

明治大学

国際日本学研究所

Global Japanese Studies Review

Meiji University

第14巻 第1号  
Vol. 14 No. 1

白戸教授 退任記念号

Special Edition in Honor of Professor Shirato

[退任者の履歴業績、論文・所感]

献呈の辞	鈴木 賢志	1
白戸 伸一 教授 近影, 履歴業績, 所感		5

[研究論文 / Articles]

Developing a Japanese-English Parallel Corpus of “Japan, the Beautiful and Myself” by Yasunari Kawabata	OYA, Masanori	13
Integrating Complex Variables in the Measurement of L2 Speech Production: Focusing on Complexity, Accuracy, Fluency, and Creativity	OZEKI, Naoko/ WANG, Wei-Tung/ OTSUKA, Kiyotaka	27
日本語学習者の作文における自動評価と教師評価の比較	小森 和子 / 伊集院 郁子 / 李 在鎬	41
まちづくり政策の現状と課題—川越市の事例を中心に—	白戸 伸一	69
英語スピーキング授業に対する学習者の評価と個人差要因の関連—対面授業とオンライン授業の比較を通じて—	廣森 友人 / 金田 慎吾 / 新井 智大 / 大石 有紗 / 大谷 温理 / 柴田 恒	89
日本語上級レベルの留学生の論証型レポートに対する自己評価・ピア評価の傾向	安高 紀子	119
Cognition or Context: How Do Teachers Approach Japanese University Speaking Classes?	GARSDIE, Paul	131

[研究ノート / Research Note]

デジタルアーカイブ・ポータル構築・運営・展開の日米欧比較—利活用における「ジャパンサーチ」の特徴と展望—	田中 絵麻 / 大江 宏子	145
上級日本語科目改革のための調査に関する報告	安高 紀子 / 黄 叢叢 / 黄 秀智 / 吴 梅	173

[資料紹介 / Reference Material]

クリエイタービジネス論における音楽視聴調査 (2018 ~ 2021)	原田 悦志	191
-------------------------------------	-------	-----

## 献 呈 の 辞

白戸伸一先生は、国際日本学部の黎明期より現在に至るまで、本学部の発展において多大なる貢献をされました。

白戸先生は、1974年に早稲田大学教育学部をご卒業後、同大学院にて商学修士号を取得され、1985年に明治大学商学研究科における単位修得の後、沖縄大学、浦和大学に勤務され、2002年には商学博士号を取得されました。

そして2008年に本学部の開設と同時に着任され、以後は日本流通史の講義を中心として、日本の流通システムに関する幅広い教育・研究活動に従事されました。またその精力的な活動の傍らで、初代の筆頭教務主任として5年間、第2代の学部長として1年間、さらに2年間のイギリスにおける在外研究から戻られてからの4年間を学科長として、本学部の発展に、多大なる貢献をされました。私を含め、学部創設時からのメンバーにとっては、白戸先生がいらっしゃる執行部は考えられないと思えるほどに、常に本学部のために力を尽くしていただいた、そんな存在でいらっしゃいます。

個人的なことを申せば、私が学部長に就任した際に無理を通して学科長へのご就任をお願いし、重責に押しつぶされそうな折にいつもそばにいて助けてくださったことに、深く感謝いたしております。15年以上後輩の私が生意気なことを言っても、いつも優しく受け止めてくださって、本当にありがとうございました。

さらには、イギリスの在外研究中に押しかけた私をロンドンのパブで出迎えてくださったり、スウェーデンや私の著書を通じてご家族の皆様にもお世話になったりと、先生には本当に様々な機会とご縁をいただいたと、今あらためて感慨を覚えております。

白戸先生の研究業績の詳細な紹介は後に譲りますが、これまでの教育・研究活動、そして本学部の発展への多大なるご貢献に対して、重ねて深く感謝を申し上げます。そして、今後ますますの活躍を祈念申し上げ、献呈の辞とさせていただきます。

2021年12月吉日

国際日本学部長 鈴木賢志





白戸伸一教授・近影



# 白戸 伸一 教授 履歴・業績

## 目次

1. 学歴・学位・職歴・受賞歴等
2. 研究活動
  - (1) 著書
  - (2) 論文・解説
  - (3) 研究発表・講演等
  - (4) 所属学会
  - (5) 社会活動等
  - (6) その他の著作物

## 1. 学歴・学位・職歴・受賞歴等

(学歴・学位)

- 1970年 3月 愛媛県立松山東高等学校卒業  
 1974年 3月 早稲田大学教育学部社会科社会科学専修卒業  
 1977年 3月 早稲田大学大学院商学研究科商学専攻博士前期課程修了・商学修士  
 1985年 3月 明治大学大学院商学研究科商学専攻博士後期課程退学（単位修得）  
 2002年 9月 博士（商学）（明治大学）

(職歴)

- 1982年 4月 埼玉県八潮市教育委員会市史編纂室嘱託（～1985年 3月）  
 1985年 4月 沖縄短期大学英語科専任講師  
 1987年 4月 沖縄大学短期大学部助教授（法経学部兼任）  
 1990年 3月 浦和短期大学経営科専任講師  
 1992年 4月 浦和短期大学経営科助教授  
 1992年 4月 城西大学女子短期大学部非常勤講師（～1997年 3月）  
 1996年 4月 浦和短期大学教授（学生部長（1997年 4月～1999年 3月）、教務部長（1999年 4月～2001年 3月）、入試企画室長（2001年 4月～2003年 3月）、FD部長（2005年 4月～2007年 3月））  
 1996年 4月 明治大学商学部兼任・兼任講師（～2010年 3月）  
 2003年 4月 浦和大学総合福祉学部兼任講師（～2007年 3月）  
 2007年 4月 浦和大学総合福祉学部教授（～2008年 3月）  
 2008年 4月 明治大学国際日本学部教授（教務主任：2008年 4月～2013年 3月。学部長：2013年 4月～2014年 3月。学科長：2016年 4月～2020年 3月）  
 2012年 4月 明治大学大学院国際日本学研究科担当（～2022年 3月）  
 2014年 4月 SOAS, University of London客員研究員（～2016年 3月）  
 2015年 4月 Brunel University London客員教授（～2016年 3月）  
 2017年 4月 明治大学史資料センター運営委員（～2022年 3月）

## (受賞歴等)

- 1988年4月 文部省科研費奨励研究A(個人):「近代日本における商業経営組織化政策の史的考察」  
 1989年4月 文部省科研費一般研究B:「沖縄における共同化による観光産業基盤整備の可能性と地域特性の検証」研究代表(～1991年3月)  
 1989年4月 宇流麻学術研究助成基金助成研究:「沖縄における共同化による観光産業基盤整備の可能性と地域特性の検証」

## 2. 研究活動

## (1) 著書

- 1985年6月 共著 『近代農民運動と支配体制』 柏書房  
 1989年7月 共著 『戦後沖縄の社会変動と家族問題』 アテネ書房  
 1990年6月 共著 『地域マーケティング』 同文館  
 2000年9月 共著 『近代日本流通政策史』 白桃書房  
 2004年4月 単著 『近代流通組織化政策の史的展開』 日本経済評論社  
 2006年2月 共著 『日英経済史』 日本経済評論社  
 2021年3月 共著 『冷戦期アジアの軍事と援助』 日本経済評論社

## (2) 論文・解説

- 1980年3月 単著 「明治前期における同業者組織化政策」『明治大学大学院紀要』第17集 pp.109-125  
 1981年3月 単著 「同業者組織化政策の展開過程」『明治大学大学院紀要』第18集 pp.69-88  
 1982年3月 単著 「1910-20年代における同業者組織化政策の一考察」『明治大学大学院紀要』第19集 pp.17-41  
 1982年3月 単著 「大都市近郊農村における「地方改良」運動」『八潮市史研究』第4号 pp.186-216 八潮市史編纂室  
 1983年5月 単著 『八潮市史調査報告書7 八潮の産業組合史料』解説 pp.347-365 八潮市  
 1983年12月 単著 「第一次大戦前後における機業地織物業の展開と勸業政策」『埼玉地方史』 pp.1-26 埼玉県地方史研究会  
 1984年6月 単著 「近郊農業と信用組合」『八潮市史研究』第5号 pp.170-198 八潮市史編纂室  
 1985年3月 単著 「地域商業の構造変化と発展方向」『沖大経済論叢』第9巻第2号 pp.75-103 沖縄大学経済学会  
 1986年8月 単著 「「流通革命」論の検討」『沖縄短大論叢』第2巻第1号 pp.47-59 沖縄短期大学  
 1987年4月 単著 「戦間期の大都市近郊農村における産業組合経営」『八潮市史研究』第6号 pp.175-196 八潮市史編纂委員会  
 1987年9月 単著 「戦後沖縄における就業構造の推移」『沖大経済論叢』第12巻第1号 pp.49-71 沖縄大学経済学会  
 1988年5月 共著 『沖縄県民及び生協組合員の生活意識調査報告書』 pp.1-162 沖縄県民生活協同組合長期ビジョン調査特別委員会  
 1989年3月 単著 「第2次大戦後日本における流通組織化政策の検討」『沖縄短大論叢』第4巻第1号 pp.1-34 沖縄大学短期大学部  
 1991年3月 単著 「マーケティング・チャンネルと流通系列化規制の関連について」『浦和論叢』

