

日本言語テスト学会 (JLTA)
第 28 回 (2025 年度) 全国研究大会発表要綱

**Handbook
of
the 28th Annual Conference
of
the Japan Language Testing Association**

Conference Theme:

Rethinking Language Assessment: The Role of Motivation

「言語評価を再考する：学習動機づけの役割」

日時： 2025 年 09 月 13 日 (土) 08:55－17:50 ～ 09 月 14 日 (日) 09:00－12:00
会場： 群馬大学 荒牧キャンパス
〒371-0044 群馬県前橋市荒牧町 4 丁目 2
<https://www.gunma-u.ac.jp/>
主催： 日本言語テスト学会 事務局
〒960-8516 福島県福島市栄町 10-6
久保田恵佑研究室
TEL: 024-581-5533 (直通)
<http://jlta.ac/>



The Japan Language Testing Association
Keisuke Kubota (Fukushima Medical University)

Address: Integrated Center for Science and Humanities /
School of Health Sciences,
Fukushima Medical University,
10-6 Sakae-machi, Fukushima City, Fukushima, 960-8516,
JAPAN

TEL: +81- 024-581-5533 (main switch number)
e-mail: kkubota86@gmail.com /
kkubota@fmu.ac.jp

URL: <http://jlta.ac/>

全国研究大会本部委員 (Annual Conference Committee)

渡部 良典 / Yoshinori Watanabe	(上智大学 / Sophia University)
久保田 恵佑 / Keisuke Kubota	(福島県立医科大学 / Fukushima Medical University)
横内 裕一郎 / Yuichiro Yokouchi	(福島大学 / Fukushima University)
松村 香奈 / Kana Matsumura	(鶴見大学 / Tsurumi University)
前田 啓貴 / Hiroki Maeda	(松山大学 / Matsuyama University)

全国研究大会運営委員 (Annual Conference Steering Committee)

熊澤 孝昭 / Takaaki Kumazawa	(運営委員長・東洋大学 / Toyo University)
高波 幸代 / Sachiyo Takanami	(実行委員長・群馬大学 / Committee Chair, Gunma University)
周 育佳 / Yujia Zhou	(東京外国語大学 / Tokyo University of Foreign Studies)
田中 洋也 / Hiroya Tanaka	(北海学園大学 / Hokkai-Gakuen University)
加藤 万紀子 / Makiko Kato	(東北大学 / Tohoku University)
前田 啓貴 / Hiroki Maeda	(松山大学 / Matsuyama University)
岡 秀亮 / Hideaki Oka	(茨城大学 / Ibaraki University)

全国研究大会実行委員 (Annual Conference Executive Committee)

高波 幸代 / Sachiyo Takanami	(実行委員長・群馬大学 / Committee Chair, Gunma University)
熊澤 孝昭 / Takaaki Kumazawa	(東洋大学 / Toyo University)
周 育佳 / Yujia Zhou	(東京外国語大学 / Tokyo University of Foreign Studies)
田中 洋也 / Hiroya Tanaka	(北海学園大学 / Hokkai-Gakuen University)
加藤 万紀子 / Makiko Kato	(東北大学 / Tohoku University)
前田 啓貴 / Hiroki Maeda	(松山大学 / Matsuyama University)
岡 秀亮 / Hideaki Oka	(茨城大学 / Ibaraki University)

研究発表審査委員 (Paper Presentation Abstract Reviewer)

熊澤 孝昭 / Takaaki Kumazawa	(東洋大学 / Toyo University)
周 育佳 / Yujia Zhou	(東京外国語大学 / Tokyo University of Foreign Studies)
田中 洋也 / Hiroya Tanaka	(北海学園大学 / Hokkai-Gakuen University)

Table of Contents (目次)

1.	Conference Schedule Overview (大会日程表)	2
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)	16
4.	Workshop Information (ワークショップ情報)	36
5.	Conference Venue (会場マップ)	40
6.	Advertisement/Commercial Exhibit Sponsors (広告・展示協賛 企業)	43

1. Conference Schedule Overview

Day 1: September 13, 2025 (Saturday)

Day 1: September 18, 2025 (Saturday)		
08:00—08:30 08:30—	Registration for Commercial Exhibitors Registration	1st Floor, GB2 (#2) Building ※Near the Student Lounge
08:55—09:10	Opening Ceremony	GB155, 1st Floor, GB2 (#2) Building
09:30—10:00	Presentation I	
10:05—10:35	Presentation II	GB101, GB154, GB203
10:35—10:50	Break	Student Lounge, GC102
10:50—12:10	Keynote Speech (Conducted in Japanese)	GB155
12:10—13:40	Lunch Break/JLTA Committee Meetings ※Lunch Room For JLTA Officers: For participants:	Student Lounge, GC102 (GC101) (Student Lounge, GC102, GB101, GB154)
13:40—14:10	Presentation III	GB101, GB154, GB203
14:15—14:45	Presentation IV	GB101, GB154, GB203
14:50—15:20	Presentation V	GB101, GB154, GB203
15:20—15:35	Break	Student Lounge, GC102
15:35—17:05	Symposium (Conducted in English and Japanese)	GB155
17:05—17:20	Break	Student Lounge, GC102
17:20—17:50	Closing Ceremony & JLTA Best Paper / Best Book Award Ceremony	GB155
17:50—18:10	JLTA General Business Meeting	GB155
18:40—20:40	Banquet	University Hall, 1st Floor (Co-op Student Cafeteria)

Day 2: September 14, 2025 (Sunday)

09:00 – 12:00 (08:30 – Registration)	Workshop (Conducted in Japanese)	GB154
--	----------------------------------	-------

Commercial Exhibitors:	Community Room (1st Floor, GB2 (#2) Building)
Lunch Room for Participants:	Student Lounge, GC102, GB101, GB154
Break Room:	Student Lounge, GC102
Family Waiting Room:	GC102
Headquarters:	GC101

✧ Complimentary refreshments are available at GC102.

Program of the 28th JLTA Conference

September 13, 2025 (Saturday)

8:00—	Registration for Commercial Exhibitors	(1st Floor, GB (#2) Building) ※Near the Student Lounge
8:30—	Registration	(1st Floor, GB (#2) Building) ※Near the Student Lounge
	Conference Attendance Fee: Students, JLTA Members & JALT TEVAL SIG Members: Free Non-members: ¥3,000 Please register ahead of time using this URL on Peatix: https://jlta28conference.peatix.com If you're not a JLTA member, kindly pay the registration fee. While on-site registration will be available on the day of the meeting, we strongly recommend registering in advance.	
8:55—9:10	Opening Ceremony	(GB155) Coordinator: Hidetoshi SAITO (Ibaraki University) Greetings: Yoshinori WATANABE (JLTA President; Sophia University)
9:30—10:35	Presentations I and II	(GB101, GB154, GB203) (Presentation: 20 minutes; Discussion: 10 minutes)
10:35—10:50	Break	(Student Lounge, GC102)
10:50—12:10	Keynote Speech	(GB155) Lecturer: Tomohito HIROMORI (Meiji University) Title: Exploring the Link Between Motivation and Assessment in Language Learning: Insights Through the Lens of Engagement Language: Japanese
12:10—13:40	Lunch Break / JLTA Meeting	※Lunch Room For JLTA Officers: (GC101) For participants: (Student Lounge, 1st Floor, GC(#3) Building, GC102, 1st Floor, GB(#2) Building, GB101 • GB154)
13:40—15:20	Presentations III, IV, and V	(GB101, GB154, GB203) (Presentation: 20 minutes; Discussion: 10 minutes)
15:20—15:35	Break	(Student Lounge, GC102)
15:35—17:05	Symposium	(GB155) Language: English and Japanese Theme: Reconsidering Learner Motivation and Assessment in Practical Contexts Chair: Sachiyo TAKANAMI (Gunma University) Panelist 1: Kiwamu KASAHARA (Hokkaido University of Education) Panelist 2: Rintaro SATO (Nara University of Education) Panelist 3: Koki SEKITANI (Toyo Eiwa University) Discussant: Tomohito HIROMORI (Meiji University)
17:05—17:20	Break	(Student Lounge, GC102)

17:20—17:50	Closing Ceremony & JLTA Best Paper / Best Book Award Ceremony (GB155) Coordinator: Takaaki KUMAZAWA (Toyo University) Best Paper Award Recipient: Keiya TANDO (Graduate School, University of Tsukuba) Best Book Award Recipient: Haruhiko MITSUNAGA (Nagoya University)
17:50—18:10	JLTA General Business Meeting (GB155) Chair: Yoshinori WATANABE (JLTA President, Sophia University) Reporter: Keisuke KUBOTA (JLTA Secretary General, Fukushima Medical University)
18:40—20:40	Banquet Coordinator: Sachiyo TAKANAMI (Gunma University) University Hall, 1st Floor (Co-op Student Cafeteria)

September 14, 2025 (Sunday)

8:30—	Registration	(GB154, 1st Floor, GB (#2) Building)
9:00—12:00	Workshop	(GB154)
	Chair: Akiyo HIRAI (University of Tsukuba)	
	Title: Utilizing Generative AI in English Education— From Lesson Preparation to Feedback—	
	Lecturer: Shunsuke TAKAGI (Seiko Gakuin Junior and Senior High School)	
	Language: Japanese	

大会日程表

第1日目：2025年09月13日（土）

08:00－08:30 08:30－	賛助会員 受付 一般会員 受付	GB棟(2号館)1階 ※学生ラウンジ側入り口
08:55－09:10	開会行事	2号館1階：GB155
09:30－10:00	研究発表 I	GB101, GB154, GB203
10:05－10:35	研究発表 II	GB101, GB154, GB203
10:35－10:50	休憩	学生ラウンジ・GB102
10:50－12:10	基調講演 (使用言語：日本語)	GB155
12:10－13:40	昼食/JLTA 委員会 ※昼食会場 役員： 一般参加者：	GC101 学生ラウンジ, 3号館1階 GC102 2号館1階 GB101・GB154
13:40－14:10	研究発表 III	GB101, GB154, GB203
14:15－14:45	研究発表 IV	GB101, GB154, GB203
14:50－15:20	研究発表 V	GB101, GB154, GB203
15:20－15:35	休憩	学生ラウンジ・GB102
15:35－17:05	シンポジウム (使用言語：英語・日本語)	GB155
17:05－17:20	休憩	学生ラウンジ・GB102
17:20－17:50	閉会行事 & 最優秀論文賞・著作賞授与式	GB155
17:50－18:10	JLTA 総会	GB155
18:40－20:40	懇親会	大学会館1階 学生食堂 (生協)

第2日目：2025年09月14日（日）

09:00－12:00 (08:30－受付)	ワークショップ (使用言語：日本語) ※GB154室前で受付を行います。
---------------------------	---

協賛企業展示： コミュニティルーム
一般参加者昼食： 学生ラウンジ・GC102・GB101・GB154
休憩室： 学生ラウンジ・GC102
家族控え室： GC102
大会本部： GC101

※ 無料の飲み物やお菓子は、GC棟(3号館)1階 102室 にございます。

日本言語テスト学会第 28 回全国大会プログラム

第 1 日目：2025 年 09 月 13 日（土）

- 8:00－ 賛助会員 受付
(2 号館 (GB 棟) 1 階, 学生ラウンジ側入り口)
- 8:30－ 一般 受付
(2 号館 (GB 棟) 1 階, 学生ラウンジ側入り口)
学会参加費：JLTA 会員・JALT TEVAL SIG 会員：無料
未会員：3,000 円
以下の URL (Peatix) から事前参加登録をしてください。
また、未会員の方は参加費をお支払いください。
<https://jlta28conference.peatix.com>
※当日会場で受付も可能ですが、できる限り事前登録にご協力ください。
- 8:55－9:10 開会行事 (2 号館 (GB 棟) 1 階 GB155)
総合司会：齋藤 英敏 (茨城大学)
挨拶：渡部 良典 (JLTA 会長・上智大学)
- 9:30－10:35 研究発表 I・II (発表 20 分, 質疑応答 10 分) (GB101, GB154, GB203)
- 10:35－10:50 休憩 (学生ラウンジ, GC102)
10:50－12:10 基調講演 (GB155)
司会：渡部 良典 (JLTA 会長・上智大学)
演題：言語学習における動機づけと評価の接点：エンゲージメントの観点から
講師：廣森 友人 (明治大学)
使用言語：日本語
- 12:10－13:40 昼食
昼食会場
JLTA 役員： (GC101)
一般参加者： (学生ラウンジ, GC102, GB101, GB154)
- 13:40－15:20 研究発表 III・IV (発表 20 分, 質疑応答 10 分) (GB101, GB154, GB203)
- 15:20－15:35 休憩 (学生ラウンジ, GC102)
15:35－17:05 シンポジウム (GB155)
テーマ：実践的な場面から再考する学習者の動機づけと評価
使用言語：英語・日本語
司 会：高波 幸代 (群馬大学)
パネリスト 1：笠原 究 (北海道教育大学)
パネリスト 2：佐藤 臨太郎 (奈良教育大学)
パネリスト 3：関谷 弘毅 (東洋英和女学院大学)
討 論 者：廣森 友人 (明治大学)
- 17:05－17:20 休憩 (学生ラウンジ, GC102)

