervise projects that require some expertise in research design. In addition to this, few EAP programmes acknowledge the importance of interdiscursive and interdisciplinary practices in academic skills/production. As student numbers on pre-sessional courses in the UK, the US and Australia continue to grow, there is further exigency to address these gaps. This study looks at a pre-masters programme based in Japan that seeks to train and assess student's academic language and introduce the research design process suitable for postgraduate courses, allowing students to gain entry to a select number of UK universities. The methodology for this study is threefold. Firstly, characteristics of the assessment criteria across both EAP and research methods are analysed for content and compared with other cases from UK universities. Secondly, samples of assessed coursework from a cohort of 43 students will be presented in view of common assessment challenges. These samples will survey student success in areas of genre specific lexis, grammar, discourse features and 'move' structuring, as well as how they produced a social science-style research project typically taught during postgraduate study in the UK. Thirdly, by interviewing teacher-examiners on their experiences of using such assessment criteria, the study will look at how academic staff have negotiated flaws and insufficiencies. The results of this presentation largely support the challenges and conclusions present in the existing literature; however, this study goes further in suggesting the research process, student-driven corpora and genre acquisition should play a much greater role in EAP/ESP assessment practices.

1303, 3rd floor, 1st building　Part III (13:40-14:10)

Keeping IELTS Fit for Purpose: The Future of Learning and Assessment

Nick SAVILLE (IELTS Cambridge Assessment English)

In this session, I will give an overview of how IELTS, the International English Language Testing System is kept relevant to the needs of stakeholders, promotes learning and incorporates new technology.

The owners of IELTS – British Council, IDP: IELTS Australia and Cambridge Assessment English – have to ensure that the test conforms to the highest international standards of language assessment. In 2018 over 3.5 million IELTS tests were taken in over 140 countries. The test is available up to 4 times a month and up to 48 times a year. Over 10,000 organisations recognize IELTS worldwide, and professional bodies, immigration authorities and other government agencies. With such huge numbers comes great responsibility because IELTS is helping millions of people each year realise their language ambitions for study, work or immigration, and recognising organisations and authorities rely on IELTS as a genuine proof of English language ability.

In an increasingly complex world, the future of assessment presents both challenges and opportunities. Cambridge Assessment English conducts ongoing research to ensure that IELTS continues to be fit for purpose and have a positive impact on test takers and users. Technological advances can help this happen. For example, a recent study by Berry et al. (2017) compared IELTS speaking exams conducted face to face and delivered by laptop, and found no significant difference in scores. Such findings have implications for equality of opportunity in that they may facilitate the delivery of exams to remote areas. Another current study is using video and data collecting software to understand rater perceptions of speaking ability and decision making. As we build greater capability with reliable automated assessment, the swifter and more flexibly we can respond to stakeholders' evolving assessment needs.

The affordances offered by technology as it becomes more sophisticated, supported by high quality research, will enable us to develop a fully integrated learning and testing system to meet the evolving assessment needs of test takers and enhance even further the positive impact of IELTS on education and society.

1303, 3rd floor, 1st building　Part IV (14:15-14:45)

Grading the Grades: An Investigation into Classroom-based Compulsory University EFL Assessment

Myles GROGAN (Kwansei Gakuin University)

Throughout the last two decades, there has been an increasing recognition within language assessment literature of the unique nature of classroom-based assessment (CBA). The research presented in this session examines the assessment of compulsory language classes in a particular Japanese university setting, exploring the similarities and differences between CBA and larger scale assessment through the actions of those directly involved in it. It also examines the impact this may have on institutions and policy more broadly. Special emphasis is placed on how grade construction is influenced not only by academic content, but by a variety of other factors both within and external to the classroom setting.

Set in a large university in Western Japan, the project began with semi-structured interviews of contract teachers. Themes identified were then followed up with interviews with participants within the broader university structure to see how teacher perceptions of grading converge and diverge from those providing the mandate for the course. A third layer of interviews was conducted with general administrators to gain further perspectives of the roles of language grades and grading against the backdrop of the university context. Through a process of coding, memoing, and constant comparison, a broad picture of the different components that influence the grading process in this setting was constructed.

Those involved showed a passion for education and language education, not just in terms of linguistic ability, but also for the personal development of the students. However, the scale of the process in such a large setting revealed a number of surprising elements. In particular, while providing academic freedom for teachers to use their own strengths, meeting expectations (both academic and pastoral) created conflict for all concerned. Another aspect of this was the isolation of those making policy, and the difficulties of achieving consensus with regard to student needs, pedagogy, and the ultimate role of the grade itself.

Organizational issues, both perceived and actual, were found to be impacting the grades. The presentation closes with suggestions of a more encompassing approach to assessment literacy, in which institution and teacher are treated more equitably. In addition, suggestions for approaches to context-based research are made, which may assist those involved in assessment to look at what grades are, how they are actually constructed, and the interpretations that may be drawn from them.