- 第6号 pp.95-123 浦和短期大学
- 1991年7月 共著 『沖縄における共同化による観光産業基盤整備の可能性と地域特性の検証』(文部省科研費一般研究B 報告書) pp.1-104 観光産業共同化研究会
- 1993年3月 単著 「地域社会における社会経済の担い手としての産業組合」『八潮市史研究』第13号 pp.96-120 八潮市立資料館
- 1993年9月 単著 「明治前期における銀行類似会社の機能」『浦和論叢』第11号 pp.85-106 浦和短期大学
- 1994年12月 単著 「戦時体制期における工業組合政策と企業活動」『浦和論叢』第13号 pp.77-110 浦和短期大学
- 1995年12月 単著 「独占禁止法の基本理念と系列問題に関する検討」『浦和論叢』第15号 pp.29-57 浦和短期大学
- 1997年1月 単著 「明治10年代における同業者組織化政策に関する覚え書」『浦和論叢』第17号 pp.47-73 浦和短期大学
- 1998年12月 単著 「首都近郊都市における商業の発展と商工団体の機能」『浦和論叢』第21号 pp.63-92 浦和短期大学
- 1999年3月 単著 「地域社会における大型店問題と商工団体の役割」『明大商学論叢』第81巻3・4号 pp.49-66 明治大学商学研究所
- 1999年9月 単著 「大型店出店規制の課題」『流通』第12号 pp.69-79 日本流通学会
- 2000年3月 単著 「戦間期における地域商工団体の構成と機能に関する若干の検討」『明大商学論叢』第82巻第3号 pp.41-60 明治大学商学研究所
- 2001年1月 単著 「戦間期における在来産業の発展と産業組合の役割」『明大商学論叢』第83巻第1号 pp.1-25 明治大学商学研究所
- 2001年12月 単著 「第二次大戦後における流通組織化政策の検討」『浦和論叢』第27号 pp.25-59 浦和短期大学
- 2002年9月 単著 「近代日本における流通組織化政策の展開と地域経済」(明治大学大学院商学研究科博士学位請求論文) pp.1-359
- 2003年6月 単著 「日本的流通チャンネルの構造変化に関する統計分析」『浦和論叢』第30号 pp.27-47 浦和大学短期大学部
- 2004年6月 単著 「デジタル時代の流通論」『浦和論叢』第32号 pp.53-79 浦和大学短期大学部
- 2005年6月 単著 「高度成長期における産地織物業の組織化と構造改善事業(1)」『浦和論叢』第34号 pp.23-68 浦和大学短期大学部
- 2010年3月 単著 「産地織物業に対する構造改善事業の歴史的役割」『明治大学国際日本学研究』第2巻第1号 pp.1-25 明治大学国際日本学部
- 2013年3月 単著 「繊維産業構造改善政策と織物業における構造改善事業の歴史的検証」『明治大学社会科学研究所紀要』第52巻第1号 pp.1-24
- 2020年3月 単著 「イギリスにおけるまちづくり推進組織(BIDs)に関する検討」『明治大学社会科学研究所紀要』第58巻第2号 pp.73-96
- 2022年3月 単著 「まちづくり政策の現状と課題」『明治大学国際日本学研究』第14巻第1号 pp.69-87 明治大学国際日本学部

## (3) 研究発表・講演等

- 1983年1月 単独 「第一次世界大戦期における機業地織物業の展開と勸業政策」 埼玉県地方史研究大会
- 1993年3月 単独 八潮市立資料館企画展記念講演「地域における社会経済の担い手としての産業組合」
- 1998年11月 単独 「大型店出店規制の課題」第12回日本流通学会全国大会（福岡大学）
- 2003年10月 単独 「戦間期産地織物業の組織編成の検討」2003年度政治経済学・経済史学会秋季学術大会
- 2006年9月 単独 平成18年度さいたま市民大学講演「埼玉の地場産業」
- 2006年10月 単独 「高度経済成長期末の産地織物業における構造改善事業の歴史的役割」2006年度政治経済学・経済史学会秋季学術大会（明治大学）
- 2016年10月 単独 「イギリスにおけるまちづくり推進組織（BIDs）に関する検討」第30回日本流通学会全国大会（専修大学）
- 2019年11月 単独 八潮市立資料館開館30周年記念講演「明治期の『模範村』を出発点としてまちづくりを考える」

## (4) 所属学会

- 1979年4月 土地制度史学会（2002年に政治経済学・経済史学会に名称変更）
- 1983年4月 埼玉県地方史研究会
- 1985年4月 日本商業学会
- 1991年4月 日本流通学会（学会誌編集委員、学会理事、関東部会幹事）
- 1993年4月 経営史学会
- 2000年4月 アジア市場経済学会（～2014年）
- 2002年4月 社会経済史学会

## (5) 社会活動等

- 1985年4月 八潮市教育委員会市史編纂室嘱託・八潮市史調査専門員（～1989年3月）
- 1988年1月 沖縄県民生活協同組合長期ビジョン調査特別委員会委員（～1988年5月）
- 1990年3月 埼玉県蕨市史編纂委員会調査員（～1995年3月）
- 1990年4月 沖縄大学地域研究所特別研究員
- 1990年4月 明治大学百年史執筆協力者（～1994年10月）
- 1991年4月 東京都北区史編纂調査会参事・調査員（～1996年3月）
- 1992年11月 埼玉県県政史料・新出重要史料刊行事業調査員（～1997年3月）
- 1993年4月 草加市史専門調査員（～2001年3月）
- 1994年7月 埼玉県『埼玉の人物』編集刊行事業調査委員（～1997年3月）
- 2008年4月 明治大学体育会剣道部部长（～2014年3月）
- 2012年4月 埼玉県文化財保護審議会委員（～現在に至る）

## (6) その他の著作物

- 1989年7月 単著 「農業生産と産業組合」、「産業組合と諸産業」、「産業組合の発展」『八潮市通史編Ⅱ』pp.268-273、pp.329-338、pp.566-588 八潮市史編纂委員会
- 1992年10月 単著 「実業・専門教育の普及」、「大学・専門部体制の整備」『明治大学百年史第3』

- 卷通史編Ⅰ』pp.467-474、pp.507-517 明治大学
- 1994年10月 单著 「大学財政の逼迫」、「明治商業学校の設立」『明治大学百年史第4巻 通史編Ⅱ』pp.27-39、pp.83-86 明治大学
- 1995年3月 单著 「蕨宿物産会社の設立」、「織物業の発展」、「織物業の近代化」、「織物業の動揺と地域社会」、「経済統制下の諸産業」『新修蕨市史 通史編』pp.533-541、pp.614-621、pp.640-647、pp.664-672、pp.722-729 蕨市
- 1997年3月 单著 「租法」史料解説『埼玉県史料叢書3』pp.5-23 埼玉県
- 2000年9月 单著 「大正期の北区域」、「都市化への行政的対応」、「産業基盤整備と商工業の発展」、「総動員体制の構築」『北区 通史編 近現代編』pp.222-229、pp.250-256、pp.281-299、pp.411-417 東京都北区
- 2001年3月 单著 「浦和県申牒布令類」史料解説『埼玉県史料叢書5』pp.29-39 埼玉県
- 2001年3月 单著 「近代産業の萌芽」、「大正・昭和戦前期の草加」、「商工業の復興と展開」、「産業の発展」『草加市史 通史編下巻』pp.143-146、pp.429-450、pp.588-604、pp.671-683 草加市
- 2005年3月 单著 「電信、郵便、水漕、陸運、渡船」史料解説pp.22-27『埼玉県史料叢書2』埼玉県
- 2006年11月 单著 「IT化の歴史」、「ICカード」日本流通学会編『現代流通事典』pp.292-293、pp.312-313 白桃書房
- 2011年3月 单著 「日米繊維交渉期における織物産地における構造改善事業の歴史的検証」『明治大学社会科学研究所年報』第50巻 pp.47-48 明治大学社会科学研究所
- 2011年11月 单著 「専門学校令下の明治大学(1903-1911)」、コラム「留学生と経緯学堂」『創立期から大学昇格期に至る明治大学財政事情—会計帳簿の分析—』pp.55-73、pp.78-79 明治大学
- 2018年3月 単独インタビュー 「セーレン川田達男会長インタビュー」『大学史紀要』第24巻 pp.112-154 明治大学史資料センター
- 2019年3月 共同インタビュー 「栗原哲也氏（元日本経済評論社社長）聞き取り」『大学史紀要』第25号 p.187-222 明治大学史資料センター
- 2021年3月 单著 「利光鶴松略年表—明治法律学校初期卒業者の足跡—」『大学史紀要』第27号 pp.122-132 明治大学史資料センター
- 2021年3月 单著 「丸善石油元社長宮森和夫（財界人編）」『白雲なびく～はるかなる明大山脈～』コラム 明治大学史資料センター

## 日本流通史研究の回顧

白戸伸一

ここでは筆者の研究過程を振り返り、流通史研究に関する自分自身の到達点を確認しておきたい。

日本経済の発達過程について知りたいという素朴な思いで大学院に入り、近代化に国家の果たした役割が大きかったという見方から経済政策に着目し、「明治政府の「原蓄期」における農業政策—米穀政策に関する一考察—」という修士論文をまとめた。それ以来、経済政策史を重視した研究を続けることになった。

博士後期課程に入学後、まずは明治前期の殖産興業政策をより深く研究したいと考えたが、当初は農業政策への関心から『明治前期勸農事蹟輯録』や『農務顛末』といった分厚い史料集を紐解きノートを取ったりしていた。しかし政策が実際にどのように展開されていったのかピンと来ない。そこで居住地の文書館に出向き行政文書を閲覧してみると、政府からの「指令」に対し出先の県や町村から「伺」、「上申」の類いが提出されていて、双方からのやりとりと付帯文書などでやっとそれぞれの政策の具体的内容や浸透過程が見えてきた。ところが、末端の史料となると膨大であり、さらには「案」どまりか正式書類なのかも判別しづらい。おまけに紙の種類もいろいろで、印刷が擦れていて読めないところもある。ともあれ、菟藪版のような印刷物で活字の史料集で分からなかった詳細を知りえた時の興奮は一入であった。後述の同業組合研究では、このようにして入手し得た史料が貴重な手がかりを与えてくれた。

歴史研究では地方の「行政文書」、「地方文書」、さらには一次史料が重要である。脱線するが、農業問題への関心から一橋大学大学院での共同研究に加えてもらい、岐阜県農村地帯での史料調査のローラー作戦に加わったことがある。家主も減多に覗かない蔵に入り、蜘蛛の巣と埃まみれになりながら毛筆の日記や帳簿を発見し、その成果が6年後に森武磨編『近代農民運動と支配体制』にまとめられている。

流通過程への関心は、明治前期における同業組合という中間組織への着目に始まる。在来産業を輸出産業に編成する上で生糸や茶、絹や綿織物等の同業組合による組織化が効果的であり、その推進力は多くの場合商業資本に依存していた。これらを博士後期課程時代に研究し、そのことが専任教員として沖縄で商業史・商業総論を担当することに繋がった。沖縄では、独自の歴史・文化と地域特性に対する関心から地域経済研究や共同研究（戦後沖縄の社会変動と家族問題）に取り組むとともに、第二次大戦後の流通問題や流通政策へと研究領域を広げ、それらが次の大学での流通産業論やマーケティング論担当としての任用に繋がった。浦和短期大学へ異動後、さらに現代の流通・マーケティング研究を継続し、マーケティング・チャンネルやデジタル化時代の流通研究、大型店問題や流通と独禁法の関連、地域商工団体などの研究を進めた。その一方で、戦前・戦後の流通組織化政策に関する事例分析を継続し、戦前部分の研究を単著にまとめた。