17:20－17:50	閉会行事 & 最優秀論文賞授与式 司会：熊澤 孝昭（東洋大学） 2024 年度 JLTA 最優秀論文賞受賞者：丹藤 慧也 (筑波大学大学院) 2025 年度 JLTA 著作賞受賞者：光永 悠彦 (名古屋大学)	(GB155)
17:50－18:10	JLTA 総会 議長選出 司会：渡部 良典（JLTA 会長・上智大学） 報告：久保田 恵佑（JLTA 事務局長・福島県立医科大学）	(GB155)
18:40－20:40	懇親会 司会：高波 幸代（群馬大学）	大学会館 1 階 (生協，学生食堂)

第 2 日目：2025 年 09 月 14 日（日）

08:30－	受付	2 号館（GB 棟）1 階 GB154 室前）
09:00－12:00	ワークショップ 司会：平井 明代（筑波大学） 題目： 「英語教育における生成 AI の活用～授業準備からフィードバックまで～」 講師：高木 俊輔（聖光学院中学校高等学校） 使用言語：日本語	(GB154)

Presentation Overview

Time	Part	GB154	GB101	GB203	GB155 (大講堂)
09:30 – 10:00	I	YOU	JOHNSON	CONNOR†	
10:05 – 10:35	II	ARAI	ALLEN	WILSON†	
10:35 – 10:50		Break			
10:50 – 12:10					Keynote
12:10 – 13:40		Lunch Break			
13:40 – 14:10	III	NISHIZAWA	WALTERS	KOIZUMI	
14:15 – 14:45	IV	TAHARA	梅谷†	GROGAN†	
14:50 – 15:20	V	小室	MILLER †	中野†	
15:20 – 15:35		Break			
15:35 – 17:05					Symposium

† Assessment practice presentation

Presentation Details

GB155 (Keynote speech & Symposium)

Time	Presenter (Affiliation)	Title (Page)
10:50 – 12:10	Keynote speech Tomohito HIROMORI (Meiji University)	Title: Exploring the Link Between Motivation and Assessment in Language Learning: Insights Through the Lens of Engagement
15:35 – 17:05	Symposium: Chair Sachiyo TAKANAMI (Gunma University) Panelist 1 Kiwamu KASAHARA (Hokkaido University of Education) Panelist 2 Rintaro SATO (Nara University of Education) Panelist 3 Koki SEKITANI (Toyo Eiwa University) Discussant Tomohito HIROMORI (Meiji University)	Title: Reconsidering Learner Motivation and Assessment in Practical Contexts 1. Encouraging Use of Small Vocabulary Quizzes to Motivate EFL Learners 2. Beyond Current Trends: Rethinking Language Activities and Assessment in EFL Classrooms through (my)WTC Research 3. Learners' Psychological Experiences in Speaking Tests: What Types of Tests Motivate Them?

GB154

Part	Presenter (Affiliation)	Title (Page)
I 9:30 – 10:00	Ga Eul YOU (University of Western Ontario, Canada) Ruslan SUVOROV (University of Western Ontario, Canada)	How does music affect test performance and anxiety? Evidence from electrodermal activity and interviews
II 10:05 – 10:35	Yuya ARAI (Chiba University) Yumi FUJITA (Waseda University)	Number of plays and anxiety in the Common Test listening test: An idiodynamic study
III 13:40 – 14:10	Hitoshi NISHIZAWA (Reitaku University)	Examining the validity of the TOEIC L&R for placement purposes
IV 14:15 – 14:45	Tatsuro TAHARA (Waseda University)	A Content Analysis of English Reading Sections in External Tests Used for University Admission Purposes in Japan
V 14:50 – 15:20	小室竜也 (東北大学) 及川凱亜 (東北大学)	日本人 EFL 学習者のための語彙サイズテスト VST-NJ8 の妥当性検証：「わからない」選択肢の追加と回答時間の関係性

GB101

Part	Presenter (Affiliation)	Title (Page)
I 9:30 – 10:00	Michelle JOHNSON (Nihon University College of Industrial Technology)	Exploring the Evolution of a Novice Instructor's Assessment Literacy in Higher Education
II 10:05 – 10:35	David ALLEN (Ochanomizu University)	Investigating the predicted washback of an English speaking exam used for university admissions in Japan
III 13:40 – 14:10	F. Scott WALTERS (Seoul National University of Science and Technology)	How to Motivate Development of Tests of Interactive Competence
IV 14:15 – 14:45	梅谷 博之 (武蔵野大学) 根岸 雅史 (東京外国語大学) 周 育佳 (東京外国語大学) ミラー マシュー (東京外国語大学) 川本 夢子 (東京外国語大学) 投野 由紀夫 (東京外国語大学)	「CEFR-J×28 多言語 Can Do テスト」 (リーディング) の開発と課題
V 14:50 – 15:20	Matthew MILLER (Tokyo University of Foreign Studies) Masashi NEGISHI (Tokyo University of Foreign Studies) Yujia ZHOU (Tokyo University of Foreign Studies) Yukio TONO (Tokyo University of Foreign Studies) James CARPENTER (Tsurumi University)	Refining the Design of CEFR-J English Listening Tests through Empirical Data and Student Perspectives

GB203

Part	Presenter (Affiliation)	Title (Page)
I 9:30 – 10:00	Stuart CONNOR (Pearson)	Dialogic Critical Thinking English Test – A New Frontier in AI-Powered Language Assessment
II 10:05 – 10:35	Gordon WILSON (Hannan University) Myles GROGAN (Osaka Ohtani University)	Evaluating the Efficacy of AI Driven Approach to the Development and Revision of a Placement Test
III 13:40 – 14:10	Rie KOIZUMI (University of Tsukuba) Yo IN'NAMI (Chuo University)	Perceptions of an AI-scored classroom-based speaking test: Comparisons between Japanese in-service and pre-service teachers
IV 14:15 – 14:45	Myles GROGAN (Osaka Ohtani University)	Commercial tests or AI-assisted streaming?: A comparison at the university level
V 14:50 – 15:20	中野聡子 (群馬大学) 浅田裕子 (昭和女子大学) 藤田元 (上智大学)	日本手話到達基準「JSL スタANDARD」におけるコミュニケーション言語活動の策定

2. From the JLTA Office: Information for Conference Participants

Registration

1. You can pre-register on Peatix at <https://jlta28conference.peatix.com>. While we primarily require pre-registration, we'll also accept participants on the day of the conference.
2. The conference registration site is in front of Student Lounge Entrance (1st Floor, 2nd Building).
3. The conference attendance fee is free for students and members and ¥3,000 for non-members.
4. If non-members apply for membership at the registration desk, the conference attendance fee will be free. 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 ¥5,000, which attendees should pay on Peatix or at the venue on the conference day. The banquet will be held at University Hall, Co-op Student Cafeteria.

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. 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.
3. 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.
4. The family waiting room is in GC102 (1st floor, GC(#3) Building). Complimentary refreshments are available in the same room. Feel free to enjoy them.

Lunch and Participants' Lounge, Etc.

1. The following rooms are available for lunch and breaks.
Lunch: Please use the Student Lounge, GC(#3) Building, Room GC102, and GB(#2) Building, Rooms GB101 and GB154.
Breaks: Please use the Student Lounge, GC(#3) Building, Room GC102.
2. Complimentary refreshments are available in GC102.
3. As it is the weekend, the cafeteria will be closed and cannot be used. Convenience stores, drugstores, and other shops are located within a 5-minute walk from the main gate.

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: kkubota86@gmail.com (Keisuke KUBOTA)

To Presenters

1. Presenters will have 20 minutes to present their paper, followed by 10 minutes for discussion.
2. There will be no chairperson in the presentation room. A timekeeper 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. We recommend using an HDMI cable for connection. Audio output is also available by default. There will not be an audio terminal connector (for PC connection through a stereo mini plug). If necessary, please prepare an adaptor. Mac users should bring their own Mini DisplayPort to VGA Adapter. Third-party adapters do not work properly sometimes.
5. Wi-Fi is not available inside the building. Gunma University participates in the international

academic wireless LAN roaming system “eduroam” through “eduroam JP” operated by NII. You can use this service with an eduroam-compatible account issued by your home institution. For details, please check with your institution.

6. Please bring handouts in case your PC or the projector does not work.
7. If you need a letter of invitation/participation, contact Keisuke Kubota (JLTA Secretary General) at kkubota86@gmail.com
8. Name tags including participation certificates will be provided on the day of the conference.

学会事務局からのお知らせ

大会参加者へのご案内

■受付

1. 参加登録は、Peatix (<https://jlta28conference.peatix.com>) にて行います。事前登録 (9/10 まで) にご協力いただけますと幸いです。大会当日の参加も受け付けております。
2. 受付は、2号館 (GB 棟) 1階、学生ラウンジ側の入り口付近にて行います。
3. 学会参加費は、学生、会員は無料、未会員 3,000 円です。
4. 未会員の方でも、受付で入会手続きを行えば学会参加費は無料となります。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。
5. 学会中は、名札をお付けください。
6. 懇親会費は 5,000 円です。参加者は研究大会当日に会場にてお支払いください。懇親会は大学生協、学生食堂にて開かれます (上限 50 名)。

■家族控室

1. 参加者のご家族 (中学生以上) のための控室を設置します。発表・講演等には参加しない家族が待機できる部屋ですが、一般参加者の休憩室も兼ねています。お含みおきください。
2. スタッフも配置しませんので、自己責任での中学生以上のご利用に限定させていただきます。
3. 発表・講演等には参加せず、家族控室のみ利用される参加者ご家族は、大会参加費は無料です。参加者が受付する際に、「参加者ご家族」用の名札を受け取ってください。
4. 家族控室は 3号館 (GC 棟) 1階、GC102 室にあります。無料の飲み物やお菓子も同室にございます。参加者ご家族の方もお召し上がりいただけます。

■昼食・休憩室等

昼食：学生ラウンジ、3号館 (GC 棟) GC102 室、および 2号館 (GB 棟) GB101 室、GB154 室をご利用いただけます。

休憩：学生ラウンジ、3号館 (GC 棟) 102 室をご利用いただけます。

2. 無料の飲み物やお菓子は、3号館 (GC 棟) 1階 102 室にございます。
3. 週末につき、学食の営業はなく、使用もできません。正門徒歩 5 分以内の場所に、コンビニエンス・ストアとドラッグストア等がございます。

■宿泊

宿泊の斡旋はいたしておりません。

■喫煙

キャンパス内は建物外も含め、禁煙です。

■緊急連絡先の E メールアドレス kkubota86@gmail.com (久保田恵佑、JLTA 事務局長)

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. 司会者はありません。タイムキーパーが残り時間を提示します。
3. 受付を済まされ、発表開始 5 分前には、発表会場にお越しください。
4. ご発表にはご自分のコンピューターを持参してください。HDMI ケーブルでの接続を推

奨めています（音声出力もデフォルトで使用可能です）。音声ケーブル（ステレオミニプラグによる PC 接続用）は発表会場にはございません。必要に応じて、変換用アダプター、並びに Mac 用のケーブルはご自身でご準備ください。持ち込んだアダプターは機能しないことが時折ございます。