1304, 3rd floor, 1st building　Part I (9:25-9:55)

Investigation into Validity of Paraphrasing Task as a Writing Performance Test Item for EFL Learners

Nan-Young JI (Korea Polytechnic University)
***KELTA delegate presentation**

Paraphrasing quality has proven to be highly associated with two major linguistic components: lexical competence and syntactic competence (McCarthy, Guess, & McNamara, 2009). Nonetheless, few attempts have been made to view the skill as a window through which learners' inter-language can be estimated. How learners manipulate sentence structures and vocabulary in retelling may represent their current level of language proficiency. Therefore, with an aim to verify whether paraphrasing tasks are legitimate as a writing test item to accurately identify the learners' productive language abilities, a correlation study was conducted with 364 test-takers ranging from grade 7 to university freshmen. The scores the learners earned from paraphrasing tasks were compared with those obtained from self-assessments of their English abilities in the case of the secondary school students and from TOEIC in the case of the university students. Paraphrase rating scales adopted in this study were developed by school teachers, considering the range of Korean secondary students' English proficiency. It has been revealed that paraphrasing task has the potential as a valid writing test item as proven by statistically significant correlation coefficients between two sets of scores.

1304, 3rd floor, 1st building　Part II (10:00-10:30)

The Effectiveness of Peer Feedback and Writing Conference as Learning-Oriented Writing Assessment

Yunjeong CHOI (Korea University)
***KELTA delegate presentation**

In second/foreign academic writing, researchers have found that peer-assessment plays a significant role in enhancing learners' writing ability. Yet, little research addressed how learners understand and negotiate feedback from peer-assessment in their writing process, and more importantly, how such process helps their learning in writing. To fill the gap, the present study examined how peer written feedback and writing conferences influence learners' writing practices and their satisfaction within the framework of learning-oriented assessment. Participants were 26 Korean undergraduate students enrolled in an academic English writing course. They were asked to provide written feedback on their peers' work in a computer-mediated environment for six weeks with three different writing assignments involved. Each time the peer-assessment was followed by having a writing conference between the writer and the reviewer where they were supposed to further interact and discuss the given feedback. Students wrote a reflection each time and participated in an end-of-semester survey to examine the effectiveness of and their satisfaction on peer-assessment and writing conference. The findings and their pedagogical implications are discussed in the context of learning-oriented writing assessment.

1304, 3rd floor, 1st building　Part III (13:40-14:10)

The Effects of Listening Instructions Using Materials with Background Noise on EFL Learners' Listening Abilities

Ryoko FUJITA (Juntendo University)

Background noise is a component that test developers should consider when they evaluate the cognitive validity of language tests. According to Field (2013), language tests become less valid if a qualitative difference between the processes assessed by the tests and conditions that test takers actually encounter in real-life contexts is found. For example, in real-life settings, listening comprehension is associated with background noise.

Language learners have difficulty with listening comprehension, especially in noisy conditions, due to their limited linguistic knowledge (Field, 2008). Past studies suggest that foreign accents negatively affect listening comprehension in the presence of noise, and the listening comprehension of low-proficiency learners is more affected by background noise than that of high-proficiency learners (Hodoshima, Masuda, Yasu, & Arai, 2009; Rogers, Dalby, & Nishi, 2004). Although past studies have examined the effects of noise from various points of view, such as proficiency levels or foreign accents, little is known about listening instructions using listening materials that have noise in the background. Therefore, the aim of the current study is to investigate how listening instructions using materials with background noise affect EFL learners' listening abilities and their perceptions about background noise.

In the study, 49 participants were assigned to either an experimental group ($n = 22$) or a control group ($n = 27$). In the experimental group, the instructor used listening materials with background noise, while she used the same listening materials without background noise in the control group. The experiment was conducted over a 10-week period, and the listening instructions were comprised of partial dictation practice. As pre- and post-listening tests, the participants answered listening comprehension questions with and without background noise. After each listening session, the participants wrote down their reflections about their listening comprehension. A three-way analysis of covariance (ANOVA) was conducted to evaluate the listening test results. The participants' reflections about their listening comprehension were analyzed qualitatively.

The test results did not show improvements in pre- and post-listening tests, and no significant interaction was observed between two groups. However, both groups' comprehension of the listening questions without noise was statistically significantly higher than their comprehension of the listening questions with background noise. Most of the students in the experimental group commented that the background noise bothered them. However, a few students commented that they were getting used to the background noise. Further implications of these results will be discussed in the presentation.

1304, 3rd floor, 1st building　Part IV (14:15-14:45)

A Validation Study of New Business Speaking Test

Keita NAKAMURA (Eiken Foundation of Japan)

Test validation has become an important part of test development because it is becoming increasingly important for test developers to conduct validation studies to ensure the proper use of tests and the interpretation of the results for a particular group of stakeholders (Chapelle et al., 2006).

This study presents the result of a series of studies to collect validity evidence of the newly developed speaking test of English for business purposes. The author starts from result of 1) needs analysis and task development, 2) trial study to check testing-time and task difficulty, 3) scale development study based on reference group, 4) concurrent validation study of the new test with other tests, and finally 5) the limitation and the implication of the study.