明治大学国際日本学部の創設に伴い再び異動する幸運に恵まれたことから、日本の流通を歴史的に跡づけながら、さらには国際的視野でその特徴を解明する研究を進めることにした。そして、戦後の流通組織化政策の研究を継続するなかで、組織化の対象が多くの場合中小企業であり地域の産業集積と密接に繋がっていることから、戦前の同業組合においても産地を形成していた織物業に関心が向かった。国内の織物生産量は1970年にピークに達して以降低下し続け、衣類・アパレル製品の輸出入を見ると、1980年代以降、フランス、イタリア、ドイツ、韓国等と比べると日本は輸出入の不均衡が極端に激しくなっていた。2008年の衣類の輸入浸透率は95%に達しており、衣類はほとんど輸入品に頼る状態になった。織物業を含む国内の繊維産業は全体的には1980年代以降衰退傾向が顕著となっており、絹織物や化合繊維物の最大の産地であった北陸地方でも90年代以降、織物生産量を大幅に減らしている。政策的には構造改善事業が対置されていたので、繊維産業の行方を見届けるためにいくつかの事例研究を進めた。そこでは、産業集積もあり高度の染色加工技術もありながら、多くの経営が大手化合繊維メーカーや商社の賃加工業者に甘んじることにより結果的には淘汰される一方で、少数ではあるが、ヨーロッパのトップブランド企業から特注を受けたり、衣料でもB to Cのハイファッション市場へ、そしてそれ以外でも自動車内装材や環境・生活資材、メディカル基材といった新市場に進出して成功している企業があることを確認した。

さらに、1990年代末に小売業に対する流通政策が「まちづくり3法」に包摂される形で大きく転換したことから、地域経済における流通の役割を「まちづくり」という視点から再検討し、在外研究先の英国での調査とも対比しながらケーススタディを試みた。英国の場合、全国にチェーン展開している大手小売企業や不動産企業等が、中心市街地活性化に積極的である。日本では、大手小売企業の中には、まるで「焼畑農業」のように参入・撤退する企業もあるようで、地域とともにある姿勢が希薄なように思われる。地域経済や地域住民に対する重要な都市インフラの構成要素である小売業について、商品・サービスの単なる提供者ではなく地域社会の担い手として、さらには豊かさや個性の演出者としての役割を追究してきたが、未だ道半ばである。

国際日本学部に在籍して意識するようになった点は、国際的視野での比較研究を通じて日本的なるものを再発見すること、そしてそれをどう生かすか、ということだろう。しかし、こうして研究を振り返ると恥じ入ることばかりだが、講義やゼミを通じて多くの学生たちと議論・歓談し、彼らの成長に関わりえたこと、優れた研究者である同僚と交わり新たな見識や技術を教わったこと、有能な職員の皆さんと一緒に大学づくりに関わりあえたことなど、多くの幸運な機会に恵まれたことに感謝あるのみである。



[Articles]

# Developing a Japanese-English Parallel Corpus of “Japan, the Beautiful and Myself”

by Yasunari Kawabata

OYA, Masanori

## Abstract

This paper is an interim report of constructing a Japanese-English parallel corpus of Yasunari Kawabata’s Nobel Prize in Literature speech “Japan, the Beautiful and Myself” and its English translation by E. Seidensticker. This parallel corpus contains the Japanese-English translation pairs which are represented as syntactic dependency trees. This paper reports the background of the text, the procedure of how to construct the parallel corpus from it, and the result of analyzing the dependency trees of the sentences of both languages in the corpus in terms of their dependency distances.

Key words: Dependency distance, dependency trees, English-Japanese translation pairs, Multilingual parallel corpus

## 1. Introduction

This paper is an interim report of developing a Japanese-English parallel corpus of Yasunari Kawabata’s Nobel Prize in Literature speech “Japan, the Beautiful and Myself (henceforth JBM)” and its English translation by E. Seidensticker (Kawabata & Seidensticker 1969). This parallel corpus contains the Japanese-English translation pairs which are represented as syntactic dependency trees, along with their structural discrepancies between them.

The reasons of the construction of a Japanese-English parallel corpus using Kawabata’s speech are as follows: There are several Japanese-English parallel corpora available at present, such as Japanese-English Bilingual Corpus of Wikipedia’s Kyoto Articles ([https://alaginrc.nict.go.jp/WikiCorpus/index\\_E.html](https://alaginrc.nict.go.jp/WikiCorpus/index_E.html)), English and Japanese parts of the Parallel Universal Dependencies treebanks (<https://universaldependencies.org/#language-u>), to name a few. These parallel-corpus data have been expected to provide us with the ideal data for comparing and

contrasting natural languages, and have been employed to test a variety of hypotheses. For example, there has been an extensive amount of research on the comparison and contrast of syntactic dependency structures languages based on multi-lingual parallel corpora, especially in terms graph-centrality measures (Oya 2014, 2021, etc.) and dependency distance (Gibson, 1998, 2000; Gildea & Temperley, 2010; Grodner & Gibson, 2005; Liu 2007, 2008; Liu et al., 2017; Oya, 2011, 2013, etc.).

These parallel corpora, however, do not include English and Japanese literary texts, and the results based on these parallel corpora can be different from those based on literary works of English and Japanese. Against this background, it can be hypothesized that Japanese sentences written by one of the excellent Japanese literary figures exemplify the expressive power of this language, that their translations are expected to open up new possibilities for the expressiveness of English itself, and that the correspondence between them may yield some results which may be different from those obtained from the English-Japanese parallel corpora available at present.

This study is an attempt to verify the hypothesis mentioned above, and JBM has been chosen as the first data for constructing a larger-scale English-Japanese parallel corpus of literary works. In other words, this study constitutes a first step toward a more comprehensive understanding of literary works of both languages from a perspective of parallel corpus research.

This paper is structured as follows: First, the background of the corpus data is explained. The background includes the introduction of Yasunari Kawabata and Edward Seidensticker, and a summary of JBM. Section 3 explains the procedure of how to construct the parallel corpus from the text. Because of the fact that Japanese and English belong to different language family, it is often the case that there is not one-to-one correspondence between the words in their sentences, even though they are translation pairs and they express the same meaning. Section 4 introduces the idea of dependency distance with reference to previous works, especially those by Gibson (2008), Gildea and Temperley (2010) and Liu (2007). Section 5 reports the results of the quantitative investigation of the dependency trees of the sentences in JBM in terms of their dependency distances. The significance of the results is discussed in Section 6, along with the limitation of this study, and Section 7 concludes this article.

## 2. The background of the corpus data

Yasunari Kawabata (川端 康成, 1899-1972) is a Japanese novelist and short story writer. His works have enjoyed broad international appeal and are still widely read today. He won the Nobel Prize Laureate for Literature in 1968, the first Japanese author to receive the award. JBM is his Nobel Lecture on December 12th, 1968. Its translation from the original Japanese text into English was conducted by Edward George Seidensticker (1921-2007). He was an American

post-World War II scholar, historian, and preeminent translator of classical and contemporary Japanese literature. He was closely associated with the work of three major Japanese writers of the 20th century: Yasunari Kawabata, Jun'ichirō Tanizaki, and Yukio Mishima.

JBM is an introduction of Japanese classical literature and culture, influenced by Buddhist philosophy. In JBM, Kawabata refers to some classical Japanese poetries by notable Buddhist monks such as Dogen (1200-1253), Myoe (1173-1232), Saigyō (1118-1190), Ryōkan (1758-1831), Ikkyū (1394-1481), and he attempts to summarize their Weltanschauung based on Buddhist philosophy. He stresses the difference between Japanese and Western viewpoints, especially that of the definition of the word nothingness ("mu" in Japanese; its Chinese character is 無). He argues that Western nothingness is emptiness, while Japanese nothingness is "a universe of the spirit in which everything communicates freely with everything, transcending bounds, limitless. (translation by Seidensticker)." He also points out the significance of Zen philosophy, or intuitive enlightenment, which cannot be taught by masters, but to be attained by oneself through Zen practice; "Enlightenment comes not from teaching but through the eye awakened inwardly (translation by Seidensticker)."

### 3. The procedure of how to construct the parallel corpus from the text of JBM

The English version of JBM is available from the official Web site of the Nobel Prize<sup>1</sup>. Japanese original of JBM is not available online as text data, but published as a book (Kawabata & Seidensticker, 1969); therefore, the printed text has been typed manually to create its text data.

Japanese original sentences in JBM and their English translations are aligned manually, and the parallel corpus obtained as a result of this alignment is called *JBM Parallel Corpus* (henceforth *JBMPC*). During the process of this alignment, special attention has been paid to observe one-to-one correspondence between Japanese sentences and English ones. There are several instances in Seidensticker's translation where one Japanese sentence corresponds to more than one English sentence because the original Japanese sentence has so complex a structure (e.g., a subordinate clause with another subordinate clause) that he found it necessary to divide it into more than one sentence. When one Japanese sentence in *JBMPC* is translated into more than one English sentence, the Japanese sentence is divided into several phrases, and each of these phrases is aligned to its English translation.

There are also some instances in his translation where the order of Japanese sentences/phrases is changed in their English translations. In order to identify such cases, each sentence in *JBMPC* is assigned with two types of IDs. First, each sentence of both languages is assigned

---

1 Adapted from "Japan, the Beautiful and Myself" by Y. Kawabata, 1968, (<https://www.nobelprize.org/prizes/literature/1968/kawabata/lecture/>). ©The Nobel Foundation

with a *language-unique* ID; when a Japanese sentence is translated into more than one English sentence, it is divided into several phrases corresponding to English translations, and each phrase is assigned with a language-unique ID. Second, each Japanese-English translation pair is assigned with an across-language ID. The order of across-language ID is based on the order of English sentences; hence its order can be different from that of the language-unique IDs of Japanese. These different IDs allow us to notice the different order of sentences across these two languages.

Language-unique IDs have the following structure; The parentheses indicate that *\_phrase* in the ID is optional.