5. 館内では Wi-fi がご利用いただけません。群馬大学は NII が運用する「eduroam JP」を介して、国際学術無線 LAN ローミング基盤「eduroam」に参加しています。各機関で発行された eduroam 対応のアカウントで本サービスをご利用いただけます。詳しくはご所属の機関にてご確認ください。
6. 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。
7. 出張依頼状などが必要な方は、kkubota86@gmail.com（久保田恵佑, JLTA 事務局長）までご連絡ください。
8. 参加証明を含むネームタグは、大会当日に配布します。

3. Abstracts (発表要旨)

Keynote Speech (GB155)

10:50－12:10

Exploring the Link Between Motivation and Assessment in Language Learning: Insights Through the Lens of Engagement

言語学習における動機づけと評価の接点：エンゲージメントの観点から

Tomohito Hiromori (Meiji University)

廣森 友人 (明治大学)

本講演では、大会テーマ「Rethinking Language Assessment: The Role of Motivation（言語評価を再考する：学習動機づけの役割）」に応え、“Engagement”（エンゲージメント）という概念を介して、言語テスト研究と動機づけ研究の接点を探ります。従来、言語テスト研究は妥当性・信頼性を重視し、学習者の認知的側面に焦点を当てて発展してきました。一方、動機づけ研究は学習者の心理的・情意的側面に注目してきました。しかし教育実践においては、テスト・パフォーマンスにも情意的要因が大きく影響するなど、両者は密接に関連しています。こうした中、エンゲージメントは、認知・行動・情意・社会的側面を統合的に捉える概念として注目を集めており（廣森・小金丸, 2024; 廣森・和田, 2024）、両領域を架橋する枠組みとして極めて有望です。

本講演では、エンゲージメント概念の理論的背景と、行動指標を通じた「見える化」という測定的観点を概観し、国内外の先行研究や発表者自身が関わった実証研究を紹介します。さらに、過去に *JLTA Journal* に掲載された研究論文も参照しながら、ペアワーク活動、スピーキングテスト、ポートフォリオ評価といった具体的実践例を通じて、エンゲージメントの視点が評価や学習支援にどのような新たな可能性を拓くかを論じます。評価と動機づけのより有機的な統合を目指し、学習者の主体的な学びの促進に貢献するとともに、公平かつ包括的な評価の可能性を考察します。

参考文献

- 廣森友人・小金丸倫隆 (編著). (2024). 『エンゲージメント×英語授業－「やる気」と「意欲」を引き出す授業のつくり方』 明治図書.
- 廣森友人・和田玲 (編著). (2024). 『エンゲージメントを促す英語授業－やる気と行動をつなぐ新しい動機づけ概念』 大修館書店.

Symposium (GB155)

15:35－17:05

Reconsidering Learner Motivation and Assessment in Practical Contexts

実践的な場面から再考する学習者の動機づけと評価

Panelists: Kiwamu KASAHARA (Hokkaido University of Education)

Rintaro SATO (Nara University of Education)

Koki SEKITANI (Toyo Eiwa University)

Discussant: Tomohito HIROMORI (Meiji University)

(Panelist 1)

Encouraging Use of Small Vocabulary Quizzes to Motivate EFL Learners

Kiwamu Kasahara (Hokkaido University of Education)

This presentation proposes an effective approach to using small vocabulary quizzes in class to enhance L2-L1 paired-associate learning (PAL), in which learners connect L2 forms with L1 equivalents. This process can be strengthened by using small vocabulary quizzes. A key to success is implementing *spaced retrieval practice*, which combines spaced learning and retrieval practice. Research has shown that spaced learning, which involves inserting intervals between encounters with the same items, is more effective for long-term retention than massed learning, where no such intervals are provided. Additionally, retrieval practice, effortfully recalling previously learned items, helps strengthen form-meaning connections in PAL.

One way to incorporate the testing effect into vocabulary learning is using cumulative tests (CTs). CTs repeatedly include previously tested items in subsequent quizzes, requiring learners to revisit vocabulary items they have already studied. Nakata et al. (2021) found that CTs were more effective than traditional tests, which only assesses newly introduced items. However, CTs may have a drawback: items introduced later tend to receive less attention, making them harder to retain.

To address this issue, Kanayama et al. (2023) proposed the *random-selection test* (RST). In this scheme, learners were given a full L2-L1 vocabulary list in advance, and a set number of items are randomly selected from the list for each small quiz. Kanayama et al. found that RSTs were more effective for vocabulary retention than CTs, likely because learners paid more attention to each item on the list.

However, both CTs and RSTs can place a heavy burden on learners, as they require preparation for an increasing number of items in CTs, or for the entire list every time in RSTs. This could lead to a decline in learner motivation. To address this issue, this presentation proposes oral vocabulary quizzes in pairs as an alternative. Learners are provided with a word list covering all the L2-L1 pairs for an entire unit in their textbook. As they progress through the unit over several lessons, they engage in short, oral pair quizzes during each lesson. Through these enjoyable and interactive activities, learners can cumulatively reinforce the meaning-form connections of L2 vocabulary items in a low-pressure environment.

References

- Kanayama, K., Iwata, A., & Kasahara, K. (2023). Effects of random selection tests on second language vocabulary learning: a comparison with cumulative tests. *The Journal of International Review of Applied Linguistics in Language Teaching*, 63(1), 391-412..
- Nakata, T., Tada, S., Mclean, S., & Kim, Y. A. (2021). Effects of distributed retrieval practice over a semester: Cumulative tests as a way to facilitate second language vocabulary learning. *TESOL Quarterly* 55(1), 248-270.

(Panelist 2)

Beyond Current Trends: Rethinking Language Activities and Assessment in EFL Classrooms through (my)WTC Research

Rintaro SATO (Nara University of Education)

This presentation offers suggestions for performance-based assessment in EFL classrooms, drawing on my recent and on-going research on Willingness to Communicate (WTC). WTC refers to the likelihood that learners will choose to use the target language when given the chance (MacIntyre, Clément, Dörnyei & Noels, 1998). My research (Sato, 2023a) shows that WTC is not a fixed trait but fluctuates depending on context, task type, and emotional factors. For instance, lower-level learners often show higher WTC when they feel secure or prepared, while advanced learners respond better to tasks that allow personal expression. I have also found that the strategic use of L1 can support WTC across proficiency levels (Sato, 2023b).

Performance tests are widely used in Japan, typically involving tasks such as speeches, interviews, and discussions. However, not all learners benefit equally from these formats. Based on my findings, I suggest designing different types of tasks for different learners: accuracy-focused tasks like reading aloud and summarizing for lower-level learners, and fluency-oriented tasks such as debates or role plays for higher-level learners. Preparation time and L1 support are helpful for both groups.

However, while motivation is often seen as essential for good performance, I question whether high motivation or high WTC is always necessary. The meaning of learner silence, which is often regarded as the lack of motivation, should also be reconsidered (Sato, 2024). A calm and stable psychological state may be more important for sustained engagement. More fundamentally, I raise concerns about whether internal factors like “motivation” or “a proactive attitude” can—or should—be assessed at all. These are subjective, context-dependent, and difficult to measure fairly. Rather than trying to quantify such qualities, we should focus on creating supportive learning environments and assessing what learners can actually do with the language.

I hope this presentation will lead to highly engaged spoken exchanges with fellow presenters and passionate live discussions with the audience.

References

- MacIntyre, P.D., Clément, R., Dörnyei, Z., & Noels, K.A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journal*, 82, 545–562
- Sato, R. (2024). Exploring the silence of Japanese EFL learners: Its relationship with the degree of willingness to communicate (WTC). *Language Teaching Research*.
- Sato, R. (2023a). Examining fluctuations in the WTC of Japanese EFL speakers: Language proficiency, affective and conditional factors. *Language Teaching Research*, 27(4), 974-994.
- Sato, R. (2023b). Japanese EFL speakers’ willingness to communicate in L2 conversations: The effects of code-switching and translanguaging. *TESL-EJ*, 27(3).

Learners' Psychological Experiences in Speaking Tests: What Types of Tests Motivate Them?

Koki SEKITANI (Toyo Eiwa University)

This presentation highlights the importance of learners' psychological experiences in English speaking tests drawing on findings from our recent studies. The first study (Sekitani & Mitsuta, 2025) compared two speaking test modes—video-conferenced (VC) and semi-direct telephone-based (SD)—to examine differences not only in learner output but also in their emotional and motivational responses. Thirty-three Japanese high school students took both tests and completed post-test questionnaires. Their speech samples, psychological responses, and washback effects (i.e., how the tests impacted learning) were analyzed.

The results showed that while the SD mode elicited more syntactically complex language, it came at the cost of fluency and accuracy. Learners strongly preferred the VC mode, exhibiting greater confidence and a sense of fully demonstrating their abilities. The interaction with the interlocutor in the VC mode was perceived as a natural conversation enhancing the perception of fairness. Participants felt that the VC mode fostered better learning behaviors, whereas the SD mode encouraged concrete learning strategies focused on accuracy and prepared templates.

Supporting this, an earlier study (Sekitani, 2022) found that learners made more perceived progress when peers actively listened and asked questions during practice. This suggests that the presence and behavior of an interlocutor—even outside testing contexts—play a key role in learning. These insights call for careful interlocutor training and thoughtful test design that balances authenticity, fairness, and motivational impact. As Galaczi and Taylor (2018) point out, we must also consider how technology transforms both delivery and the nature of what is being assessed. These considerations are essential for ensuring that speaking tests remain meaningful and fit for purpose.

References

- Galaczi, E., & Taylor, L. (2018). Interactional competence: Conceptualisations, operationalisations, and outstanding questions. *Language Assessment Quarterly*, 15(3), 219–236. <https://doi.org/10.1080/15434303.2018.1453816>
- Sekitani, K. (2022). The importance of a peer who plays a listener's role in English as a foreign language speaking practice: Effects of teaching active listening and questioning skills. *JACET Journal*, 66, 75–96. https://doi.org/10.32234/jacetjournal.66.0_75
- Sekitani, K., & Mitsuta, R. (2025). *Comparing videoconferencing and human-to-machine modes in speaking assessment: Holistic ratings, analytical measures, affective factors, and washback effects* [Manuscript in preparation]. Department of Human Sciences, Toyo Eiwa University.

Paper Session

Room GB154 Part I (9:30-10:00)

研究発表

How does music affect test performance and anxiety? Evidence from electrodermal activity and interviews

Ga Eul YOU (University of Western Ontario, Canada)
Ruslan SUVOROV (University of Western Ontario, Canada)

Many language learners experience significant anxiety during L2 assessment tasks (Huang, 2018), which can undermine their motivation to engage in L2 learning (MacIntyre, 2003). Anxiety is particularly detrimental to listening (Elkhafaifi, 2005), the most complex language skill to develop and assess due to its transient nature (Field, 2019). One way to reduce anxiety and facilitate L2 learners' task performance is music listening (Degrave, 2019). Previous research has shown that listening to calm music can lower blood pressure and heart rate (Lilley et al., 2014) and reduce subjective stress levels (Linnemann et al., 2015). However, the exact effect of such music on L2 learners' assessment outcomes remains unclear due to the inconsistent findings, which can be partly attributed to reliance on survey-based methods that are prone to various biases. One way to gather more objective data about L2 learners' anxiety is by measuring their electrodermal activity (EDA). EDA can detect subtle changes in skin conductance resulting from sweat gland activation that is caused by emotional arousal such as anxiety or stress. To our knowledge, no studies in the field of language testing have utilised EDA to examine L2 learners' test anxiety.

To address this gap, our study investigated whether listening to calm, andante (moderately slow) music could alleviate L2 learners' anxiety and improve their performance on L2 listening assessments. Specifically, we employed a mixed-methods, within-participants design with 39 L2 learners of English completing two sections of the IELTS Listening test: one before and one after a 6-minute music intervention. Anxiety was measured using EDA—including skin conductance level (SCL) and skin conductance response (SCR) metrics—as well as through semi-structured interviews. Statistical analyses revealed that, contrary to expectations, test scores significantly declined following the music intervention ($t(39) = 5.81, p < .001, 95\% \text{ CI } [1.92, 3.98], \text{Cohen's } d = 1.88$), indicating that music did not enhance test performance. However, EDA data told a different story: The frequency of emotional arousal responses (SCR event count) and mean SCL were both significantly lower after the music intervention (i.e., SCR event count: $t(34) = 4.45, p < .001, 95\% \text{ CI } [8.88, 23.81], \text{Cohen's } d = 21.74$; mean SCL: $t(34) = 1.93, p = .03, 95\% \text{ CI } [-.03, 1.29], \text{Cohen's } d = 1.93$), suggesting a reduction in test anxiety. Interview responses largely corroborated these findings, with many participants reporting reduced anxiety after listening to music, even if this did not translate into improved scores.