Based on the needs analysis, the new test was designed to have three parts, part 1, 2, and 3. In part 1, test takers were asked to provide their job-related basic information such as what they do and likes/challenges of their current job. In part 2, test takers were asked to read both texts and graphs and summarize the issue and give a possible solution. In part 3, test takers were asked to read text and graphs to give their opinions to the given topic. Through those parts, interviewer was asked to rate test takers' both English Language Skills (ELS) and Business Performance Skills (BPS).

A total of 39, 398, and 626 adult learners of English took part in the study 2), 3) and 4), respectively. In study 2), each participant took the prototype task and filled in the questionnaire which asked them their responses to the testing time, task instruction, and task difficulty. In study 3), test reliability and item-level factor structure of the new test were investigated using Mplus 7.4 (Muthen & Muthen, 2015). In study 4), relationship with EIKEN and BULATS speaking were investigated in terms of correlation coefficient. In study 3), test takers from various fields of work (e.g. IT, education, or service industry) participated in this study. The test reliability was 0.99, while the correlation between the two rating criteria, Business Performance Skills (BPS) and English Language Skills (ELS) was 0.96.

In this presentation, the details of study results would be presented and, the limitation and the implication of the study would be discussed.

1305, 3rd floor, 1st building　Part III(13:40-14:10)

How Japanese Students Conceptualize and Experience University Assessment

Paul WICKING (Meijo University)

Since the turn of the century there has been increasing interest in the ways that educational assessment can encourage or discourage effective learning processes. Much of this interest has centered around formative assessment, and those practices which can promote student learning through the act of being assessed. In order to do that, formative assessment necessarily taps into the affective and cognitive dimensions of learning, touching upon students' feelings, motivations, beliefs, attitudes and conceptions surrounding learning and assessment. The purpose of the present study is to gather and analyze Japanese students' conceptions and experiences of assessment, in order to lay the foundation for assessment practices that can better promote learning.

This study is mixed methods research, integrating data from both quantitative and qualitative sources. Original data was gathered at Japanese universities via two instruments. The first was a multidimensional self-report survey, which was administered on a volunteer sample of 613 Japanese university students, of which 552 valid cases were drawn. The second data collection instrument was a narrative frame. The use of narrative frames for eliciting qualitative data is a method first developed by Barkhuisen and Wette (2008) to explore university English teachers' experiences in China. In essence, a narrative frame is a series of sentence starters, connectives and sentence modifiers which scaffold the writer and guide him/her to concentrate on certain features of his/her narrative story. The narrative frame data were drawn from eight intact EFL classes held in three different Japanese universities. In total, 219 students completed the narrative frame.

Analysis of the survey results was performed with IBM SPSS version 22. To begin with, a table of descriptive statistics was generated. Next, a factor analysis was conducted, which revealed a 7 factor solution to explain how these students conceived of assessment. The narrative frame data was put through a process of qualitative content analysis. A coding frame was first developed, following a strategy of subsumption, after which it was input into NVivo for Mac and then triangulated with the survey data.

The results indicate that Japanese students approach formative assessment tasks in a way that is at odds with popular Confucian categories. Students did not seem to be highly competitive, they valued practical skill over book knowledge, and familial obligation was not a strong factor in educational motivation. The presentation concludes with pedagogical implications for teachers seeking to conduct formative assessment with Japanese students.

1305, 3rd floor, 1st building　Part IV(14:15-14:45)

The Impact of Japanese-L1 Loanword Knowledge in Tests of English-L2 Lexical Knowledge

David ALLEN (Ochanomizu University)

From a psycholinguistic perspective, cognates are words that overlap in orthographic and/or phonological form and meaning across languages (e.g., heart and hart, in English and Dutch, respectively). Loanwords in Japanese that are borrowed from English (e.g., アドレス /adoresu/ 'address') also share some degree of phonological form and meaning, making them cognate under the psychological definition.

In second language (L2) studies, cognates have been shown to provide an advantage over noncognates in terms of response speed and accuracy, an observation referred to as the cognate facilitation effect. This cognate effect has been observed in many languages and in many tasks (see Dijkstra & van Heuven, 2018). Moreover, studies with participants whose L1 differs in script from their L2, such as Japanese speakers of English, have shown comparable findings (Allen & Conklin, 2013; Miwa et al., 2014), suggesting that the overlap in phonological form generates an observable processing advantage over noncognates.

In language assessment, there is a growing body of research demonstrating that cognates impact participants' accuracy on tests of lexical knowledge. These studies have investigated the cognate effect in same-script languages (e.g., Batista & Horst, 2016; Cobb, 2000; Meara, Lightbown & Halter, 1994) and in different-script languages (e.g., Elgort, 2013; Laufer & McLean, 2016). In this presentation, I will describe two separate studies (Allen, 2018, 2019) which demonstrate the cognate effect with Japanese learners of English in two different tests of English receptive vocabulary knowledge, the Vocabulary Levels Test (VLT; Schmitt, Schmitt & Clapham, 2001) and the Vocabulary Size Test (VST; Nation & Beglar, 2007).