*paragraph\_sentence(\_phrase)*

For example, the language-unique ID “2\_2” is assigned to the second sentence of the second paragraph of a text, and the language-unique ID “3\_4\_5” is assigned to the fifth phrase of the fourth sentence of the third paragraph. With this dual assignment of IDs to the bilingual sentences, we can deal with the issue of one-to-many correspondence between, and reordering of, translation pairs across languages.

The first sentence of *JBM* provides us with a typical example of one-to-many correspondence between a Japanese sentence and English sentences, along with their reordering. Example (1) is the Japanese original, (2) is the romanized version of (1), and (3) is its English translation by Seidensticker:

(1)

春は花夏ほととぎす秋は月  
冬雪さえて冷しかりけり

道元禪師(一二〇〇年–五三年)の「本来ノ面目」と題するこの歌と、

雲を出でて我にともなふ冬の月  
風や身にしむ雪や冷たき

明恵上人(一一七三年–一二三二年)のこの歌とを、私は揮毫をもとめられた折りに書くことがあります。

(2)

haru wa hana natsu hototogisu aki wa tsuki  
fuyu yuki saete suzushikarikeri

Dogen zenshi (1200-53) no "honrai no menmoku" to daisuru kono uta to,

kumo wo idete ware ni tomonau fuyu no tsuki

kaze ya mini shimu yuki ya tsumetaki

Myoe shonin (1173-1232) no kono uta to wo, watashi wa kigou wo motomerareta ori ni kaku

koto ga arimasu.

(3)

"In the spring, cherry blossoms, in the summer the cuckoo.

In autumn the moon, and in winter the snow, clear, cold."

"The winter moon comes from the clouds to keep me company.

The wind is piercing, the snow is cold."

The first of these poems is by the priest Dogen (1200-1253) and bears the title "Innate Spirit". The second is by the priest Myoe (1173-1232). When I am asked for specimens of calligraphy, it is these poems that I often choose.

The alignment of the phrases/sentences above is summarized in the table below.

**Table 1**

*The alignment of the first seven Japanese-English translation pairs in JBMPC*

ID (Across)	ID (Spec J)	Japanese sentences	ID (Spec E)	English translations
1	1_1_1	haru wa hana natsu hototogisu	1_1	"In the spring, cherry blossoms, in the summer the cuckoo.
2	1_1_2	aki wa tsuki fuyu yuki saete suzushikarikeri	1_2	In autumn the moon, and in winter the snow, clear, cold."
3	1_1_4	kumo wo idete ware ni tomonau fuyu no tsuki	2_1	"The winter moon comes from the clouds to keep me company.
4	1_1_5	kaze ya mini shimu yuki ya tsumetaki	2_2	The wind is piercing, the snow is cold."
5	1_1_3	Dogen zenshi (1200- 53) no "honrai no menmoku" to daisuru kono uta to,	3_1	The first of these poems is by the priest Dogen (1200- 1253) and bears the title "Innate Spirit".
6	1_1_6	Myoe shonin (1173- 1232) no kono uta to wo,	3_2	The second is by the priest Myoe (1173-1232).
7	1_1_7	watashi wa kigou wo motomerareta ori ni kaku koto ga arimasu.	3_3	When I am asked for specimens of calligraphy, it is these poems that I often choose.

Note. ID (Across) = across-language IDs, ID (Spec J) = language-unique IDs for Japanese, ID (Spec E) = language-unique IDs for English.

The language-specific IDs for both languages indicate several aspects of the alignment of the sentences in JBMPC: First, the first paragraph of the original Japanese consists of one sentence and it is divided into seven phrases, each of which corresponds to an English sentence. Second, its English translation has three paragraphs; the first and the second paragraphs contain two sentences each, and the third contains three sentences. The phrase with the ID “1\_1\_3” in Japanese is reordered to follow the phrase with the ID “1\_1\_5,” and it is aligned with its English sentence with the ID “3\_1.” This indicates that it corresponds to the first sentence of the third paragraph of English translation. All the other Japanese phrases are aligned with their English counterparts accordingly.

JBMPC contains 256 sentence pairs, with 145 Japanese sentences and 253 English sentences. On average, one Japanese sentence is translated into 1.77 English sentences.

The data in JBMPC will be used for a variety of linguistic research, such as structural divergence between Japanese and English (Oya, 2017, 2020) or quantification of connectedness of sentences within a text (Oya, 2021a). It is expected that we can obtain interesting insights into the difference(s) of these two languages based on JBMPC, due to the uniqueness of its data (literary work by a notable novelist and its translation by a first-rate translator), along with its quality of hand-coded bilingual alignment which may still be beyond the capacity of modern computer. Among other research topics including those mentioned above, this study has chosen to focus on the quantificational analysis of the syntactic dependency structure of these languages, especially in terms of dependency distance.

#### 4. Dependency distance

This section introduces the idea of dependency distance (henceforth DD), which is one of the measures for structural complexity of a sentence (Gibson, 1998, 2000; Gildea & Temperley, 2010; Grodner & Gibson, 2005; Liu, 2007, 2008; Liu et al., 2017; Oya, 2013).

DD is the number of words between a dependency head and its dependency tail. Consider the example sentence (1):

(1)

Sarah has written this thesis.

The noun “thesis,” the 5th word in the sentence, depends on the verb “written,” the 3rd word. In this dependency relation, “written” is the dependency head (the word on which another word depends) and “thesis” is the dependency tail (the word which depends on another word). The DD between them is two ( $5-3=2$ ).

As such, we can calculate the DD of each dependency in a sentence. In the example sentence (1) above, the dependencies and their DDs are as follows:

"Sarah" depends on "has," and the DD is 1. The auxiliary "has" does not depend on any other word in this sentence; this is the Root of the sentence. The present participle "written" depends on the auxiliary "has," and the DD is 1. The demonstrative "this" depends on the noun "thesis," and the DD is 1. The noun "thesis" depends on the present participle "written," and the DD is 2. Then, we can calculate the average DD (henceforth ADD) of a sentence. In the example sentence above, it is  $(1+1+1+2) / 4 = 1.25$ .

Sentences with the same word count can have different ADDs. For example, consider the following example:

(2)

Her thesis has been reviewed.

The ADD of this example sentence is 1, because all the dependency distances in this sentence are 1; "Her" depends on "thesis," "thesis" depends on "has," The past participle "been" depends on "has," "reviewed" depends on "been," and the word "has" does not depend on any other word in this sentence. The ADD of the example sentence (2) is smaller than the ADD of the example sentence (1). The word count of these two sentences is the same, but their ADDs are different.

Sentences with larger word counts do not necessarily have longer DDs. For example, consider the following example:

(3)

The reviewers must have evaluated it positively.

The ADD of this example is approx. 1.167: The word "The" depends on "reviewers," and the DD is 1; "reviewers" depends on "must," and the DD is 1; "have" depends on "must," and the DD is 1; "evaluated" depends on "have," and the DD is 1; "it" depends on "evaluated," and the DD is 1; "positively" depends on "evaluated," and the DD is 2; and the auxiliary "must" does not depend on anything in this sentence. The ADD of (3) is  $(1+1+1+1+1+2) / 6 \approx 1.167$ , which smaller than the ADD of (1), even though the word count of (3) is 6, which is larger than that of (1).

Gibson (2000) proposed a theory of linguistic complexity, the Dependency Locality Theory (DLT). DLT assumes that sentences become more complex in proportion to the distance between dependency heads and their dependency tails. This assumption is based on the insight concerning the process of natural language comprehension. In the process, each of the words in a sentence is assumed to be stored in the memory, and they are integrated into the syntactic structure of the sentence. With respect to the dependency of words in a sentence, the integration of words implies the establishment of a dependency relationship between the new word and its dependency head or tail within the structure in the memory.

For example, imagine that we are going to comprehend a sentence "David has written this thesis." First, we hear the word "David," and it will be stored in our memory. We expect that

something follows after this word which can be its dependency head or tail. Second, we hear the word “has,” and then we know that this word is the dependency head of the word “David,” and the structure based on the sequence “David has” will be stored in the memory. Next, we hear the word “written,” and we know that this word depends on “has,” and we have the structure based on the sequence of words “David has written.” Based on our knowledge that the verb “written” is a transitive verb, we further expect that we will hear its object. Then, we hear the word “the,” yet we cannot find anything in the structure in our memory which can establish a dependency relationship with this “the,” and we wait for something more to hear. Then we hear the word “thesis” and the word “the” depends on this “thesis” and also this “thesis” depends on “written” in the structure “David has written.” As a result, we have the structure of the whole sentence “David has written this thesis.” in our memory. As such, the words in a sentence are supposed to be integrated into the structure one by one, resulting into a larger, more complex structure.

In this context, the complexity of the structure of a sentence can be measured in terms of the distance (or locality) between a dependency tail and its dependency head, because a longer dependency distance means a heavier memory load than shorter ones.

Natural languages tend to prefer shorter dependency distances because they are considered to require smaller burden on memory than longer ones (Gibson, 2000; Futrell et al., 2015; Liu, 2008, etc.), and different languages show different preferences for dependency distances (Oya, 2021b). Oya (2021b) investigated the ADDs of the languages in a multi-lingual parallel corpus (Parallel Universal Dependency) (Zeman, 2015; Zeman et al., 2020) and found that on average Japanese ADD is smaller than English ADDs, and different dependency types show different preferences for DDs; more specifically, Oya (2021b) found that the ADDs of the dependency type *nominal subject* (the dependency type between the main verb of a clause and the noun phrase functioning as its subject) is much longer in Japanese (approx. 9.99) than in English (approx. 3.01), while the ADDs of the dependency type *object* (the dependency type between the main verb of a clause and the noun phrase functioning as its object) is only slightly longer in Japanese (approx. 2.82) than in English (approx. 2.21).

The insights obtained by these previous studies should be further examined by a variety of genres of texts, which include literary texts, and it is natural to assume that JBM is one of such texts. Based on this assumption, the research questions of this study are stated as follows:

- (1) Do Japanese sentences in JBMPC have shorter dependency distances than English sentences in JBMPC?
- (2) Do different dependency types of the two languages in JBMPC show different tendencies in terms of dependency distances?

## 5. Analysis on the dependency trees of the sentences in JBMPC in terms of their dependency distances

This section describes the procedure through which the research questions stated in the previous section are examined, and reports its results.