Keywords

anxiety, L2 listening, music, electrodermal activity, EDA

Number of plays and anxiety in the Common Test listening test: An idiodynamic study

Yuya ARAI (Chiba University)

Yumi FUJITA (Waseda University)

The Common Test for University Admissions (CT) in Japan introduced single-play items in its English listening test in 2021. Despite the CT's high-stakes nature, little is examined on how the number of plays affects test-takers' anxiety in the listening test. Listening is an anxiety-provoking skill (Ji et al., 2022) and negatively related to anxiety (Brunfaut & Révész, 2015; In'nami et al., 2022; Li, 2022). The number of plays in listening is linked to anxiety (Buck, 2001; Field, 2019, 2023), and previous studies have suggested that double plays reduce anxiety (Field, 2015; Holzknecht & Harding, 2024). Affective factors, including anxiety, are also important in language testing as part of test-taker characteristics affecting test performance (Bachman & Palmer, 2010; O'Sullivan, 2000; O'Sullivan & Green, 2011). While some previous studies have examined anxiety in foreign language listening tests (Chang & Read, 2008; Elkhafaifi, 2005; Golchi, 2012; In'nami, 2006; Winke & Lim, 2017), a test validation study is needed on the relationship between the number of plays and anxiety in the CT listening test. Furthermore, previous studies on listening anxiety have mainly used questionnaires, which have difficulty capturing the dynamic change of anxiety during the tasks/tests.

With this background, this study explored six Japanese female high school students' anxiety during the CT listening test and its relationship to the number of plays by employing the idiodynamic method (MacIntyre, 2012; MacIntyre & Legatto, 2011). This method aimed to capture the moment-to-moment change of anxiety during the test and the reasons for the change. The participants took a trial version of the CT listening test while their test performance was video-recorded. Immediately after the test, they rated the intensity of anxiety on a moment-by-moment basis using a six-point Likert scale while watching the video. Then, a stimulated recall interview was conducted to understand the reasons for their rating.

Six themes were generated by conducting reflexive thematic analysis (Braun & Clarke, 2006, 2022). They suggested that single plays caused anxiety and poorer comprehension. There was also a bidirectional relationship between test performance and anxiety. Meanwhile, the number of plays was not the single factor affecting them (e.g., task-related factors). Furthermore, the relationship between the number of plays and anxiety was mediated by learner-related factors, which also contributed to the students' various preferences for the number of plays in the CT listening test. Based on these findings, this presentation will discuss implications for CT developers, teachers, and researchers.

Keywords

The Common Test for University Admissions, listening test, single/double play, anxiety, idiodynamic method

Examining the validity of the TOEIC L&R for placement purposes

Hitoshi NISHIZAWA (Reitaku University)

The TOEIC L&R is widely used for placement in Japanese universities. Despite its popularity, little has been done to examine the validity of the use. The current study presents a case study at a university, where the total score of TOEIC L&R has been used for placement to keep the proficiency homogenous within a class as much as possible. In the university, there are two English majors. One major has six different levels, while the other major has three. Both majors have three compulsory English courses. One focuses on reading and writing, while two focus on speaking and listening. While the TOEIC L&R does not assess speaking and writing, all the three courses share some elements with the test (i.e., reading, listening). For the present study, test data from 154 students were used.

Among many aspects of validity, the present study examines the consequence of the TOEIC L&R (Chapelle, 2020; Kane, 2013; Messick, 1987). To do this, data were taken from three sources for triangulation. One was a student questionnaire, which asked the appropriateness of their placement for each class and affective factors (e.g., anxiety). Second was teacher judgment, in which the instructors assessed the appropriateness of their students' placement. The student questionnaire and teacher judgment were conducted after one semester to allow ample time and experience to assess the appropriateness of the placement. Third was the LANGX (Saeki et al., 2024), an AI-based speaking test which produces scores ranging from zero to 7. The TOEIC L&R and LANGX were administered at the beginning of the semester, which allows the direct comparison. Note that the placement was decided based on the TOEIC. The LANGX results were used to estimate how the placement might have differed had LANGX been used instead.

Preliminary analysis found a small correlation, calculated by Spearman's rho, between the TOEIC score (mean = 340.42, SD = 118.13, median = 320) and LANGX (mean = 1.51, SD = 0.80, median = 1.33; total = .35, $p < .001$; listening = .34, $p < .001$; reading = .30, $p < .001$). The weak association was reflected in the placement, where a total of 95 students hypothetically placed differently with LANGX. Among them, 52 students showed the difference by two levels. These results will be interpreted in light of the student survey and teacher judgment.

Keywords

TOEIC, placement, validity

A Content Analysis of English Reading Sections in External Tests Used for University Admission Purposes in Japan

Tatsuro TAHARA (Waseda University)

Since the 2010s, the role of English proficiency tests in university entrance exams has sparked debate in Japan, particularly regarding external commercial tests (e.g., Haebara, 2018). While score conversion has been widely discussed, less attention has been given to differences in the reading constructs assessed by these tests. To address this gap, this study analyzes the reading sections of five major tests used in Japanese university admissions: two national standardized tests—the Common Test for University Admissions (Common Test) and the National Center Test for University Entrance Examination (NCT)—and major external tests, including the EIKEN Test (Grades 2 and Pre-1) and the Test of English for Academic Purposes (TEAP).

Drawing on the frameworks of Bachman (1990) and Kunnan and Carr (2017), a multi-method approach was employed. Textual features were analyzed using Coh-Metrix (McNamara et al., 2014) for readability (e.g., Flesch–Kincaid Grade Level) and LexTutor (Cobb, n.d.) for vocabulary levels (e.g., coverage of the New General Service List [NGSL] and the New Academic Word List [NAWL]). All items were manually coded for aspects of reading assessed and the scope of textual engagement. For robust inferences, group differences in the textual measures were tested using the Kruskal–Wallis test and evaluated using 95% bootstrap confidence intervals, while coding results were analyzed solely using 95% bootstrap confidence intervals.

The analysis showed notable differences in construct coverage across tests. In terms of readability, EIKEN Pre-1 was the most difficult, while the Common Test was the easiest. A clear contrast in vocabulary was also observed: EIKEN Pre-1 and TEAP emphasized academic vocabulary (NAWL), whereas EIKEN Grade 2 focused on basic vocabulary (NGSL 1k). Coding results revealed additional differences. EIKEN Pre-1 primarily functioned as a vocabulary test (51.2% of items), whereas the Common Test emphasized information synthesis across extended texts. TEAP and the NCT focused on local-level reading, with over 40% of items targeting narrowly defined segments.

These findings highlight critical differences in test design, with critical implications for university admissions in Japan. Universities using external tests—particularly when applying conversion scales such as the EIKEN Common Scale for English (CSE) score—should recognize that EIKEN Grades 2 and Pre-1 differ not only from each other but also fundamentally from the Common Test in terms of the reading constructs they assess. Moreover, the findings suggest that the perceived difficulty of the Common Test, compared to the NCT, derives not from textual complexity but from its emphasis on synthesizing information across multiple segments and modalities. These insights can guide alignment between test constructs, admissions policies, and high school instruction, promoting more appropriate test uses for university admission in Japan.

Keywords

test content analysis, external tests, university entrance exams in Japan, university admission tests in Japan

日本人 EFL 学習者のための語彙サイズテスト VST-NJ8 の妥当性検証：
「わからない」選択肢の追加と回答時間の関係性

小室竜也 (東北大学)
及川凱亜 (東北大学)

本研究は New JACET List of 8,000 Basic Words に基づいて作成された VST-NJ8 (Hamada et al., 2021) について、2つの実験を通してその妥当性を検討した。実験 1 では I don't know 選択肢を追加した際の正答率および回答時間の関係性を検討し、実験 2 では I don't know を含まないコンピュータ適応版 (CAT) で出力される受験者能力との関係を検討した。

VST-NJ8 は 8 つの頻度レベルから英単語が 20 問ずつ提示され、正しい日本語訳を I don't know を含まない 4 つの選択肢から選ぶ形式の語彙サイズテストである。Hamada et al. (2021) では 3 パラメタロジスティックモデル (3PLM) の IRT を用いて、推測パラメタも含めたモデリングを行っている。多肢選択式の語彙サイズテストは、当て推量によって結果が変動してしまうため I don't know を加えて信頼性を高めることが検討されてきた (Stoeckel et al., 2016; Zhang, 2013)。しかし 3PLM に基づく語彙サイズテストでは、推測パラメタが項目作成の段階でモデリングに含まれているため、I don't know を含めることが正答率および回答時間にどのような影響を与えるのかは明らかではない。

そこで実験 1 では、I don't know を含む VST-NJ8 と含まない VST-NJ8 を用いて、81 名 (I don't know 選択肢あり群 41 名、なし群 40 名) の日本語を母語とする大学生の回答を比較した。テストはコンピュータ上で実施され、回答に要した時間 (RT) も収集された。その結果、弁別力が高い項目ほど、I don't know 選択肢があることで正解率が高く推定された ($OR = 1.25$, 95% Credible Interval = [1.04, 1.51])。一方、RT において I don't know 選択肢の有無による差は見られなかった ($\beta = 1.09$, 95% Credible Interval = [0.86, 1.38])。

実験 2 では、Hamada et al. (2021) で得られているパラメタに基づいて作成したコンピュータ適応型テスト (CAT) による受験者能力の出力 ($n = 82$) と、実験 1 で推定された受験者能力の関係を検討した。その結果、CAT の θ は紙とペンを使ったテストの θ よりも有意に高かった ($t = 6.97$, $p < .001$, $d = 1.09$)。これらの結果から、VST-NJ8 は紙で実施した場合には I don't know 選択肢の有無は正答率にほとんど影響しないが、I don't know を含まない CAT 版では受験者能力を過剰に高く推定する傾向が見られた。

Keywords

語彙サイズテスト, I don't know 選択肢, コンピュータ適応型テスト

Exploring the Evolution of a Novice Instructor's Assessment Literacy in Higher Education

Michelle JOHNSON (Nihon University College of Industrial Technology)

Assessment literacy is widely acknowledged as an important part of teaching well in higher education. However, there are persistent challenges in defining the concept clearly, accurately evaluating instructors' assessment proficiency, and supporting the actual use of new assessment ideas in practice (Gaikwad et al., 2023; Pastore, 2022; Medland, 2018). While researchers have proposed various frameworks, such as Pastore and Andrade's (2019) three-dimensional model of assessment literacy and Clarke and Hollingsworth's (2002) interconnected model of professional growth, more research is needed that connects theory to what teachers actually do, especially among novice instructors in diverse educational contexts.

This qualitative case study investigated the development of assessment literacy in a novice EFL university instructor in Japan over an extended period. The study focused on the instructor's use of both formative and summative classroom assessment and traced how her beliefs, knowledge, and assessment practices evolved across multiple semesters. Using classroom observations, interviews, and archival materials such as lesson plans and assessment artifacts, the study explored how and why the instructor made specific assessment decisions and how those decisions reflected shifts in her professional thinking and understanding of effective assessment.

The conceptual framework for the study combined Pastore and Andrade's (2019) dimensions of assessment literacy—knowledge, practices, and principles—with Clarke and Hollingsworth's (2002) model of professional growth, which examines the interplay between a teacher's beliefs, practice, and reflection. By analyzing the instructor's reflective comments, teaching artifacts, and classroom behaviors, the study identified patterns of change and professional learning in relation to classroom assessment.

The findings revealed that the instructor's assessment literacy developed through repeating processes of trying things out and thinking them through. Influenced by institutional context, student feedback, and growing familiarity with assessment principles, her assessment practices shifted from being task-oriented to more learning-oriented. This growth was accompanied by increased confidence, clearer alignment between learning objectives and assessment tasks, and a more nuanced understanding of student learning processes.