In total, 183 L1-Japanese undergraduates took part in the studies. Whether English items had loanword equivalents in Japanese was determined using the Balanced Corpus of Contemporary Written Japanese (NIJLL, 2013). Moreover, rather than using a simple binary classification of cognate/noncognate based on whether a loanword exists in the corpus, L1 cognate frequency was used as a continuous predictor of accuracy. Using generalized linear mixed effects models, the findings revealed a significant effect of L1 cognate frequency on selection accuracy in both bilingual and monolingual versions of the VST, and for both target selection accuracy and distractor rejection accuracy in the VLT. Taken together, knowledge of loanwords appears to support the development of English lexical knowledge.

The implications of these findings for language assessment in the Japanese context are discussed, focusing not only on the validity of the above tests but also the validity issues that arise for standardized testing in specific contexts.

Parts V: 賛助会員発表 (Institutional Member Presentations) (14:50-15:20)

1301, 3rd floor, 1st building　Part V (14:50-15:20)

スピーキングテスト TSST を使用した「高校生の英語スピーキング力と学習実態に関する 3 年間追跡調査」及び「アルク語学教育研究支援制度」について

平野 琢也（株式会社アルク　アルク教育総合研究所）
木下 あおい（株式会社アルク　アルク教育総合研究所）

アルク教育総研では、語学教育や語学マーケットに関する調査・研究を実施しております。今回は「高校生のスピーキング力と英語学習実態に関する調査」について発表いたします。大学に入学してくる学生さんは、高校で、どのような学習を積み、どの程度のスピーキング力を身につけているのでしょうか。この疑問に答えるため、2015 年度から 3 年間に渡り、高校 3 校の協力を得て、アルクのスピーキングテスト TSST（Telephone Standard Speaking Test）を使用して高校生 275 人のスピーキング力を測定し、英語学習実態をアンケートにより調査いたしました。その結果、高校 1 年次から 3 年次にかけてスピーキング力が向上した生徒の多かった学校や、同一学校内でスピーキング力が向上した生徒は、「学習時間」や「学習内容」に一定の特徴があることが分かりました。高校生のスピーキング力の推移やその背景を調べた本調査が、大学の先生方や学会関係者の皆様の、お役に立てば幸いです。

また、2019 年に創立 50 周年を迎えたアルクでは、語学教育の研究に従事する方々を支援させていただくことで語学教育のさらなる発展に貢献したいと考え、「アルク語学教育研究支援制度」を創設いたしました。これは、研究に必要な資金や、研究にご利用いただくための教材・サービス等を提供する制度で、これから研究実績を積みたいと考えていらっしゃる研究者や大学院生の方々、現場での実践を研究発表としてまとめたいとお考えの中学・高校の先生方などにご活用いただきたいと考えております。その概要についてもお知らせいたします。

1302, 3rd floor, 1st building　Part V (14:50-15:20)

日本語能力試験聴解「即時応答」のプロトコル調査の報告

堀川 有美（国際交流基金日本語試験センター）
板橋 貴子（国際交流基金日本語試験センター）
田渕 咲子（国際交流基金日本語試験センター）

「日本語能力試験」は、1984 年より日本語を母語としない人の日本語能力を測定し認定する試験として実施されている。「聴解」試験に含まれる 3 枝選択式の「即時応答」は、2010 年の改定後に導入された問題形式の一つで、短い発話を聞き、それに対する適切な応答を即時に判断できるかどうかを問うことをねらいとしている。

妥当性検証の一環として、「即時応答」の解答行動を探るプロトコル調査を実施した。日本語学習経験のある非日本語母語話者 45 名を対象に、一人ずつに「即時応答」の問題 2 題に解答してもらい、どのように解答したかインタビューを行った。

本発表では、誤答選択枝の選択率が想定より高かった 1 題の分析結果を報告する。インタビューのデータを、英語の聴解タスクの解答ストラテジーを分析した Ross（1997）を参考に、解答ストラテジーの観点から分類し、調査対象者がどのように解答していたかを分析した。

分析の結果、正答を選んだ調査対象者の多くは、音声テキストの意味を正しく理解した上で正答を選んでいたことが確認できた。逆に、誤答を選んだ調査対象者は、音声テキストの意味を正しく理解できていなかったことが確認できた。つまり、「即時応答」の当該問題について、聴解問題としての妥当性をプロトコル調査の面から確認できたと考えられる。

誤答を選んだ解答行動の中には、問題をほとんど聞き取れないといった理解度の低いケースもあれば、一語の聞き誤りのために意味を誤解しているケースなどもあった。聴解能力に段階があり、「即時応答」の短い音声テキストの理解においてもその段階が反映されている可能性が示唆された。

1303, 3rd floor, 1st building　Part V (14:50-15:20)

The Impact of IELTS in Japanese Higher Education

Nick SAVILLE (IELTS Cambridge Assessment English)

This session looks at the impact of IELTS on learning in higher education in Japan. I will present findings from a study which investigates whether IELTS exerts a positive impact on productive language skills, study habits and motivation.

Traditional approaches in Japan have been criticised for placing too much emphasis on rote learning and not enough on skills development, with speaking skills being particularly neglected. Therefore, one of the report's most important washback hypotheses concerned productive skills, and whether using IELTS for higher education in Japan might foster better learning of speaking and writing, including greater spoken fluency and more effective interactive communication.