### 5.1 Procedure

The dependency distances of the sentences in JBMPC are calculated in the following procedure. First, the Japanese sentences in JBMPC are automatically parsed by JUMAN/KNP (Kawahara & Kurohashi, 2006), and the parse output is converted automatically by an original Ruby script into the format of Universal Dependency (UD) (Zeman, 2015; Zeman et al., 2020). Second, the English sentences in JBMPC are automatically parsed by Stanford Parser (de Marneffe, McCarthy & Manning, 2006). Unlike the parse output of JUMAN/KNP, the output of the current version of Stanford Parser follows the format of UD, therefore we need not to convert it into UD.

Based on the parse output of the Japanese and English sentences, their dependency distances are calculated and compared with each other. Not only the ADDs of both languages, the ADDs of the following dependency types are also calculated; the ADDs between verbs and their subjects (nominal subjects, henceforth NSUBs), those between verbs and their objects (OBJs), those between nouns and adjectives modifying them (ADJs), those between verbs and adverbial clauses modifying them (ADVCLs), and those between nouns and relative clauses (RELCLs). These dependency types focused on in this study are the same as those in Oya (2021b).

### 5.2 Results

Table 2 shows the descriptive statistics of all the dependency distances in Japanese and English in JBMPC.

**Table 2**  
*The descriptive statistics of all the dependency distances in Japanese and English in JBMPC*

	All	Sum of DD	ADD	Median	Mode	Min.	Max.	SD
Japanese	2271	7059	3.108	1	1	1	34	3.987
English	4763	16232	3.407	2	1	1	57	4.749

The average dependency distance (henceforth ADD) of all the dependencies in Japanese ( $Mdn=1$ ) is shorter than that in English ( $Mdn=2$ ). A Mann-Whitney U test shows that the difference between the ADDs of Japanese and those of English is statistically significant,  $U(N=2271, N=4763)=79632.49, z=8.52, p < .01$ .

Table 3 shows the descriptive statistics of the dependency distances between verbs and their subjects in Japanese and English in JBMPC. The ADD of this type of dependency in Japanese ( $Mdn=2$ ) is longer than that in English ( $Mdn=2$ ). A Mann-Whitney U test shows that the difference between the ADDs of this type of dependency in Japanese and those in English is statistically slightly significant,  $U(N=184, N=430)=34436.5, z=2.54, p < .05$ .

**Table 3**

*The descriptive statistics of the dependency distances between verbs and their subjects in Japanese and English in JBMPC*

	NSUBJs	Sum of DD	ADD	Median	Mode	Min.	Max.	SD
Japanese	184	670	3.641	2	1	1	20	3.402
English	430	1439	3.347	2	1	1	37	4.476

Table 4 shows the descriptive statistics of the dependency distances between verbs and their objects in Japanese and in English in JBMPC.

**Table 4**

*The descriptive statistics of the dependency distances between verbs and their objects in Japanese and English in JBMPC*

	OBJs	Sum of DD	ADD	Median	Mode	Min.	Max.	SD
Japanese	129	220	1.710	1	1	1	13	1.637
English	176	432	2.455	2	1	1	29	2.842

The ADD of this type of dependency in Japanese ( $Mdn=1$ ) is shorter than that in English ( $Mdn=2$ ). A Mann-Whitney U test shows that the difference between the ADDs of this type of dependency in Japanese and those in English is statistically significant,  $U(N=129, N=176)=7796.5, z=4.67, p < .01$ .

Table 5 shows the descriptive statistics of the dependency distances between nouns and the phrases which modify these nouns in Japanese and in English in JBMPC.

**Table 5**

*The descriptive statistics of the dependency distances between nouns and the phrases which modify these nouns in Japanese and in English in JBMPC*

	NMODs	Sum of DD	ADD	Median	Mode	Min.	Max.	SD
Japanese	828	1537	1.856	1	1	1	20	2.264
English	735	2831	3.852	3	3	1	44	3.515

The ADD of this type of dependency in Japanese ( $Mdn=1$ ) is shorter than that in English ( $Mdn=3$ ). A Mann-Whitney U test shows that the difference between the ADDs of this type of dependency in Japanese and those in English is statistically significant,  $U(N=828, N=735)=102496.5, z= 22.66, p < .01$ .

Table 6 shows the descriptive statistics of the dependency distances between verbs and the adverbial clauses modifying these verbs in Japanese and in English in JBMPC.

**Table 6**

*The descriptive statistics of the dependency distances between verbs and the adverbial clauses modifying them in Japanese and in English in JBMPC*

	ADVCLs	Sum of DD	ADD	Median	Mode	Min.	Max.	SD
Japanese	223	910	3.910	2	1	1	26	4.15
English	71	664	9.350	7	7	1	46	9.12

The ADD of this type of dependency in Japanese ( $Mdn=2$ ) is shorter than that in English ( $Mdn=7$ ). A Mann-Whitney U test shows that the difference between the ADDs of this type of dependency in Japanese and those in English is statistically significant,  $U(N=223, N=71)=4432.5$ ,  $z=7.6$ ,  $p < .01$ .

Table 7 shows the descriptive statistics of the descriptive statistics of the dependency distances between nouns and the relative clauses modifying these nouns in Japanese and in English in JBMPC.

**Table 7**

*The descriptive statistics of the dependency distances between nouns and the relative clauses modifying them in Japanese and in English in JBMPC*

	RELCLs	Sum of DD	ADD	Median	Mode	Min.	Max.	SD
Japanese	119	199	1.670	1	1	1	16	2.37
English	87	352	4.450	3	2	1	14	3.01

The ADD of this type of dependency in Japanese ( $Mdn=1$ ) is shorter than that in English ( $Mdn=3$ ). A Mann-Whitney U test shows that the difference between the ADDs of this type of dependency in Japanese and those in English is statistically significant,  $U(N=119, N=87)=1598.5$ ,  $z=8.7$ ,  $p < .01$ .

## 6. Discussions

The result in Table 1 shows that Japanese sentences in JBM have the shorter ADD than English sentences in JBM. This means that Japanese sentences tend to be simpler than English ones, as far as the ADDs of their sentences are considered to be one of the measures for the complexity of syntactic structure. The result arguably suggests that Japanese language can express some ideas more simply than English can. More specifically, Zen philosophy—the main topic of JBM—is expressed quite concisely in Japanese, and it sometimes requires translators to express it with English sentences with more complexity.

The results shown in Table 2 through 6 indicate that different dependency types of the two languages in JBM show different tendencies in terms of dependency distances. The DD of dependency type NSUBJ of Japanese is longer than that of English, which is outstanding because all the other dependency types show that Japanese language prefers shorter DDs to longer ones compared to English. This result reflects the typological difference between Japanese, which is an SOV languages, and English, an SVO language. It is natural that the ADD of SOV languages tends to be longer than that of SVO languages, because there is O between S and V in SOV languages, while there is nothing between S and V in SVO languages. This result also supports the insight obtained in Oya (2021b) that the ADDs of different dependency types reveal more detailed differences across languages.

There are some issues revealed in the results which are not directly concerned with the research questions of DDs. First, the ratio of adverbial clauses to all the dependencies is about seven times larger in Japanese sentences in JBM ( $223/2271 \approx 0.098$ ) than that in English ( $71/4763 \approx 0.015$ ). This difference of usage is also observed in PUD (Oya 2021b), which has 28,784 dependencies in Japanese, and 21,168 dependencies in English, and the ratio of adverbial clauses to all the dependencies in Japanese in PUD ( $903/28784 \approx 0.031$ ) is about three times larger than that in English in PUD ( $293/21168 \approx 0.013$ ). This result may suggest the preference of Japanese language for adverbial clauses compared to English. Second, the ratio of relative clauses to all the dependencies in Japanese sentences in JBM ( $119/2271 \approx 0.052$ ) is about 2.8 times larger than that in English ( $87/4763 \approx 0.018$ ), yet this result does not go along with the result in PUD, in which the number of relative clauses in Japanese ( $211/28784 \approx 0.007$ ) is about *seven times smaller* than that in English ( $1119/21168 \approx 0.052$ ). This discrepancy of the use of relative clauses between JBM and PUD may suggest the stylistic characteristics of the Japanese sentences in JBM and also stylistic judgment of translators between English and Japanese; the Japanese sentences in PUD are translated from English, and the translators of PUD might have preferred not to use relative clauses when translating the English sentences into Japanese. Seidensticker, on the other hand, might have preferred not to use relative clauses when translating the original Japanese sentences into English. The issue of translators' preference across languages will be one of the research questions using the data of parallel corpus.

As a final remark, it is worth mentioning that the data in JBMPC can be analyzed quantitatively, using measures for the structural properties of dependency structures, such as degree centralities, closeness centralities, or graph density (Oya 2014, 2021a), and also qualitative analyses of dependency structures, such as syntactic divergence patterns between English and Japanese (Oya 2017, 2020). These topics will be further investigated in future research.

## 7. Conclusion

This paper reported the development of JBMPC, a Japanese-English parallel corpus of Yasunari Kawabata’s Nobel Prize in Literature speech “Japan, the Beautiful and Myself (JBM)” and its English translation by E. Seidensticker. JBMPC contains the syntactic dependency trees of the Japanese-English translation pairs of JBM. This paper explained the background of JBM and the procedure of how to develop the parallel corpus from it, with special attention to the alignment of the translation pairs across these two languages. Through the analyses of the dependency distances of the sentences in JBMPC, it was found that Japanese sentences in JBMPC have shorter dependency distances than English sentences in JBMPC, and different dependency types of the two languages in JBMPC show different tendencies in terms of dependency distances.