The study offers implications for the design of assessment literacy professional development programs in higher education, particularly those aimed at supporting early-career university instructors. While limited to a single case, some findings may be transferable to similar teaching contexts and contribute to the broader conversation about realistic, long-term ways to help teachers grow in assessment literacy, depending on their teaching context.

Keywords

Assessment literacy, Foreign language assessment literacy, Higher education

Investigating the predicted washback of an English speaking exam used for university admissions in Japan

David ALLEN (Ochanomizu University)

The use of English exams for admission into schools and universities in Japan, and the washback of these exams, has received increasing attention (Allen & Tahara, 2021). One of the key points of discussion is the potential washback of speaking exams on education in mainstream schools.

This presentation describes research into the use of a speaking test for university admissions. The goals of the study were 1) to determine the expected washback of using the test if it were used broadly across Japan, and 2) to produce a theory of action for the use of the test in this context (Chalhoub-Deville & O'Sullivan, 2020; Saville & Khalifa, 2016).

The study focuses on the BCT-S, which is a tablet-based speaking test developed by the British Council and Tokyo University of Foreign Studies, where it is currently used as part of the general admissions process. The BCT-S involves a variety of tasks, including responding to personal questions, picture description, and an extended monologue.

Nine senior high school teachers from different regions of Japan were recruited for the study. First, each participant was familiarized with the test by reading documentation, watching videos, and taking a sample test. Then, they completed a background survey and took part in a 90-minute semi-structured interview.

After eliciting information about teachers' current educational micro-context, the interviewer invoked a hypothetical situation in which the BCT-S was to be used as part of the Common Test of English. Teachers were asked to imagine how, if at all, teaching English at their school would change due to the introduction of the test. Interview questions targeted perceived washback effects and mediating factors.

The findings revealed a range of washback effects that varied by teaching context. The overarching mediating factors were the perceived goals of English education and stakeholder expectations. These determined the teaching approach in each school and consequently the extent to which current teaching was perceived to be aligned with the test, and the likelihood that teachers could cope with the test demands.

In particular, teachers at prestigious high schools that adopted 'traditional' teaching methods expected intense washback. Conversely, teachers at schools where English speaking was already a focus of instruction expected only minimal washback.

Based on these findings, a theory of action was produced to support positive washback. The specific actions listed therein include potential modifications to the test, and the provision of additional training, support and information to teachers.

Keywords

washback, speaking assessment, entrance exams, university admissions

How to Motivate Development of Tests of Interactive Competence

F. Scott WALTERS (Seoul National University of Science and Technology)

As Huth (2021) observes, the testing of interactional competence (or IC; Galaczi and Taylor, 2018; Hall, 2018; Pekarek Doehler, 2018, 2021) has been limited by conceptions of language and language proficiency, such as those embodied in the standards of the American Council for the Teaching of Foreign Languages (ACTFL Proficiency Standards, 2012) and those of the Common European Framework of Reference for Languages (Council of Europe, 2018), which are largely at variance with the findings of conversation analysis (CA; e.g., Sacks, Schegloff, and Jefferson, 1974; Heritage, 1984). CA researchers since the 1970s have analyzed a large body of video- and audio-taped conversational data, in both non-test and test settings, which explicates a wide variety of conversational practices and co-constructed language-sequences (e.g., Koshik, 2002; Bernsten, 2002; Schegloff, 2007; Tominaga, 2013; Okada and Greer, 2013; Kunitz, 2021). However, since these practices and sequences are often considered irrelevant or opaque to classroom teachers trained to view a second language (L2) in “ACTFL-like,” individualistic terms, such empirical language phenomena may become largely invisible to traditional assessments of L2 pragmatics. Such invisibility therefore threatens those protocols’ validity-of-use (Kane, 2006; Chapelle and Lee, 2022).

To address this test-development problem, this study examines a dataset ($n = 136$) of audio- and video-recorded tester-examinee interactions collected from a series of CA-informed tests (CAITs) of L2 English pragmatics, with two research questions in mind: (1) How might the raw data from CAIT interactions be separated into meaningful High, Middle, and Low proficiency levels? (2) How might content-related evidence for validity (Messick, 1989) be ascertained from such an arrangement?

The following methodology was used: Interactions were first finely transcribed from the recordings according to CA conventions (Antaki, 2017). Next, participants’ responses were compared with native-speaker norms as described in CA research literature on assessment responses (Pomerantz, 1984), compliment responses (Pomerantz 1978), and pre-sequence responses (Schegloff, 2007) to isolate tentative proficiency levels. Third, Kane’s (2011) validity inference-scheme was applied to the data, with its focus on Data, Claim, Warrant, Backing, and Exceptions, and an interpretive argument for validity-of-use was then drafted. Finally, the resultant tentative operationalized norm was interpreted with reference to principles enunciated by Stivers (2015), Schegloff (1993), and Bachman and Palmer (1996 [2010]). It is here argued that content-related evidence for validity of such CAITs can be adduced via the above process and can be used to motivate the construction of useful tests of IC.

Keywords

conversation analysis, language norms, rating scale development, validity

「CEFR-J×28 多言語 Can Do テスト」 (リーディング) の開発と課題

梅谷 博之 (武蔵野大学)

根岸 雅史 (東京外国語大学)

周 育佳 (東京外国語大学)

ミラー マシュー (東京外国語大学)

川本 夢子 (東京外国語大学)

投野 由紀夫 (東京外国語大学)

本発表では、東京外国語大学で進められている「CEFR-J×28 多言語 Can Do テスト」のうち、リーディングテストの開発過程に焦点を当て、そこで明らかとなった課題と、今後に向けた改善策について報告する。

東京外国語大学では、外国語教育プログラムの改革にあたり、CEFR-J を英語以外の専攻言語にも適用し、専攻言語間の学習成果を可視化・比較可能とする言語共通の評価枠組みを構築している。その中で、学習者が各レベルで「何ができるか」を測定する Can Do テストを開発している。本プロジェクトでは、すでに開発されていた英語版テストをモデルに、2021 年度より各言語への翻訳を進めた。さらに、英語版のテストスペックを基に、作問者向けの説明会を開催し、各言語において独自のリーディングテストを作成した。2023 年度末には、27 言語においてリーディングテストを実施した。

テスト実施後の学生アンケート調査では、リーディングテストの結果として自身の CEFR-J レベルがフィードバックされることが学習の動機づけにつながり、今後の学習に役立つという回答が 9 割近くあった。一方で、テスト作成における課題として、作問者の行動指向アプローチに対する理解不足、Can Do ディスクリプターに忠実に作問することの困難さ、言語の構造や文化的背景の違いによる、多言語間で共通のスペックを適用することの難しさが挙げられる。

今後は、言語の違いに配慮したテストスペックの改善、分析結果に基づく作問者への効果的なフィードバック体制の構築、また作問および運用プロセスの効率化が求められる。さらに、体系的な作問支援の整備や、全学的な実施体制の確立も重要な課題である。本発表では、これらの課題と現時点での対応策、そして今後の展望について考察を行う。

Keywords

CEFR-J、多言語対応、Can Do テスト、リーディングテスト開発

Refining the Design of CEFR-J English Listening Tests through Empirical Data and Student Perspectives

Matthew MILLER (Tokyo University of Foreign Studies)

Masashi NEGISHI (Tokyo University of Foreign Studies)

Yujia ZHOU (Tokyo University of Foreign Studies)

Yukio TONO (Tokyo University of Foreign Studies)

James CARPENTER (Tsurumi University)

The CEFR-J×28 Project, initiated by Professor Tono and his collaborators at Tokyo University of Foreign Studies (TUFS), seeks to enhance the assessment of English proficiency among Japanese learners by aligning testing practices with the CEFR-J framework—Japan’s localized version of the Common European Framework of Reference for Languages (CEFR) for ELT. The CEFR-J provides more detailed proficiency levels, making it well-suited for both diagnostic purposes and curriculum development. While previous project cycles focused on CEFR-J English Reading tests, the 2024 initiative shifted its attention to Listening proficiency, an area underexplored in standardized testing for advanced learners in Japan.

This study aims to (1) evaluate the effectiveness of the CEFR-J English Listening Test, (2) explore how different item contextualizations influence learners’ listening test-taking strategies, and (3) inquire about learners’ real-life listening contexts and their perceptions of context necessity in listening tests. The ultimate goal is to refine the test’s design to ensure test quality and real-world relevance, thereby supporting more accurate proficiency assessment and greater learner engagement. This year’s project included the administration and analysis of CEFR-J Listening tests (the English test targeted levels B2.1 to C2). In addition, qualitative feedback from test-takers’ questionnaires was collected and examined. Three item formats with different degrees of contextual embedding were created and administered to eleven English-major students. Post-test interviews were conducted to gain deeper insights into their test experiences and everyday listening contexts.

Analysis of the Listening test results highlighted encouraging findings for B-levels (B2.1 to B2.2), with items showing appropriate difficulty and strong discrimination. However, C-level items were not challenging enough, indicating a misalignment needing revision. Test-taker comments in the post-test questionnaire provided valuable feedback on the quality and relevance of contexts.

In the experiment, unlike previous findings with Reading, varied contexts in Listening items did not alter students’ test-taking strategies. A likely explanation is the one-minute preparation time before the audio began, allowing test takers ample time to preview items regardless of the type of context. Interviews revealed a general consensus that while the items allowed students to demonstrate their listening abilities, improvements could be made in authenticity and clarity of the contexts and audio delivery. Interview responses also illustrated how and why students engage with English listening in daily life, suggesting a need to align test content more closely with real-world usage.

From these findings, we propose key revisions: recalibrating C-level items, diversifying and clarifying listening contexts, modifying audio scripts and recordings, and enhancing communication about the test’s purpose. These revisions would help improve test validity and encourage test-takers’ motivation—crucial for the success of language assessment.

Keywords

CEFR-J, Listening Tests, Test Design, Contextualization, EFL Learners

Dialogic Critical Thinking English Test – A New Frontier in AI-Powered Language Assessment

Stuart CONNOR (Pearson)

This presentation introduces a groundbreaking prototype in English language assessment: the Dialogic Critical Thinking English Test (DCTET), an AI-powered, voice-interactive module designed to evaluate not just language proficiency but also critical thinking in real-time, dialogic contexts. Developed as an add-on module to the Versant English Speaking & Listening Test, DCTET leverages agentic AI to simulate authentic customer interactions and assess how well candidates think, reason, and communicate under pressure.

The session will begin by contextualizing the need for such an assessment in today's world, where language users must navigate increasingly complex, emotionally nuanced, and cognitively demanding conversations that cannot easily be replaced by AI. Uniquely, the test captures both language and reasoning performance domains—ranging from clarity and fairness to inference and self-regulation—through an immersive, interactive “ambient” assessment experience.

The presentation will introduce the test's architecture, which includes a voice user interface powered by a low-latency speech language model that modulates emotional tone and speech rhythms in real time. The test flow includes a pre-briefing with an AI instructor, a five-minute role-play with a simulated customer, and an automated scoring phase. The scoring engine uses a multi-agent workflow: each agent evaluates a specific criterion, generates evidence and rationale, and contributes to a visualized score report within 30 seconds.

Key features include scenario configurability (e.g., customer service, academic advising), adjustable difficulty (e.g., CEFR-aligned speech complexity), and real-time feedback. Unlike traditional speaking tests that rely on monologues or short responses, DCTET captures the richness of real conversation—turn-taking, repair strategies, and emotional shifts—making it uniquely suited for enterprise and academic applications.

The presentation will also share results from internal and external pilots, highlighting score distribution patterns and inter-agent reliability. Early feedback from user testing described the experience as “absolutely natural and realistic,” underscoring the test's potential to redefine how we assess decision-quality in spoken English while also building speaking confidence through practical application of skills in relevant, hyperreal simulations.

Attendees will leave with a clear understanding of how dialogic, AI-mediated assessment can bridge the gap between language accuracy and critical thinking—an essential leap for workforce readiness and academic success in the age of conversational AI.