In the research design, about 200 undergraduate students were recruited to take IELTS as the measure of language proficiency, with the test administered on two occasions to investigate learning gains. A mixed-methods approach with survey and interviews was used to collect relevant contextual information, including test-takers' experiences and perceptions.

In summary, the report sheds light on the potential benefits of using IELTS – a four-skills test with an emphasis on communication skills – in a Japanese educational context. It appears that preparing for IELTS not only provides clear goals and motivation for Japanese learners of English, but also fosters good study habits without excessive cramming or test preparation activities (i.e. an absence of negative washback). The report suggests that there is indeed positive washback of the kind originally suggested by the developers of IELTS. It demonstrates that IELTS encourages Japanese students to develop their productive skills, and provides clear evidence that they do make measurable proficiency gains.

1304, 3rd floor, 1st building　Part V (14:50-15:20)

新 4 技能 CBT Linguaskill リンガスキルの活用報告

青山 智恵 (ケンブリッジ大学英語検定機構)

ケンブリッジ英語検定、IELTS に次ぐケンブリッジ大学英語検定機構が作問する試験で、4 技能を測ることができる「Linguaskill リンガスキル」の特長は、「速い」「安い」「正しい」こと。リスニング＆リーディングがコンピューター適応型テスト、ライティングは AI による自動採点、スピーキングは AI と採点官によるハイブリッド・スコアリングで採点されており、日本では、高校生を中心に活用が拡がっています。新しいテストですが、既に大学入試での活用も始まりました。

ケンブリッジが採用する Cambridge English スケールを介して、日本で展開中のケンブリッジが作問する試験（ケンブリッジ英語検定、IELTS）と比較検討できるため、目的に応じて多様な活用方法が考えられます。また、JLTA でも継続してご紹介させて頂いている Write & Improve や、昨年末発表された Speak & Improve は繰り返し練習できる Linguaskill リンガスキルのテスト準備にぴったりの無償のセルフ・プラクティス用ツールとなっています。

実践報告として、Linguaskill リンガスキルの試験センターである学校法人河合塾のデータを引用し、離島を含めた高校生の活用実績、実際の声をご紹介します。その他、実際活用頂き、正しい試験結果にご満足頂いている事例についても、ご紹介させて頂ければと考えています。

4. Workshop Information（ワークショップ情報）

題目：「ベイズ統計とその外国語教育研究への応用」（日本語で実施）

講　師　草薙　邦広（広島大学）
司　会　平井　明代（筑波大学）

日時：　2019 年 9 月 11 日（水） 11 時 00 分～12 時 30 分、13 時 30 分～15 時 00 分
（昼食をはさむ）
場所：　新潟青陵大学（4 号館、4108PCL 教室）
参加費：　1,000 円
定員：　30 名（申し込み順）
参加申し込み期間：7 月 1 日（月）～9 月 4 日（水）　［先着 30 名になり次第締め切り］
参加条件：ハンズオンセミナーでは以下の計算環境を前提とします。

（a）OS Windows 7 以上のバージョン、または MacOS X 以上のバージョン、または近年の Linux OS
（b）Microsoft Excel やそれに準じる表計算ソフトウェア
（c）R 3.0.0 以上のバージョン；以下からダウンロード可能
（https://cran.ism.ac.jp/bin/windows/base/）
（d）統計ソフトウェア JASP；以下からダウンロード可能（https://jasp-stats.org/）

目的：

1. ベイズ統計の基本を特に頻度主義との違いから理解する。
2. 従来の t 検定、分散分析、重回帰分析、相関分析といった基礎的な手法のベイズ統計的代替法を習得する。

手順：

1. 講義 1: ベイズ統計の基本
2. ハンズオンワークショップ 1: ベイズ統計を試してみる
3. ハンズオンワークショップ 2: ベイズ統計を使った外国語教育研究
4. 講義 2: より高度なモデリングへ

申し込み方法：

1. 定員に達するまで申し込み可能です。ワークショップを円滑に進めるため、下記 URL にアクセスし、以下の情報をご登録ください。
   https://goo.gl/forms/BMsDd0XorDye7Ip13
2. 上記の申込方法で不具合がある場合はメールでの受け付けも致します。下記の情報を藤田亮子（順天堂大学）r-fujita@juntendo.ac.jp まで e-mail でご連絡ください。

（1）氏名・所属・e メールアドレス
（2）ベイズ統計に興味があるのであれば、それはどうしてか教えてください。
（3）ベイズ統計を使用して、どのような研究実践をしたいか教えてください。
（4）講師へのご質問（希望者のみ）
（5）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）

【ご連絡】学会当日、学生食堂は営業しておりません。大学近隣の飲食店（3～4 店）でご食事いただくか、学内の 5 号館 1 階大学生協の売店（パンやおにぎり等を販売）や大学近隣のコンビニエンス・ストアでお買い求めいただき、6 号館の学生ホールでご飲食ください（ワークショップ会場での飲食禁止です）。