## References

- de Marneffe, M.C., B. MacCartney, & C.D. Manning. (2006). Generating typed dependency parses from phrase structure parses. *LREC 2006*.
- Futrell, R., K. Mahowald, & E. Gibson. (2015). Large-scale evidence for dependency length minimization in 37 languages. *Proceedings of Natural Academy of Science*, 112(33), 10336-10341. <https://doi.org/10.1073/pnas.1502134112>
- Gibson, E. (1998) Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, 1–76.
- Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. P. Marantz, Y. Miyashita, & W. O’Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium* (pp. 95–126). MIT Press.
- Gildea, D. & Temperley, D. (2010). Do grammars minimize dependency length? *Cognitive Science*, 34(2), 286–310. <https://doi.org/10.1111/j.1551-6709.2009.01073.x>
- Grodner, D. & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science*, 29(2), 261–290. [https://doi.org/10.1207/s15516709cog0000\\_7](https://doi.org/10.1207/s15516709cog0000_7)
- Kawabata, Y & E. Seidensticker. (1969). *Utsukushii Nihon no Watashi* [Japan, the Beautiful and Myself] Tokyo: Kodansha.
- Kawahara, D. & S. Kurohashi. (2006). A fully-lexicalized probabilistic model for Japanese syntactic and case structure analysis. *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL2006)*, 176–183.
- Liu, H. (2007). Probability distribution of dependency distance. *Glottometrics*, 15, 1–12.
- Liu, H. (2008). Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, 9 (2), 159–191.
- Liu, H., C. Xu, and J. Liang. (2017). Dependency distance: A new perspective on syntactic

- patterns in natural languages. *Physics of Life Reviews*, 21, 171–193. <https://doi.org/10.1016/j.plrev.2017.03.002>
- Oya, M. (2011). Syntactic dependency distance as sentence complexity measure. *Proceedings of the 16th International Conference of Pan-Pacific Association*, 313–316.
- Oya, M. (2013). Degree centralities, closeness centralities, and dependency distances of different genres of texts. *Selected papers from the 17th Conference of Pan-pacific Association of Applied Linguistics*, 42–53.
- Oya, M. (2014). *A study of syntactic typed-dependency trees for English and Japanese and graph-centrality measures* [Ph.D. dissertation, Waseda University]. Waseda University Repository, <http://hdl.handle.net/2065/45597>
- Oya, M. (2017). Syntactic divergence patterns among English translations of Japanese one-word sentences in a parallel corpus. *English Corpus Studies*, 24, 19–40.
- Oya, M. (2020). Structural divergence between root elements in English-Japanese translation pairs. *Global Japanese Studies Review, Meiji University*, 12(1), 107–126.
- Oya, M. (2021a). Analysis and quantification of the network of lexical connectedness within a text based on metrics used in network analysis. *Global Japanese Studies Review, Meiji University*, 13(1), 15–38.
- Oya, M. (2021b). Three types of average dependency distances of sentences in a multilingual parallel corpus. *PACLIC 35*. Shanghai.
- Zeman, D. (2015). Slavic languages in universal dependencies. *Slovak 2015: Natural Language Processing, Corpus Linguistics, E-learning*. Slovakia.
- Zeman, D., Nivre, J., Abrams, M., Ackermann, E., Aeppli, N., Aghaei, H., Agić, Ž., Ahmadi, A., Ahrenberg, L., Kennedy Ajede, K., et al. (2020). *Universal Dependencies 2.7*, LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. <http://hdl.handle.net/11234/1-3424>.

[Articles]

# Integrating Complex Variables in the Measurement of L2 Speech Production: Focusing on Complexity, Accuracy, Fluency, and Creativity

OZEKI, Naoko  
WANG, Wei-Tung  
OTSUKA, Kiyotaka

## Abstract

This study focuses on how the speaking performance of Japanese learners of English changes over time regarding creativity, as well as complexity, accuracy, and fluency (CAF). By adopting the perspectives from dynamic systems theory (DST), this study describes the nonlinear changes of learners' performance over a period of time. Moreover, focusing on creativity, in addition to CAF, enables us to investigate speaking performance from a different viewpoint.

The participants of this study are Japanese university students who learn English as a foreign language. They were assigned tasks to tell stories based on cartoons. Their speaking performance was recorded and transcribed. The data was then analyzed regarding the CAF of their English production. In addition, the creativity of participants' performance was evaluated based on a rubric containing categories such as image, characterization, voice, and storyline. The results indicated that learners' speaking performance did not improve linearly in terms of CAF which also explained language development from the perspectives of DST. Furthermore, although learners' speaking performance in terms of CAF fluctuated, the creativity showed a rather stable pattern. Based on these results, methods for providing speaking instruction to different developmental patterns is also discussed.

Keywords: speaking performance, complexity, accuracy, fluency, dynamic systems theory, creativity

## Introduction

For second language (L2) learners, it is important to become proficient in the production of the target language. Without productive proficiency, L2 communication can break down easily. Although written communication is an essential aspect of L2 use, spoken communication is

equally important, as it is required in a broad range of situations, such as everyday conversations, workplace interactions, and lectures at schools.

In L2 education, teaching how to write and speak the target language necessarily involves the evaluation of such performances. Thus, product-oriented approaches are common in L2 teaching and learning, and when evaluating learners' performance, teachers often use rubrics, with categories related to particular linguistic or communicative aspects, for assessment.

Although a product-oriented approach can explain the improvement of learners' performance in a linear way, L2 learners' language use often fluctuates and it may be unrealistic to claim that a teaching intervention directly leads learners to develop their proficiency. As viewpoint of dynamic systems indicates, learners' developmental paths are nonlinear (Larsen-Freeman & Cameron, 2008).

Consequently, L2 studies have come to focus on the dynamic quality of language learning. From such perspectives, it is possible to explain changes in language proficiency more in detail by investigating the multiple factors involved in L2 learning longitudinally as well as cross-sectionally. In other words, researchers might focus more on the process of individual learners' language learning. For example, focusing on the change of one learner's spoken language over a certain period of time may result in particular implications for dealing with the stagnation of learners' development.

Much of the focus of assessment in product-oriented approaches focuses strictly on linguistic aspects. And, as noted earlier, common criteria for evaluating learners' performance includes the complexity, accuracy, and fluency (CAF) (Crookes, 1989) of their products. Nonetheless, these criteria do not necessarily cover every aspect of language use. For example, even if speakers perform well in terms of linguistic factors, they may not perform well in other areas that may also be important, in various situations. Therefore, rubrics for evaluating productive proficiency need to be expanded in terms of how learners effectively make themselves understood with limited linguistic resources. One important aspect to which some attention has been paid is to the levels of creativity evident in language production. Evaluation of creativity in productive performance focuses on aspects such as the story's details, the teller's voice, and the expression of the story's purpose (Mozaffari, 2013).

Based on these conceptions, this study focuses on the changes of L2 learners' speaking performance from the perspectives of dynamic systems theory (DST) by evaluating four components: CAF and creativity. It is expected that the combination of the analysis regarding the dynamic fluctuation of learners' performance and multiple aspects of spoken language will lead to a deeper understanding of how L2 learners' productive proficiency can be improved. The next section reviews the previous studies related to this study.

## Literature review

This section introduces research related to a) complexity, accuracy, and fluency; b) dynamic systems theory; and, c) creativity evaluation criteria.

### Complexity, accuracy, and fluency

Measuring productive proficiency of L2 learners has been one of most controversial issues in L2 education. As for speaking proficiency, it is usually measured by using certain tasks which require participants to produce the target language under various conditions.

As Crookes (1989) implied, the three major aspects in measuring speaking proficiency are CAF. According to Crookes (1989), complexity includes both syntactic complexity and grammatical complexity such as subordinate clauses per utterance. Accuracy focuses on the number of errors in the speakers' production. The rate of errors to the total number of units or words is commonly used as an indicator of accuracy. Fluency is related to the quantity of production—measuring the number of syllables and pauses—in a given amount of time. Since CAF have been used as the indicators of analysis, the base unit used for analysis varies among researchers. Compared with written data, which are usually rather organized, spoken data contain features such as false starts, repetition, and self-corrections. Different units have been developed for analyzing spoken data. Foster, Tonkyn, and Wigglesworth (2000) proposed a unit called an "Analysis of Speech Unit" (AS-Unit) for analyzing spoken data. An AS-Unit takes the features of spoken text into consideration. It includes an independent clause or sub-clausal unit with any subordinate clauses (Foster, Tonkyn, & Wigglesworth, 2000).

Although CAF has been used in a number of studies to measure L2 production and performance, the literature has yet to reach an agreement as to whether CAF are solely valid measures. Skehan (2009) points out that a lexical aspect should be included in the measurement of L2 production because lexical choice may influence syntax, which, in turn, affects complexity and accuracy. Additionally, the influence of task conditions on speakers' performance should not be ignored. Ozeki and Hiromori (2016) investigated CAF based on ten spontaneous speaking tasks from one university EFL learner. Pre- and post-tests were also included. The results indicated that learner's speaking performance improved. By doing the same type of task over a period of time, the effect of task repetition might influence learners' performance in terms of CAF.

However, the results of the effect of task repetition are rather mixed (Hsu, 2019; Kim, & Tracy-Ventura, 2013). It is difficult to draw a clear conclusion regarding the effect of task repetition in terms of CAF in spoken performance since there are various factors involved. In order to track how learners' speaking performance changes over time, DST is useful. The details

and the theoretical background of DST will be introduced in the next section.

### **Dynamic systems theory**

Regarding L2 speaking proficiency studies, research methodologies focusing on a large amount of data, collected from a number of participants, are common (Dörnyei, 2007). Such methodologies are based on the conception that the effects of certain teaching/learning approaches on learners share common tendencies. Thus, it is necessary to gather an adequate amount of data in order to measure the effectiveness of each approach.

On the other hand, as Larsen-Freeman and Cameron (2008) state, relatively new research methodologies on L2 speaking proficiency include a DST perspective. Methodologies grounded in DST assume that various factors inside and outside of L2 learners interconnectedly influence learners' performance and development of their proficiency in dynamic ways. Thus, the same event has different meanings for each individual learner and even one learner's improvement cannot be described in a linear way. The emergence of DST indicates that L2 development can be better depicted as a continuous and multidimensional process (Lowie, 2012). Larsen-Freeman (2014) also listed nonlinearity as one characteristic of DST because it is difficult to predict a cause and its effect. Furthermore, the nonlinear features of DST indicate that research methods have undergone a transition from a traditional, product-oriented focus to a more process-oriented focus.

To investigate the complexity in L2 learning, then, it is necessary to avoid focusing on simple cause-effect relationships between a few variables. Rather, we need to look carefully into how various parameters are interconnected and to consider that such relationships themselves are in constant fluctuation. In recent years, there have been a few important studies looking at the relationship between DST and language development.

De Bot, Lowie, and Verspoor (2007) collected writing journals from eleven low-intermediate high school EFL learners over six weeks. After analyzing the length of the written sentences, they discovered that learners' writing development was nonlinear and interaction between cognitive and social factors influenced complexity of writing. Not only written performance, but also oral performance has been examined by using the DST framework. Ozeki and Hiromori (2016) analyzed one intermediate EFL university learner doing ten speaking tasks over two weeks in terms of CAF. With quantitative and qualitative analyses, the results indicated that learner's spoken performance fit the features of DST. However, since it was a short span of two weeks and there was only one participant, a certain number of participants and longitudinal studies over longer periods are essential for future research.