Keywords

Conversational AI, Agentic AI, AI scored testing

Evaluating the Efficacy of AI Driven Approach to the Development and Revision of a Placement Test

Gordon WILSON (Hannan University)
Myles GROGAN (Osaka Ohtani University)

Creating tests is time-consuming. Finding the skill, resources, and time to develop or maintain a test may be challenging in some contexts. However, easily available artificial intelligence (AI) tools like ChatGPT allow institutions to re-examine this situation in terms of development of instruments and the assessment of results. This assessment practice report shows the describes how one department in the developed, trialed, and implemented a placement instrument with the aid of AI, showing the approach and results of the process.

The department had been using a 50-item multiple choice placement instrument developed internally many years previously. Roughly 300 students annually needed to be divided among seven classes. Teachers' impressions were that the test seemed to work reasonably well, but little consideration had been given to statistical performance indicators. An investigation of the most recent sitting showed Cronbach's Alpha of 0.80, SD = 6.8, and the 20 items returning less than 0.3 on item discrimination. This performance showed the need for a new instrument to place students.

After considering the curriculum and several commonly available tests, seven sections were developed for the new instrument (four listening sections and three reading sections). Each section had a prompt used to generate questions within AI. A pilot test was run at the end of the academic year with 256 students at all levels taught in the department. This produced a stronger result ($\alpha = 0.842$, SD = 7.58, ID<0.3 = 13). Items were identified for modification before the implementation of a final version for use as a placement instrument. Following such modifications, the results were $\alpha = 0.870$, SD = 8.72, ID<0.3 = 13).

The role of AI in creating this test was pivotal. Prompts were used to create multiple items quickly. Because AI is not perfect, several items were rejected, but this led to a clarification of objectives for each of the sections. This reflects ideas of test specification or blueprints. Finally, the review process created a modest interest in assessment literacy among several parties at the university. This led to several useful observations relating to test practice, such as booklet size or ways to make sure students answered all questions.

AI has therefore proved both directly and indirectly beneficial to test creation in this department. This success leads presenters to believe that AI can be effective in creating local tests tailored to specific department needs and student demographics while enhancing teachers' assessment literacy.

Keywords

AI, Placement, Revision

Perceptions of an AI-scored classroom-based speaking test: Comparisons between Japanese in-service and pre-service teachers

Rie KOIZUMI (University of Tsukuba)
Yo IN'NAMI (Chuo University)

This study examines how in-service and pre-service teachers perceive an AI-scored classroom-based speaking test to identify its promises and challenges for formative and summative assessments. Technological advances in AI have been used to administer and score second language (L2) English-speaking tests, and provide detailed feedback to students and teachers in Japanese classrooms (e.g., ELSA, PROGOS, and AISATS). While teacher assistance to students may be needed for effective assessment, an AI-based speaking assessment could improve formative and summative assessments in the classroom. However, the effectiveness of such assessments may vary depending on the teachers' and students' perceptions of such technology. According to the technology acceptance model (Davis, 1989), whether people use and are ready to use a certain technology is affected by their perceptions of (a) its usefulness for their learning and (b) the ease of using the technology (e.g., Liu et al., 2024). Drawing on this model, this study examines (a) perceived usefulness and (b) perceived ease of use from the perspectives of in-service and pre-service teachers. Although similar studies exist in other countries, to the best of our knowledge, no such studies have been conducted in Japan.

The participants were 20 Japanese senior high school English teachers and 17 Japanese undergraduate and graduate students at Japanese universities. The latter group either pursued teaching certificates or were enrolled in L2 English language education programs. They offered perspectives as both learners and future educators. All participants took AISATS (formerly Speechace; Potential Plus Co., Ltd., 2025) by responding to nine English questions online. After receiving a report containing test scores and other details, the participants answered an online questionnaire for this study.

The results show that over 90% of the participants found the test easy to take. Additionally, 80% of in-service and 94% of pre-service teachers considered the test scores useful for grading when tasks are teacher-developed and practiced in class. Moreover, 88% of pre-service teachers preferred the use of such scores for grading. While both groups perceived AISATS as simple and useful, some expressed concerns about its application for grading. These included potential technological issues with speech recognition and automated scoring, classroom noise and WiFi stability, cost, interpretability of scoring rubrics, and availability of technical support from the test provider. Additionally, some emphasized the importance of ensuring test scores that are reasonably valid, reliable, and fair across classes, and the accountability of those scores to students and parents. Addressing these concerns would help teachers and students use AI for summative assessments. Future studies should incorporate the perspectives of additional stakeholders, such as parents and students, regarding the use of AI.

Keywords

formative assessment, summative assessment, learning-oriented assessment, AI technology

Commercial tests or AI-assisted streaming?: A comparison at the university level

Myles GROGAN (Osaka Ohtani University)

AI has given institutions such as universities new choices with regard to test creation and implementation. With AI, placement tests can be created using prompts for each section, and successive iterations of tests can be tailored to a specific population using the feedback from each item—an option that may not be possible with commercial tests. This assessment practice report explores the presenter's experiences in trying to create a flexible and cost-effective assessment using AI, while also aiming to improve diagnostic precision for incoming students.

In autumn 2024, a seven-section test was created at the presenter's institution, using AI prompts. The first pilot test used 30 paid student volunteers from each of the faculties and courses at the university. The test items worked well, and reliability was high. The 2025 cohort of students was tested with the usual commercial test, but the AI-created test was also administered to 144 participants from five classes. The newly created test proved internally reliable (Cronbach's $\alpha = 0.89$), and it showed a reasonable correlation with the commercial test (Pearson $r = 0.88$).

Several problems remained. While about 76% of students scored above the level of chance, the test showed a left-skewed distribution; a flatter distribution would be preferred to reduce error in placement. In addition, teachers reported that the test was too difficult when compared with the previous test. The absence of illustrated items was considered problematic, particularly given the level of the students. Finally, the audio, which had been made using text-to-speech technology, was considered too fast.

The experience of creating the test highlighted the need for the “soft-skills” of assessment literacy, such as getting buy-in from other teachers and stakeholders. Administrative costs for the test also need to be considered. While these are cheaper than commercial tests, they are not free in financial terms or in terms of administrative time.

Although this report shows a successful case, it comes with some caveats. Some institutions may still consider commercial tests to be a better fit. While AI offers scalable item creation, successful implementation still depends on professional expertise, stakeholder engagement, and resource planning. However, for the presenter's university, the arrival of AI and its use in test creation show great promise in terms of increasing assessment literacy and a continued process for focused feedback for all stakeholders.

Keywords

Artificial Intelligence, Placement, Commercial testing

日本手話到達基準「JSL スタANDARD」におけるコミュニケーション言語活動の策定

中野聡子 (群馬大学)
浅田裕子 (昭和女子大学)
藤田元 (上智大学)

本研究は、中野・浅田(2023)、中野・浅田・藤田(2024)が開発を進めてきた日本手話到達基準「JSL スタANDARD」を、コミュニケーション言語活動(以下、活動と省略する)に拡大するものである。

日本手話(Japanese Sign Language: JSL)は、自然言語の1つであり、日本語とは異なる言語体系を有している。しかし、L2日本手話・手話通訳教育においては、学術的包括性、明確性、一貫性を兼ね備えたスタンダードが確立されていない。そこで著者らはL2日本手話教育の質の向上を目的として、CEFRに準拠した到達指標であるJSLスタンダードの開発に取り組んでいる。

本発表は、JSLスタンダードで能力記述文を必要とする活動とは何であるかを提案し、同提案にかかる背景や経緯を明らかにするものである。本研究では、手話教育における活動にはどのようなものが含まれるべきであるかという観点から、CEFR-CV(Council of Europe 2020)で提示されている活動の内容を整理することを試みた。CEFR-CVでは、活動が、受容・産出・やりとり・仲介という4つの観点で下位分類されている。さらに、4つの観点のそれぞれについて、具体的な活動や方略が設定され、それぞれに能力記述文が設定されている。

JSLスタンダードも、上記の4つの分類を導入することとした。一方で、CEFR-CVの個別の活動を、JSLにそのまま適用できないことも明らかとなった。この点に関する先進的な事例として、チェコ共和国におけるFramework of Reference for Sign Languages(FRSL)及びReference Level Descriptors for Czech Sign Language(RLDCSL)がある。これらは、CEFR-CVとは異なるが、学習者を含む手話使用者にとっては有用な具体的活動を設定している。

JSLスタンダードは、CEFR-CVを尊重しつつ、能力記述文が設定される個別具体的な活動については、FRSL・RLDCSLも取り入れて、どの活動について能力記述文を設定すべきかを決定した。具体的には、FRSL・RLDCSLにおける活動を基本として、そこに設定されている活動をCEFR-CVと対応付けることとした。その結果、受容的活動については6つ、産出的活動については5つ、やりとりにかかる活動については6つの具体的な活動が設定された。

Keywords

日本手話, CEFR, コミュニケーション言語活動

Workshop Information (ワークショップ情報)

題目：「英語教育における生成 AI の活用～授業準備からフィードバックまで～」
(日本語で実施)

講師 高木 俊輔 (聖光学院中学校高等学校)
司会 平井 明代 (筑波大学)

日時： 2025 年 9 月 14 日 (日) 9 時 00 分～12 時 00 分 (休憩を含む)

場所： 群馬大学荒牧キャンパス (2 号館 GB154 教室)

参加費： 会員...無料、非会員...1,000 円

※JLTA の会員資格は 9/3 (水) までに学会費の納入があったかどうかで判断します。

定員： 50 名 (申し込み順)

申込期限： 9 月 3 日 (水) [定員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環境を前提とします。(c)～(e)について、各 AI サービスのアカウントをお持ちでない場合は、事前に無料アカウントをご作成ください。(f)についてはアカウントのご準備は不要です。

(a) PC 端末 (OS は問いません。キーボードのあるものをご持参ください。)

(b) ウェブブラウザ (Google Chrome を推奨。Safari、Microsoft Edge でも可。)

(c) ChatGPT (<https://chatgpt.com/>) の無料アカウント

(d) Claude AI (<https://claude.ai/>) の無料アカウント

(e) ElevenLabs (<https://elevenlabs.io/ja>) の無料アカウント ※音声生成 AI

(f) Google Gemini (参加者向けに当日利用可能なデモアカウントをご用意します。)

目的：

1. 生成 AI を活用し、授業の素材を作成するためのスキルを獲得する
2. 英語教育における生成 AI の効果的な活用に向けて、必要な知識と考え方を習得する

手順：

1. 講義 1: 生成 AI の活用ガイドラインと、活用に向けた基本的な考え方
2. ハンズオンワークショップ 1: プロンプトの書き方と生成 AI を活用した教材作成
3. 講義 2: 生成 AI を織り込んだ授業デザインを考える
4. ハンズオンワークショップ 2: シナリオに基づいたプロンプトの調整
5. グループディスカッション: 各校での活用に向けたブレインストーミング

申し込み方法：

1. 下記 URL または QR コードより申し込みフォームにアクセスし、必要な情報をご登録ください。

URL: <https://forms.gle/BmVCyUBharbc5b7K7>

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。
下記の情報を久保佑輔 (山梨大学) [kubo\(at\)yamanashi.ac.jp](mailto:kubo(at)yamanashi.ac.jp) まで e-mail でご連絡ください。

(1) 氏名・所属・メールアドレス（全員）

※Google Gemini のデモアカウントをお送りするため、講師の高木先生にお名前とメールアドレスが共有されます。デモアカウントは不要の場合は、その旨をメールにお書きください。

(2) ワークショップの参加理由（全員）

(3) 講師からの質問（③は任意）

①以下の各 AI について、使用頻度を教えてください。「全く使ったことがない」「誰かが使っているのを見たことがある」「使ったことがある」「日常的に使っている」の中から 1 つずつお選びください。

ChatGPT、Claude AI、ElevenLabs、Google Gemini

②「授業で AI を活用する」と言われた際、授業（50 分）における使用時間はどれくらいだと想像しますか？「10 分以内」「10～20 分」「20～30 分」「30～40 分」「40～50 分」の中から 1 つお選びください。

③生成 AI を英語学習で活用する際に不安に感じるがあれば教えてください。（自由記述）

(4) 講師へのご質問（希望者のみ）

Workshop Information

Utilizing Generative AI in English Education **— From Lesson Preparation to Feedback—** (Conducted in Japanese)

Lecturer: Shunsuke TAKAGI (Seiko Gakuin Junior and Senior High School)

Chair: Akiyo HIRAI (University of Tsukuba)

Date: Sunday, September 14, 2025, 9:00 – 12:00 (including break)

Venue: Gunma University, Aramaki Campus (Building 2, Room GB154) **Fee:** Free (1,000 yen for non-members)

Maximum number of participants: No limit

Application deadline: September 29th

Prerequisite: Participants are expected to be familiar with basic statistical knowledge such as correlation coefficients and regression analysis. The workshop will include a detailed explanation of machine learning with demonstrations using R. However, basic data handling in R is not required.