Bayesian Statistics and its Application to Foreign Language Education Study

(Conducted in Japanese)

Lecturer: Kunihiro KUSANAGI (Hiroshima University)
Chair: Akiyo HIRAI (University of Tsukuba)

Date: September 11, 2019 (Wednesday), 11:00－12:30, 13:30－15:00 (with a 60-minute lunch break)
Venue: Niigata Seiryo University (Building #4, 4108PCL)
Attendance Fee: ¥1,000
Maximum Number of Participants: 30 (first-come, first-served basis)
Registration Period: July 1st (Mon.) – September 4th (Wed.)
(The registration will be closed after the number of participants reaches the limit.)
Prerequisite: All participants must bring a computer with:

（a） a later version of Windows 7 or MacOS X, or recent version of Linux OS.
（b） Microsoft Excel or other spreadsheet software.
（c） R 3.0.0 or later; downloadable from (https://cran.ism.ac.jp/bin/windows/base/)
（d） statistical software JASP; downloadable from (https://jasp-stats.org/)

- Objectives

1. To understand the basics of Bayesian statistics by contrasting it with frequentism.
2. To learn Bayesian statistics as an alternative to classical statistical methods such as t test, ANOVA, multiple regression, and correlation analysis.

- Procedure

1. Lecture 1: Basics of Bayesian statistics
2. Hands-on Workshop 1: Trying Bayesian statistics
3. Hands-on Workshop 2: Foreign language education study using Bayesian statistics
4. Lecture 2: Implications for more advanced modelling

- How to register

1. Registration is open until the maximum capacity is reached. To facilitate the workshop process, please go to the following website and fill in your name, affiliation, and e-mail address. Also, please answer these questions.
   https://goo.gl/forms/BMsDd0XorDye7Ip13
2. If you cannot register using the above website, please email Ryoko FUJITA (Juntendo University) at r-fujita@juntendo.ac.jp

   Let us know the following information when you register the workshop.
   (1) Your name, affiliation, and email address.
   (2) Reason(s) for your interest in Bayesian statistics.
   (3) Kinds of research that you would like to conduct in the future using Bayesian statistics.
   (4) Questions to the instructor, if you have. (Optional)
   (5) Requests for this workshop, or JLTA workshops in general. (Optional)

4. NOTE: The school cafeteria is closed throughout the conference. Participants can have lunch at restaurants near the campus, or buy foods from the CO-OP on campus or a nearby convenience store. Food and drinks are not allowed in the workshop venue. They are only permitted at the Student Hall in the 6th building.

5. Conference Venue

交通アクセス（Access to the University）

To JR Niigata Station

About 2 hours by Joetsu Bullet Train from Tokyo Station. Get off at Niigata Station, the last stop.

Bus (Niigata Kotsu)

Take a bus bound for Niigata Seiryo University - Aquarium Front at the bus terminal at Bandai Exit of Niigata Station and get off at Niigata Seiryo University Front. It takes about 20 minutes.

Or, take a Hamaura-cho Cycle (West Circulation), Bound for Seibu Eigyosho via Hamaura-cho or Bound for Shinano-machi via Hamaura-cho bus and get off at Matsunami-cho 1-chome. It's a 20-minute bus ride and a 4-minute walk from the bus stop.

Train

Get off at Hakusan Station (JR Echigo Line) and walk for 15 minutes.

Car

About 40 minutes by car from Niigata Airport.

JR 新潟駅へ

東京駅より上越新幹線で約 2 時間。

JR 新潟駅万代口バスターミナルから **バス利用**

「浜浦町線＜新潟青陵大学前経由＞水族館前ゆき」 バスで約 20 分、新潟青陵大学前下車。
または、「浜浦町線＜浜浦町経由＞信濃町ゆき」「浜浦町線＜浜浦町経由＞西部営業所ゆき」
バスで約 20 分、松波町 1 丁目下車、徒歩 4 分。