Recently, Yu and Lowie (2020) investigated ten Chinese EFL learners' oral performance

in terms of complexity and accuracy based on DST. Twelve speaking tasks based on language proficiency tests were recorded through one semester. Participants had one minute to prepare before the task and the task was around two or three minutes. The results showed that participants' performance of accuracy and complexity had up and downs among participants over time which indicates the nonlinearity, as one characteristic of DST.

This study builds on these previous ones by examining how production changes longitudinally, in terms of CAF from a DST perspective. Moreover, the additional factor of creativity is also explored in this study.

### **Creativity evaluation criteria**

As mentioned in the previous section, CAF have been used as the criteria to evaluate learners' performance. Since CAF are analyzed based on linguistic and grammatical features, the evaluation framework is more analytical. On the other hand, measuring creativity is not as simple as measuring CAF. The issue of creativity has been discussed mainly in the field of writing education. However, the evaluation of learners' creativity tends to depend on teachers' subjective judgement, which has led to a desire for researchers to develop more valid and reliable criteria.

Mozaffari (2013) notes that teachers' subjective judgement is a drawback in assessing creativity. She also argues that theories of creativity sometimes lack validity and that attributes that are included in some rubrics for evaluating creativity are not truly related to creativity. As a result, she developed her own analytical rubric to measure creativity in writing. The rubric consists of descriptors for four major aspects of creative writing: *image*, *characterization*, *voice*, and *story*. Image refers to levels of significant details of imaginative writing. In other words, creative writing with rich image avoids using abstractions, generalizations, and judgements by relying on literary tropes and speaking/thinking voices. Voice focuses on the kind of vocabulary being used, as choices in syntax can enrich the previous category of image. Characterization refers to how characters in creative writing are animated. If a character is described rather indirectly, with descriptive details about the character's physical appearance, actions, thoughts, etc., it is regarded as having more creativity. And finally, story is related to the purpose of the writing. These aspects construct the rubric for measuring creativity. In this study, the author also tests the rubric and finds it to be both reliable and valid for assessing creativity in writing.

Maryam and Saeed (2018) also developed a rubric for creative writing with input from previous research and current professors of creative writing. They argue that the nine-category rubric not only helps teachers set course objectives, design the syllabus, and develop curriculum but also provides learners with accurate feedback which leads to deeper learning.

While creativity in writing has been discussed, and such assessment criteria has been proven to facilitate both learning and teaching, the evaluation of creativity in speaking has not yet been explored thoroughly. Based on the results of Ozeki and Hiromori (2016), learners' speaking performance in terms of CAF showed dynamic patterns. However, the relationship between CAF and creativity has not been examined yet. Therefore, this study attempts to evaluate learners' creativity in speaking performance while investigating CAF. By doing so, creativity of learners' spoken performance can be examined analytically and objectively, and the results can be applied to the teaching and learning of speaking skills.

To sum up, this study will address the following research questions by analyzing learners' speaking performance.

1. How does learners' speaking performance change over a period of time in terms of complexity, accuracy, and fluency?
2. How is the change in learners' creativity in spoken production related to the changes in CAF?

## **Methodology**

### **Participants**

Fifteen undergraduate students in a private university in Tokyo participated in the study. Their majors were related to international affairs and their English proficiency was at a B2 level, judging from their TOEFL iBT scores.

### **Procedure**

The participants were assigned 20 speaking tasks. The speaking task consisted of three to ten pictures and they improvised a story based on the pictures within a three-minute time limit. A set of pictures were taken from cartoons of Heaton (1995, 1997) and those of Watterson (2005). Because they had to work on 20 speaking tasks and they were told to do one speaking task per day, it took them 20 days to finish the tasks. They recorded their speaking in voice recorders which had been provided by researchers just before this study began.

Each of the participants received a package of 20 cartoons, but the order of the cartoons was different for each student in order to avoid the effect of each cartoon's difficulty level on participants' performance. For example, the order for Student A consisted of Cartoon 17, Cartoon, 1, Cartoon 8, and continued randomly, whereas the order for Student B's package consisted of Cartoon 2, Cartoon,19, Cartoon 4, and continued with a different random order. By doing tasks in a random order, the process of development of each student's speaking ability could be described authentically.

The following instructions were given in Japanese for the speaking task:

*As an assignment, you are going to tell a story based on a set of three to ten pictures. Each picture has a number. Tell the story according to the numerical order. Before you start a speaking task, you have 30 seconds to study them. You have to make at least three sentences for each picture. Continue talking and finish your story within three minutes. Work on one speaking task per day. Don't do more than one task a day. In addition, do not forget to record your story in the voice recorder.*

The participants were not given sufficient preparation time since they had only 30 seconds to look at the pictures and think of a story. By instructing them to make at least three sentences and continue talking, they were also not given time to think about what they would say, so that the participants were forced to complete the task under pressure (Ellis & Yuan, 2005).

### **Data analysis**

The recorded speaking data was transcribed and analyzed in terms of two components: CAF and creativity. As for CAF, a useful analytical framework (Ozeki & Hiromori, 2016) with the following six items was employed: C1 (Clause/AS-Unit), C2 (Verb tenses and voices), A1 (Error/Tokens), A2 (Error/AS-Unit), F1 (Syllable/Minute), and F2 (Pause/Tokens). In terms of complexity, C1 shows how many clauses there are per AS-Unit and C2 demonstrates how many verb tenses (such as present, past, present progressive) or voices are used. If the numbers are high, it means that participants can produce complex sentences. Regarding accuracy, A1 stands for the rate of errors appearing in tokens while A2 signifies the rate of errors in each AS-Unit. The higher the numbers, the less accurate the sentences are. Regarding fluency, F1 refers to the number of syllables per minute. If this number is high, it means that participants produce sentences fluently. On the other hand, F2 reveals how many pauses in tokens. If participants frequently pause while narrating the stories, the numbers will be high.

Regarding creativity, a rubric for assessing creativity was employed, based on Mozaffari (2013) (see *Table 1*). Although originally designed for creative writing, the speaking tasks in this study share similar task features with creative writing so the rubric was appropriate for adaptation for our analysis of creativity. Each of the rubric's four criteria assesses participants' performance on a four-level scale from excellent (4) to poor (1). Three raters, two Japanese speakers and one Taiwanese speaker who teach English at universities in Japan, evaluated creativity based on the rubric and discussed the differences to reach an agreement. If the ratings were the same among three raters, it was counted as one point. The points of the same ratings were added. The points of the same ratings were divided by the total number of ratings. The inter-rater reliability is 83.17%.

## Results and discussion

This section will discuss the results based on the two research questions.

1. How does learners' speaking performance change over a period of time in terms of complexity, accuracy, and fluency?
2. How is the change in learners' creativity in spoken production related to the changes in CAF?

As for participants' speaking performance in terms of CAF, it fluctuated among participants so it is difficult to tell that the speaking performance improved over time. Moreover, participants' performance varied among each category and they were intertwined (See *Table 2*). Sometimes even though participants performed well in terms of accuracy, his/her performance of fluency and complexity was low. There is also a pattern of which participants achieved high fluency but low accuracy and complexity. With the assumption that participants' speaking performance will improve by task repetition, the order of cartoons were arranged randomly so as to avoid the effect of task difficulty (Robinson, 2001). The results of the development of CAF differed from each other and the results of Student 2, 7, 12, and 15 will be discussed as cases of different patterns. Since the tasks were put in random order, the horizontal axis in each graph shows different tasks chronologically. *Figure 1* describes errors in token while *Figure 2* shows the number of errors in AS-Unit in terms of accuracy. Both of them varied among tasks. These indicate that task difficulty influences participants' accuracy performance. Regarding fluency, *Figure 3* illustrates that participants produced more syllable per minute compared to the beginning although there were ups and downs among tasks; on the other hand, familiarity with the topics swayed participants' fluency performance in terms of pauses in tokens (See *Figure 4*). If participants knew the topic very well, it was not difficult for them to come up with words for production. However, if the topic was rather new or unfamiliar to them, they would need time to search for vocabulary and organize the sentences, which led to lower performances in terms of fluency. In order to probe the details about fluency, analyzing vocabulary and participants' reflection on the speaking tasks will enrich qualitative data in further studies. As for complexity, *Figure 5* and *Figure 6* demonstrated that participants' ability to create complex sentences varied from time to time. DST says that it is essential to analyze learners' performance through process rather than focusing on its product. DST also claims that learners' language development does not just improve over time, and this seems evident in the data because the participants' performance of CAF fluctuated among tasks. So, this is how DST explains this non-linear development.

Table 1 Rubrics for Assessing Creativity in Speaking Tasks

Criteria	4 = Excellent	3 = Good	2 = Fair	1 = Poor
Image	Maximal use of significant details (there is no or just 1 abstraction, generalization, and judgment)	Several use of significant details (they are significantly more than abstractions, generalizations, and judgments)	Minimal use of significant details (they are significantly less than abstractions, generalizations, and judgments)	No use of significant details (sole use of abstractions, generalizations, and judgments)
Characterization	Maximal use of characters' physical appearance, action, thought, symbol, etc. to reveal characters (complete indirect characterization)	Several use of characters' physical appearance, action, thought, symbol, etc. to reveal characters	Minimal use of characters' physical appearance, action, thought, symbol, etc. to reveal characters	No use of characters' physical appearance, action, thought, symbol, etc. to reveal characters (complete direct characterization)
Voice	Maximal use of images to make the voice appealing	Several use of images to make the voice appealing	Minimal use of images to make the voice appealing	No use of images to make the voice appealing
Storyline	Maximal use of narrative to convey purposes	Several use of narrative to convey purposes	Minimal use of narrative to convey purposes	No use of narrative to convey purpose (purpose is conveyed through formal statement)

Note. This table is based on Mozzaffari (2013).