Fee: Free (¥1,000 for non-members)

※ Membership status will be determined based on whether the annual fee has been paid by Wednesday, September 3.

Capacity: 50 participants (first come, first served)

Registration Deadline: Wednesday, September 3

※ Registration will close once capacity is reached.

Participation Requirements

This hands-on seminar assumes the following setup. For items (c) to (e), please create a free account in advance if you do not already have one. No account is needed for item (f).

- (a) A PC device (any OS is acceptable; please bring a device with a keyboard)
- (b) A web browser (Google Chrome recommended; Safari and Microsoft Edge are also acceptable)
- (c) A free account with ChatGPT (<https://chatgpt.com/>)
- (d) A free account with Claude AI (<https://claude.ai/>)
- (e) A free account with ElevenLabs (<https://elevenlabs.io/ja>) — voice generation AI
- (f) Google Gemini (demo accounts will be provided for participants on the day)

Objectives

1. To acquire skills for creating teaching materials using generative AI
2. To gain the necessary knowledge and mindset for effectively utilizing generative AI in English education

Agenda

1. **Lecture 1:** Guidelines for utilizing generative AI and key concepts
2. **Hands-on Workshop 1:** Writing prompts and creating teaching materials with generative AI
3. **Lecture 2:** Designing lessons that incorporate generative AI
4. **Hands-on Workshop 2:** Adjusting prompts based on specific scenarios
5. **Group Discussion:** Brainstorming ideas for implementation in individual schools

How to Register

1. Please access the registration form via the URL or QR code below and fill out the required information:

URL: <https://forms.gle/BmVCyUBharbc5b7K7>

2. If you have any issues with the online form, you may register by email. Please send the following information to Yusuke KUBO (University of Yamanashi) via email: y.kubo(at)yamanashi.ac.jp
- (1) **Name, Affiliation, and Email Address** (*required*)
- ※ Your name and email address will be shared with the lecturer, Mr. Takagi, for the purpose of sending the Google Gemini demo account. If you do not need a demo account, please indicate that in your email.
- (2) **Reason for participating in the workshop** (*required*)
- (3) **Questionnaire from the lecturer** (*required except for ③*)
- ① For each of the following AI tools, how frequently do you use them? Please choose one:
Never used / Have seen someone use it / Have used it / Use it regularly
- ChatGPT:
 - Claude AI:
 - ElevenLabs:
 - Google Gemini:
- ② When you hear “using AI in class,” how much time do you imagine it would be used in a 50-minute lesson? Please choose one:
Up to 10 minutes / 10–20 minutes / 20–30 minutes / 30–40 minutes / 40–50 minutes
- ③ If you have any concerns about using generative AI for English learning, please share them. (*Free response*)
- (4) **Any questions for the lecturer** (*optional*)

5. Conference Venue

交通アクセス (Access to the University)

(公共交通機関をご利用の皆さまへ)

※新幹線は前橋駅には停車しません。東京方面からお越しの場合は高崎駅で在来線（両毛線）にお乗り換えください。

※荒牧キャンパスの最寄りとなるバス停は「群馬大学荒牧」です。正門すぐ横（出て左側）にあります。

※土日祝は「群馬大学荒牧」発の 17:20 が最終となります。それ以降は「前橋自動車教習所」発で 20:52 が最終となります。詳細は下記をご参照ください。

◎前橋駅から荒牧キャンパスまで（約 30 分）

JR 前橋駅北口 2 番乗り場 【関越交通バス】 前橋渋川線

<https://kan-etsu.net/pages/24/>

■前橋駅発（渋川駅方面）バス時刻表（土日祝）

<https://kan-etsu.net/files/libs/12299/202405231450492417.pdf>

◎荒牧キャンパスから前橋駅まで（約 30 分）

※土日祝はバス（関越交通バス）の本数が非常に限られております。以下をご確認のうえ、ご乗車ください。最寄りのバス停は「群馬大学荒牧」または「前橋自動車教習所」となります。

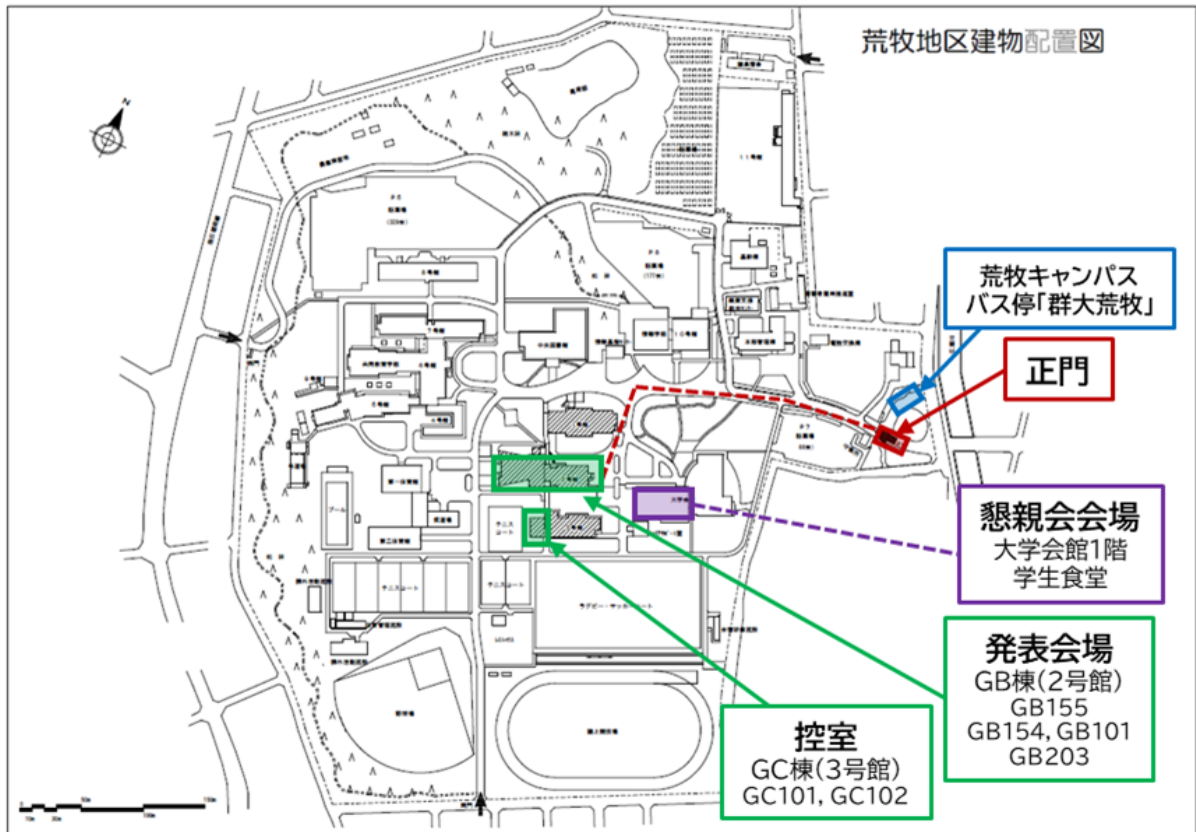
■渋川駅発（前橋駅方面）バス時刻表（土日祝）

<https://kan-etsu.net/files/libs/13409/202507141424138055.pdf>

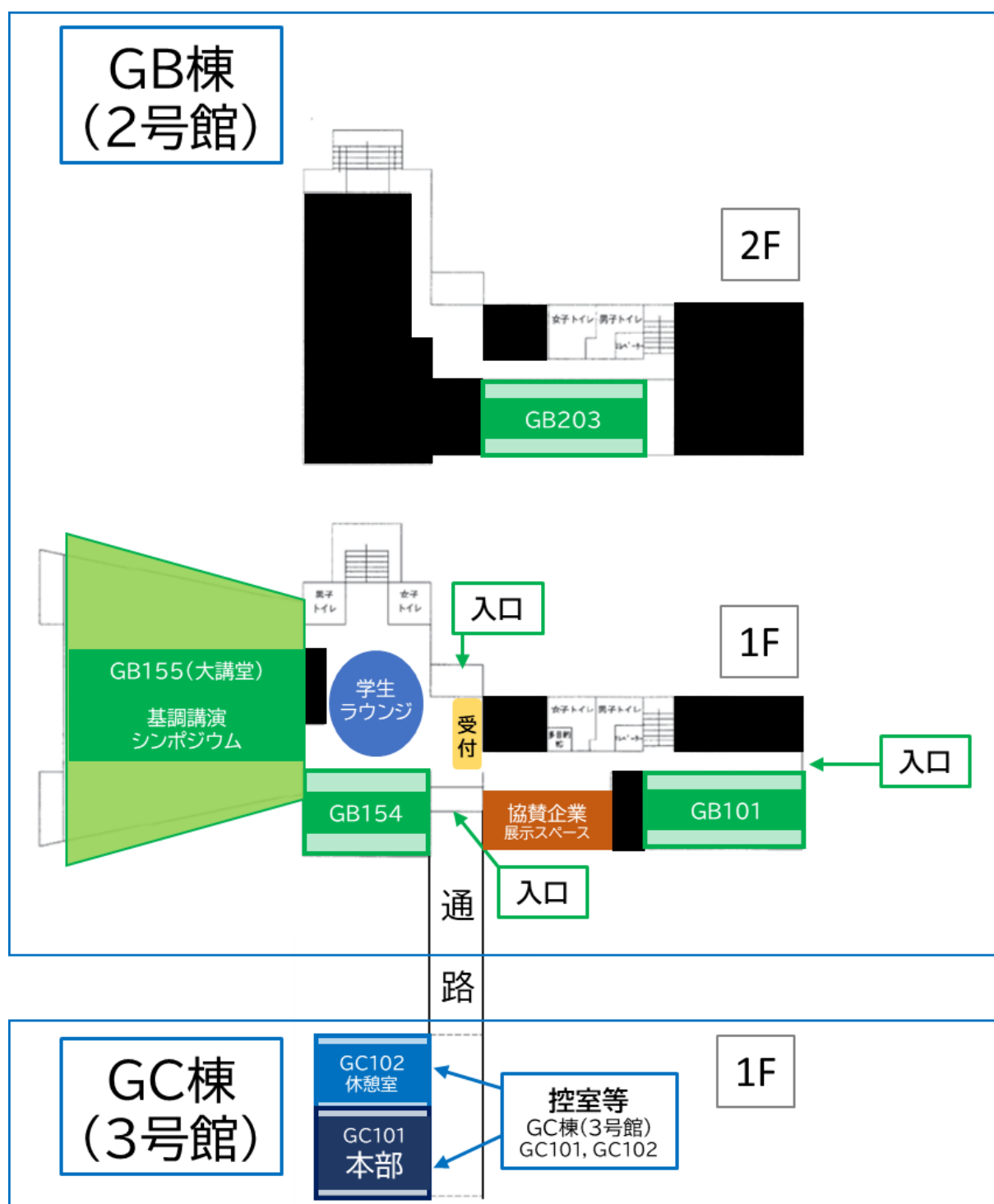
※参考資料：バス停位置確認用



キャンパスマップ (Campus Map)



GB棟（2号館）・GC棟（3号館）フロアマップ
(Floor Map of the GB (#2) and GC (#3) Building)