JR 越後線 白山駅から徒歩

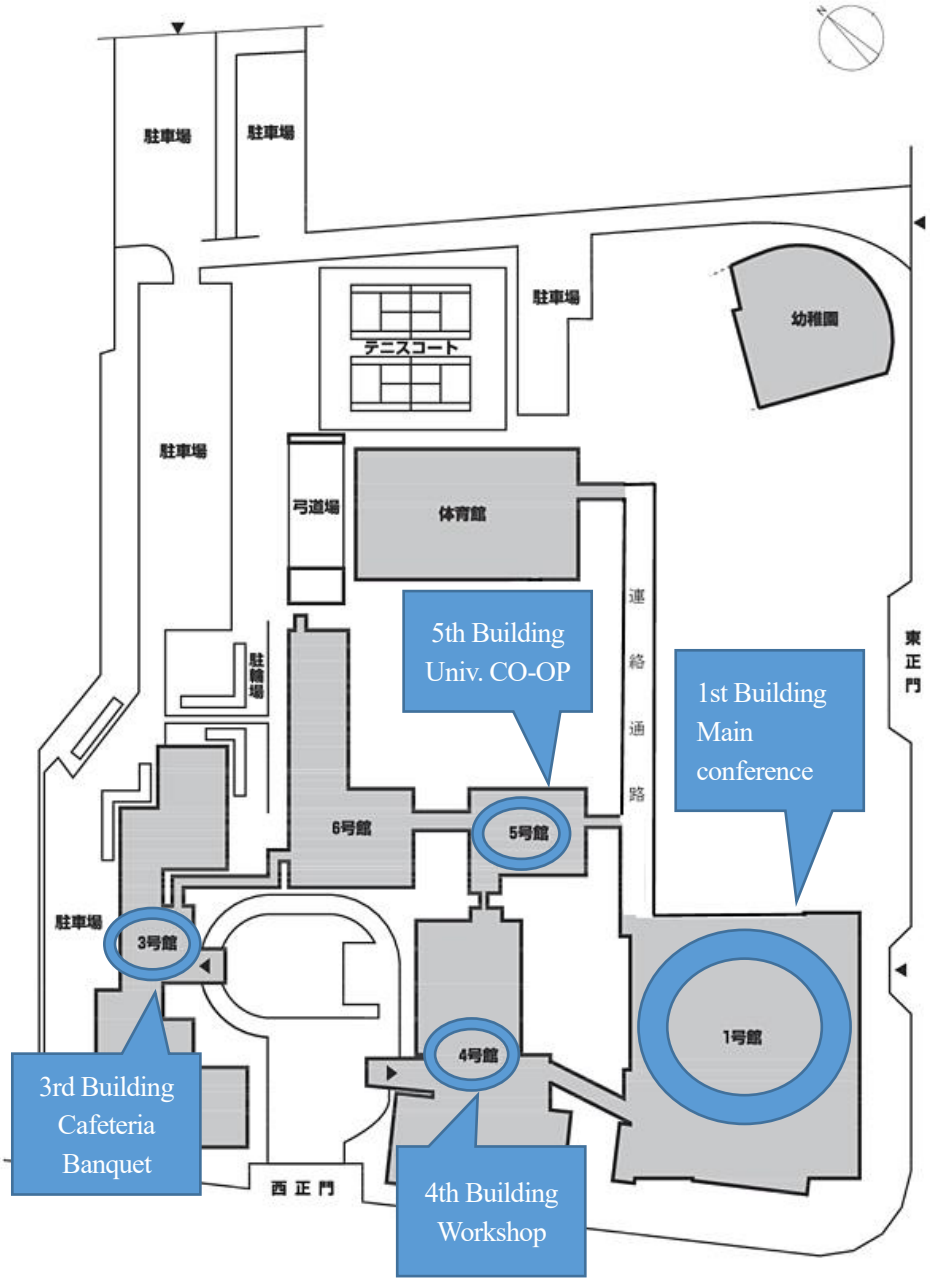
徒歩 15 分。

新潟空港より車

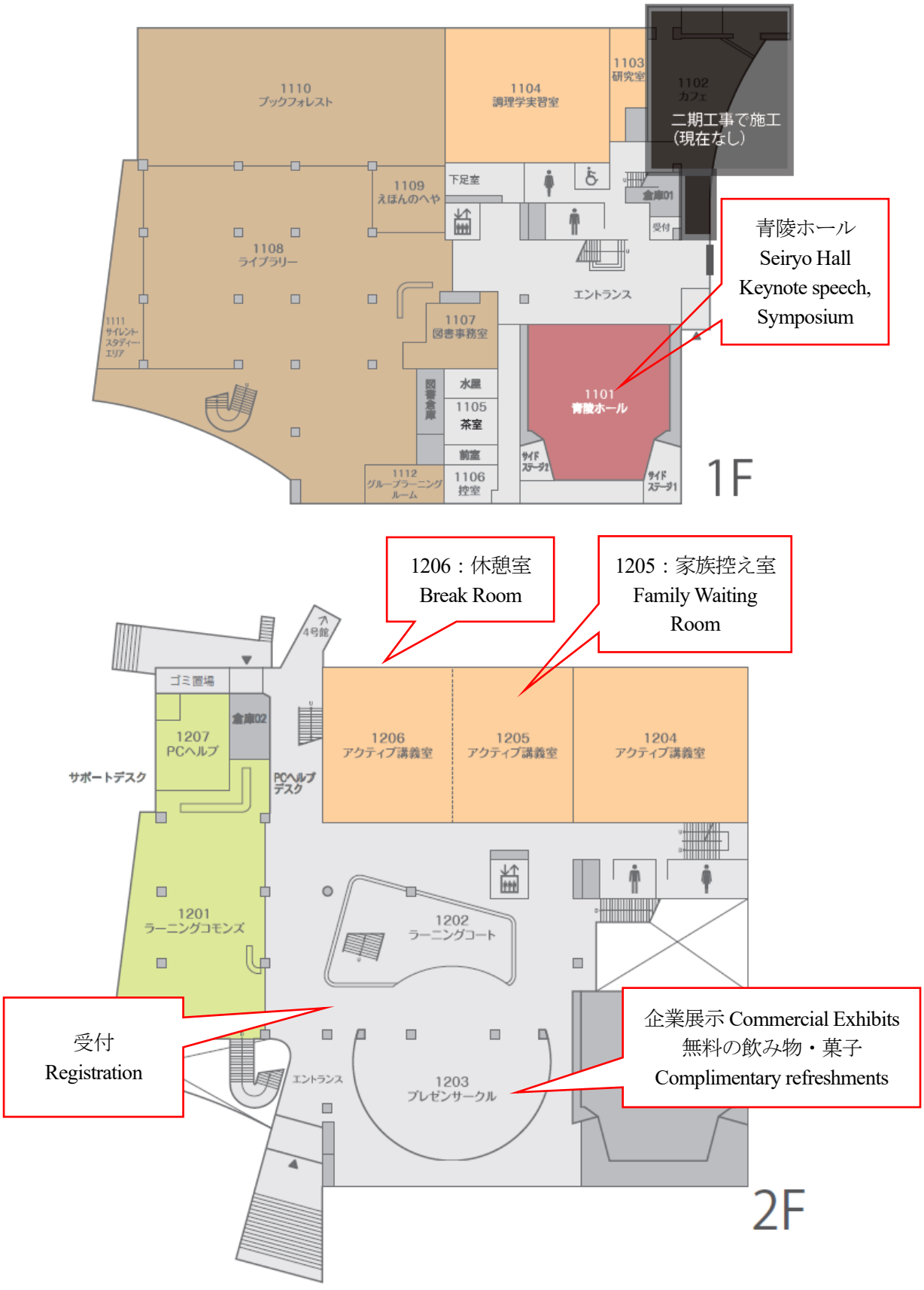
約 40 分。

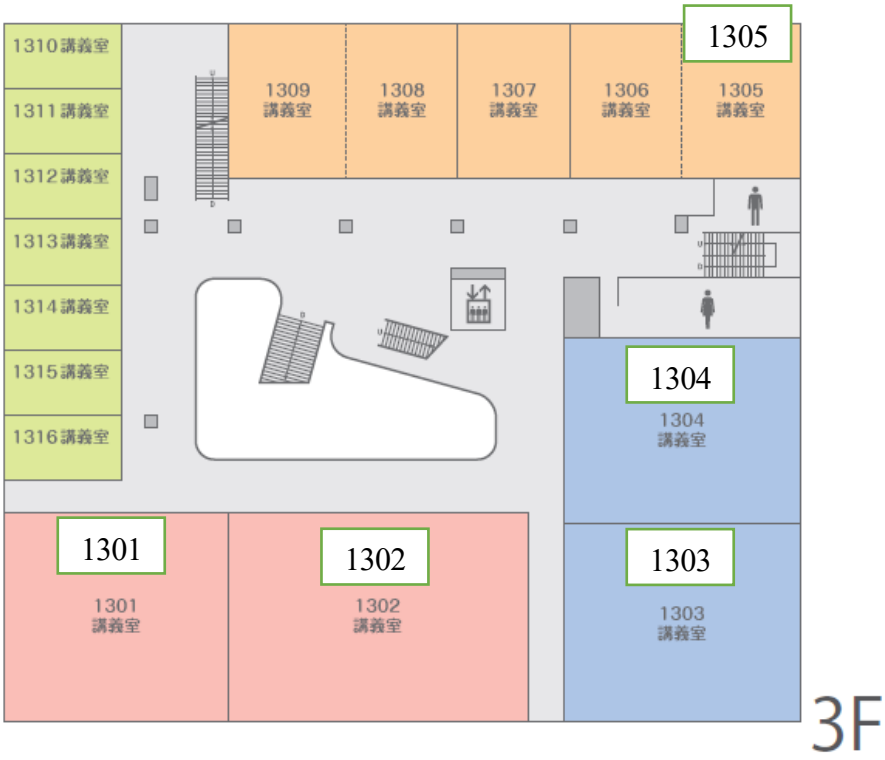
キャンパスマップ（Campus Map）

校舎配置図



1 号館フロアマップ（1st Building Floor Map）




1310講義室
1311講義室
1312講義室
1313講義室
1314講義室
1315講義室
1316講義室
1309 講義室
1308 講義室
1307 講義室
1306 講義室
1305 講義室
1305
1304
1304 講義室
1301
1301 講義室
1302
1302 講義室
1303
1303 講義室
3F

日本言語テスト学会 (JLTA) (2019 年度) 全国研究大会発表要綱
Handbook of the 22nd Conference of the Japan Language Testing Association
発行日：2019 年 8 月 1 日
発行：日本言語テスト学会 (JLTA)　会長：渡部良典 (上智大学)
事務局：〒270-1695　千葉県印西市平賀学園台 1-1　順天堂大学
さくらキャンパス　小泉利恵研究室　TEL: 0476-98-1001 (代表)
FAX: 0476-98-1011 (代表)　E-mail: rie-koizumi@mwa.biglobe.ne.jp
編集：全国研究大会実行委員会
印刷：株式会社国際文献社　〒162-0801 東京都新宿区山吹町358-5
TEL: 03-5937-0249　FAX: 03-3368-2822