※ 休憩の際は、GC棟（3号館）102室および学生ラウンジをご利用ください。

TOEIC® Program

英語コミュニケーション能力を公平公正に評価する世界共通の基準

TOEIC® Tests

TOEIC L&R



TOEIC S&W



日常生活やグローバルビジネスにおいて効果的にコミュニケーションを図るうえで欠かせない、英語で「聞く・読む・話す・書く」能力を測るテストです。

TOEIC Bridge® Tests

TOEIC Bridge
L&R



TOEIC Bridge
S&W



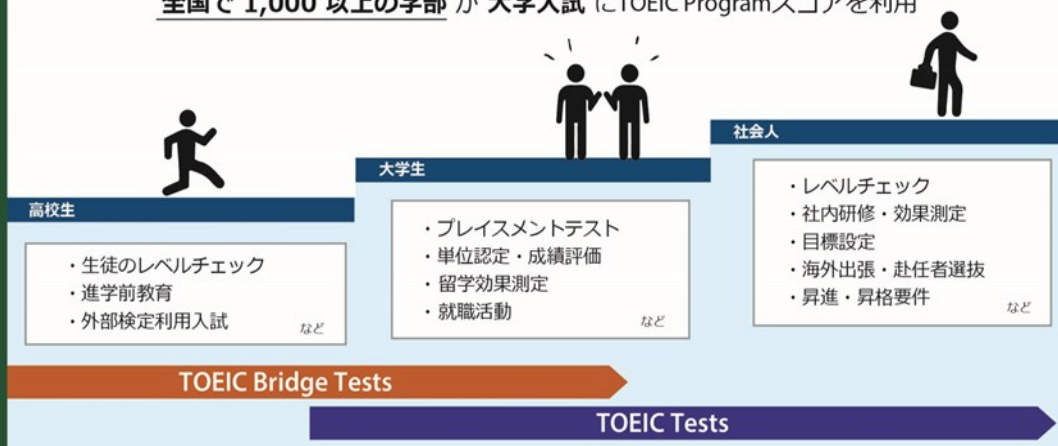
初・中級レベルの基礎的な英語力（聞く・読む・話す・書く）について、日常生活におけるコミュニケーション能力を比較的良好に出会う場面やトピックを用いて測るテストです。

TOEIC Programに関する詳細はこちら



学校でのTOEIC® Program活用例

全国で 1,000 以上の学部 が 大学入試 にTOEIC Programスコアを利用



TOEIC Programに関する各種研究論文

IIBCでは、TOEIC Programと英語教育、英語試験、教科書などに関する調査・研究を行っています。ぜひご覧ください。

各種研究論文はこちら



PICK UP 研究論文の1つをご紹介します！

TOEIC Bridge L&Rで使われる語彙は
中高英語教科書の学習語彙で 約 91% をカバー
中高英語教科書とTOEIC Bridge L&Rの使用語彙の類似性を調査しました。

令和4年度の新課程英語教科書コースおよび
共通テストコースに基づくTOEIC L&R, Bridge L&Rの語彙の特徴
柳城 祐司 (現 日本国際大学教授・筑波大学名誉教授)
小室 竜也 (現 東北大学)



VELC Test[®] [ベルクテスト]

Visualizing English Language Competency Test

大学生のための
英語力診断テスト

オンライン受験可能な VELC Test[®]Online 登場！

詳しくは
ホームページ



テストの
概要

VELC Test[®]PP

受験料金	800円(税抜) / 1人
試験時間	70分(音声CDで時間管理)
試験問題	リスニング・リーディング 各60問 計120問
解答形式	ペーパーテスト(マークシート形式)

VELC Test[®]Online

受験料金	900円(税抜) / 1人
試験時間	通常版: 70分 短縮版: 55分
試験問題	通常版: リスニング・リーディング 各60問 計120問 短縮版: リスニング・リーディング 各45問 計90問
解答形式	オンライン(選択形式)



VELC Test[®]Online の特徴

自宅のパソコンやスマートフォンを使って受験することが可能になりました。年間受験者数が500人以上の場合は、800円(税抜) / 1人に割引します。



迅速な 結果通知

受験結果はweb上のeポートフォリオにてご覧いただけます。最短でOnline版は試験終了日の翌日中、PP版は返却資料到着の翌日中には結果を通知しています。土日・祝日も対応可能です。



可視化された スコアレポート

eポートフォリオで通知されるスコアレポートでは

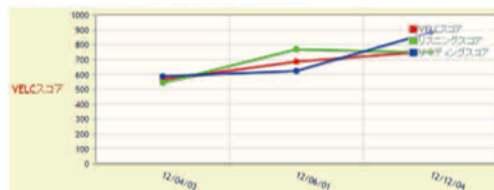
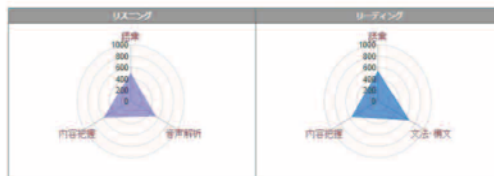
「実際に英語で何ができるかを記載した
10段階で評価される Can Do レベル診断

英語の知識やスキルを細分化し、それぞれの能力を
診断するスキル別正答率と学習アドバイス

VELCスコアから予測されるおおよその目安となる
TOEIC[®]L&R TESTのスコアを表示

など、受験者の英語能力を細部まで可視化しています。

VELC Test [®] スコアレポート		学籍番号: 1002	VELCスコア 564	リスニングスコア 343	リーディングスコア 387
VELCスコア	564	8 8 8 8 8 8 8 8 8 8			
リスニングスコア	343	7 7 7 7 7 7 7 7 7 7			
読解	301	6 6 6 6 6 6 6 6 6 6			
音訳能力	335	7 7 7 7 7 7 7 7 7 7			
内容把握	373	8 8 8 8 8 8 8 8 8 8			
リーディングスコア	387	9 9 9 9 9 9 9 9 9 9			
読解	331	7 7 7 7 7 7 7 7 7 7			
文法・構文	677	10 10 10 10 10 10 10 10 10 10			
内容把握	343	8 8 8 8 8 8 8 8 8 8			



動画でわかる VELC Test
<http://www.velctest.org/movie>

貴学のライティング指導を効率化し、劇的な学習効果を生み出します！

特長

- ①クリック1秒、添削完了！ 提出後、学生自身が即座に振り返り学習
- ②ライティング評価を数値化！ ランク化やスコアリングでらくらく成績管理
- ③アカデミック用途に最適化！ 生成AI特有のハルシネーション等を排除

桐原書店 X PLAY SQUARE

採点時間は約1秒！ 人の手を借りず即時採点、即時添削
英語エッセイライティング自動採点サービス



桐原 AI エッセイライティング

powered by PLAYSQUARE

エッセイ自動採点システム (AEES : Automated Essay Evaluation System) 搭載

学習に喜びと感動を

桐原書店



お問い合わせ・詳細はこちら ▶▶

olbd@kirihara.co.jp



スペリングや文法だけでなく、
書かれた内容の質まで採点！

桐原AIエッセイライティング
(ChatGPT連携版)

評価の理由・よく
できた部分・改善
すべき部分を分析
して具体的にコメ
ントを提示

ChatGPTがより好ましい
解答例を提供

論理構成について
詳細にスコアリング

総合評価	評価理由	改善	文法	表現	構成	採点理由
論理構成に関する改善点						
論理構成スコア	22.5 / 25					
文章構成	1 / 1					
内容の一意性	10 / 10					
エッセイ構成	1 / 1					

評価の理由: エッセイは、読者も自分自身に読み取らなければならない主題と論理に基いており、その構造を重視しています。しかし、内容は一貫性に欠け、読者に読者の興味を惹きつけることができていません。

よくできた部分: 文法は正しいと見えますが、その構造を重視する必要があります。特に、ロジックの構造や単語・文法との関係性を重視する必要があります。

改善すべき部分: 読者の興味は、最初の二つの段落とつながりません。ロジックの構造と論理の関係を重視する一方で、文法と文法を重視するという観点から見て、この点について、より明確な論理構造を構築する必要があります。

■ 中略 (Introduction)

無料体験アカウントを発行できます。上記メールまたはQRコードでお気軽にお問い合わせください。

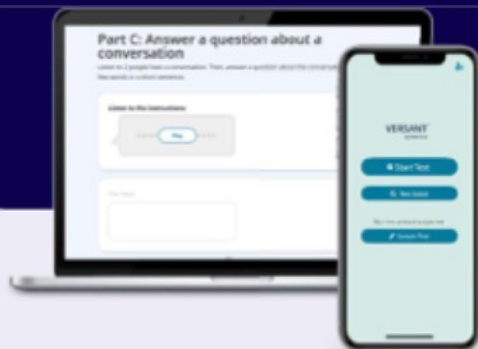
<https://kirihara.sakura.ne.jp/information/aees-ne/>



世界基準の英語能力テスト

Versant by Pearson English Speaking and Listening Test

英語で聞いて理解して瞬時に応答する
実用的なコミュニケーション力を測定



Accurate

最先端のAIを使った言語認識 & 自動採点システムが、客観的で公平に英語力を測定



Fast

パソコンやスマホで簡単に受験ができ、テスト後は、すぐにスコアレポートを開覧可能



Reliable

大学等の教育機関他、世界160ヵ国、2000社以上（国内約500社）が導入！

Versant English Speaking and Listening Test 概要

VERSANT[™]
by Pearson

用途

- 英語コミュニケーション力の測定
- 留学前後の効果測定

測定スキル

- リスニング
- スピーキング（話の内容/話し方）

テスト時間

- 約20分（結果取得 約5分）

対応デバイス

- スマートフォン
- コンピュータ

採点

- Global Scale of English
- CEFR

テスト内容（全40問）

- Questions (8問) 質問に対して簡潔に答える
- Repeat (16問) 復唱する
- Conversations (6問) 会話を聞いて質問に答える
- Passage Comprehension (6問) 物語を聞いて質問に答える
- Story Retellings (2問) ストーリーリテリング
- Open Questions (2問) 自分の意見を話す自由回答

スコアレポート

- ☒ 総合点 (GSE/CEFR)
- ☒ リスニング
- ☒ スピーキング
- ☒ 語法能力
- ☒ 現在の能力
- ☒ レベルアップに向けたアドバイス

リモートモニタリング機能（追加オプション）

受験者の音声と顔を認証する「なりすまし防止」機能で、テスト管理者の負担を軽減！
受験者の映像と音声にアクセスができるので、受験者へのアドバイスにも役立ちます。



先着3校限定！Versant English Speaking and Listening Test 無料トライアル校を募集！

実際のテストを通じて、導入の効果を実感していただけます。ご興味のある方は今すぐお問い合わせを！
無料ライセンス数：各校最大50ライセンス お申込み期限：2025年10月31日

お問い合わせ ピアソン・ジャパン株式会社 カスタマーサービス Email: elt.jp@pearson.com

＜協賛企業・法人・団体の一覧（50 音順）＞

(財) 日本英語検定協会

ETS Japan 合同会社

一般財団法人国際ビジネスコミュニケーション協会 *

英語能力測定・評価研究会（VELC 研究会）事務局 *

株式会社エル・インターフェース

株式会社教育測定研究所

株式会社桐原書店 *

国際交流基金・日本語試験センター

特定非営利活動法人 英語運用能力評価協会

ピアソン・ジャパン株式会社 *

British Council

Cambridge University Press & Assessment

We are deeply grateful to the following companies, corporations, and organizations for their support of our association.

Four companies, corporations, and organizations have Commercial exhibits at the JLTA 28th Annual Conference. We would like to express our sincere gratitude for their contribution to the conference.

当学会は以上の企業・法人・団体様にご支援を頂いております。
第 28 回全国研究大会においては、12 の企業・法人・団体様に展示協賛をいただいております。本大会へのご貢献に厚く御礼申し上げます。

Commercial Exhibits（展示協賛企業） *

Exhibits are located in the Community Room
on the 1st floor in the Building 2 (GC).

展示は、2 号館 1 階コミュニティルームで実施いたしております。

日本言語テスト学会 (JLTA) (2025 年度) 全国研究大会発表要綱
Handbook of the 28th Conference of the Japan Language Testing Association

発行日：2025 年 8 月 26 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：〒960-8516 福島県福島市栄町 10-6

久保田恵佑研究室

TEL: 024-581-5533 (直通)

編集：日本言語テスト学会 全国研究大会実行委員会