Table 2 Average Levels of Complexity, Accuracy, and Fluency in Speaking Tasks

	C1	C2	A1	A2	F1	F2
Student 1	0.68	4.11	0.06	0.55	49.83	10.83%
Student 2	0.77	5.00	0.04	0.39	72.74	7.38%
Student 3	0.56	4.25	0.04	0.37	68.46	6.89%
Student 4	1.08	5.00	0.02	0.25	98.48	4.24%
Student 5	1.08	4.40	0.03	0.41	75.92	4.44%
Student 6	1.19	4.30	0.04	0.55	77.52	4.57%
Student 7	1.58	5.60	0.01	0.19	95.18	3.98%
Student 8	1.64	5.00	0.02	0.19	62.45	11.41%
Student 9	1.35	5.05	0.01	0.18	86.45	6.07%
Student 10	1.10	4.70	0.01	0.18	65.43	10.47%
Student 11	1.79	5.50	0.01	0.18	87.22	5.10%
Student 12	0.99	4.75	0.03	0.40	85.95	5.25%
Student 13	1.24	4.74	0.02	0.28	74.22	8.95%
Student 14	1.36	4.90	0.02	0.30	82.34	3.89%
Student 15	1.34	5.30	0.02	0.25	77.50	6.10%
Average	1.18	4.84	0.03	0.31	77.31	6.64%

Figure 1. Errors in tokens (A1)

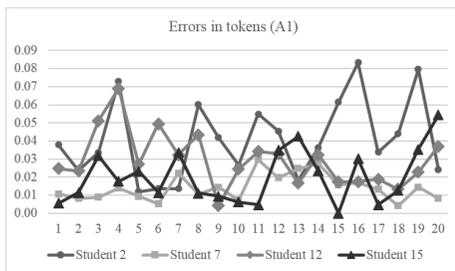


Figure 2. Errors in AS-Unit (A2)

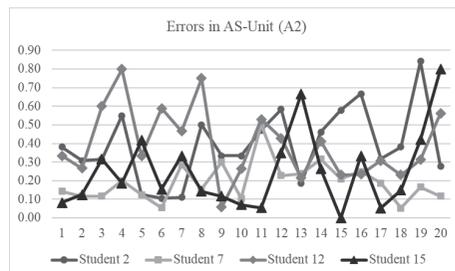


Figure 3. Syllable per minute (F1)

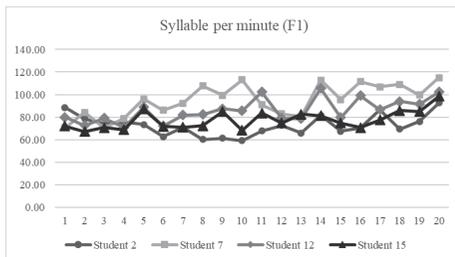


Figure 4. Pause in tokens (F2)

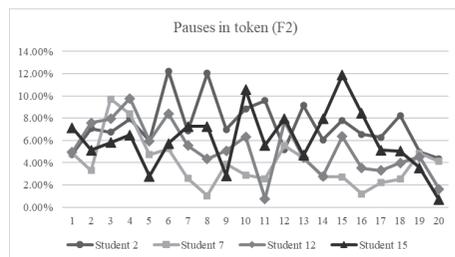


Figure 5. Clause per AS-Unit (C1)

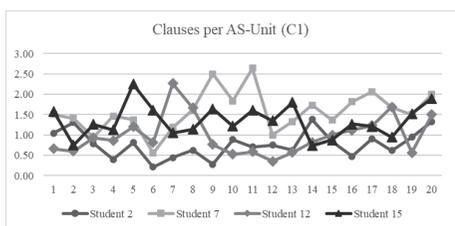
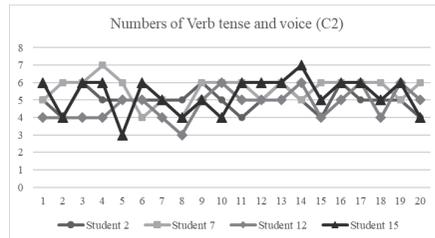


Figure 6. Numbers of Verb tense and Voices (C2)



Regarding the second research question, Table 3 shows the average levels of participants' creativity performance in speaking tasks. On a scale of 0-4, the average levels are 2.94 for Image, 2.96 for Characterization, 3 for Voice, and 2.91 for Storyline. The results indicate that participants' creativity is relatively high, although there are slight deviations among participants. To have a closer look at each category, *Figures 7 to 10* show levels of Image, Characterization, Voice, and Storyline for Student 2, 7, 12, and 15. The results show that in each category the score in task 20, compared to task 1, was either maintained or improved upon, despite occasional fluctuations throughout the process of completing all 20 tasks. What is more, the changes of creativity are similar to CAF, there are variations among participants. For instance, performance of Student 15 descended several times in different tasks. This also explains task difficulty could have influence on creativity sometimes.

Although the performance of CAF was not perfect, the participants' stories can be understood properly. This reveals that even when participants are not good at speaking, they can still express their ideas or opinions through simple sentences. This also means that mistakes are sometimes acceptable when it does not affect listeners' comprehension. For future applications in the classroom, it is necessary for teachers to design detailed rubrics of creativity which fit their task goals or purposes since the general descriptions only show some rough guidelines.

**Table 3 Average Levels of Creativity in Speaking Tasks**

	Image	Characterization	Voice	Storyline
Student 1	2.74	2.11	2.63	2.58
Student 2	3.10	3.20	3.05	3.05
Student 3	3.30	3.20	3.15	3.40
Student 4	3.00	3.05	2.80	2.95
Student 5	2.65	2.55	2.55	2.60
Student 6	3.25	3.35	3.35	3.35
Student 7	3.05	3.15	2.75	2.85
Student 8	2.85	3.15	2.90	2.80
Student 9	3.05	3.05	3.70	3.05
Student 10	2.85	2.90	3.10	2.90
Student 11	3.30	3.35	3.70	3.40
Student 12	3.00	3.10	3.40	3.05
Student 13	3.05	3.05	3.11	2.84
Student 14	2.55	2.70	2.40	2.50
Student 15	2.40	2.50	2.40	2.35
Average	2.94	2.96	3.00	2.91

Figure 7. Image in Creativity

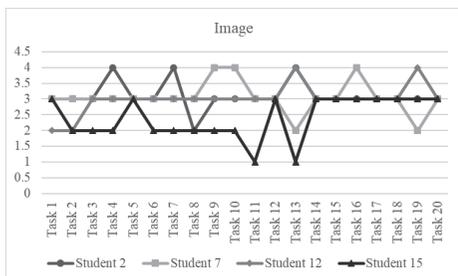


Figure 8. Characterization in Creativity

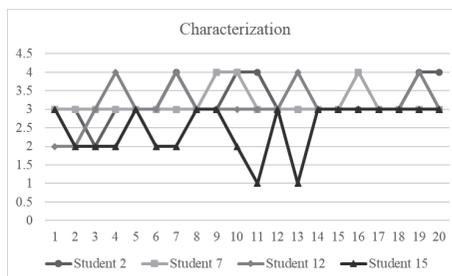


Figure 9. Voice in Creativity

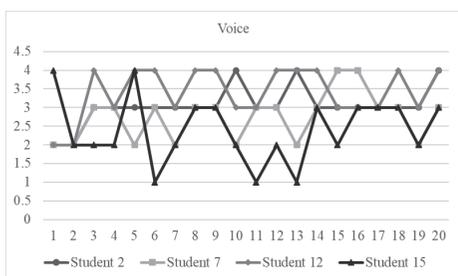
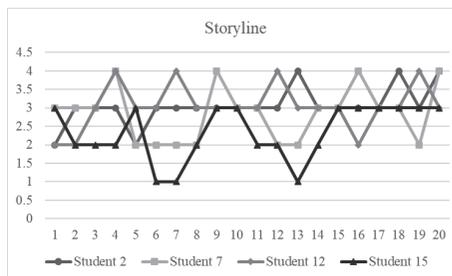


Figure 10. Storyline in Creativity



## Conclusion

This study explored the development of speaking performance of 15 Japanese learners of English based on creativity as well as CAF. The results showed that learners' speaking performances do not necessarily improve linearly, which is a result best understood through the lens of DST. Moreover, even when learners' levels of CAF are not high, they can still convey the stories to some extent. Maybe some learners cannot speak fluently but they still can express their feelings or opinions by using simple sentences. With regard to implication, it is suggested that teachers can give some attention to creativity in addition to linguistic features such as CAF. For future studies, investigating additional features such as vocabulary or learners' reflections on each task may help researchers gain richer data to help deepen their understanding of the influence of DST on L2 spoken production.

## References

- Crookes, G. (1989). Planning and interlanguage variability. *Studies in Second Language Acquisition*, 11, 367-383.
- De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and cognition*, 10(1), 7-21.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford University Press.

- Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167-192). John Benjamins.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21, 354-375.
- Heaton, J. B. (1975). *Composition through pictures (English as a 2nd language book)*. Pearson Education.
- Heaton, J. B. (1997). *Beginning composition through pictures (skills)*. Pearson Japan.
- Hsu, H. C. (2019). The combined effect of task repetition and post-task transcribing on L2 speaking complexity, accuracy, and fluency. *The Language Learning Journal*, 47 (2), 172-187.
- Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction?. *System*, 41 (3), 829-840.
- Larsen-Freeman, D. (2007). On the complementarity of chaos/complexity theory and dynamic systems theory in understanding the second language acquisition process. *Bilingualism: Language and cognition*, 10(1), 35-37.
- Larsen-Freeman, D. (2014). Ten 'Lessons' from Complex Dynamic Systems Theory: What is on Offer. In *Motivational dynamics in language learning* (pp. 11-19). Multilingual Matters.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.
- Lowie, W. (2012). Dynamic Systems Theory Approaches to Second Language Acquisition. In Carol Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd.
- Maryam, V., & Saeed, R. (2018). Development of a rubric for evaluating creative writing: a multi-phase research. *New Writing*, 16(3), 303-317.
- Mozaffari, H. (2013). An analytical rubric for assessing creativity in creative writing. *Theory and Practice in Language Studies*, 12 (3), 2214-2219.
- Ozeki, N., & Hiromori, T. (2016). Study of L2 Speech Production from a Dynamic Systems Approach. *Memoirs of the Institute of Cultural Sciences, Meiji University*, 79, 1-28.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied linguistics*, 22(1), 27-57.
- Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510-532.
- Verspoor, M., de Bot, K., & Lowie, W. (2011). *A dynamic approach to second language development: Methods and techniques*. John Benjamins.
- Watterson, B. (2005). *The Complete Calvin and Hobbes*. Andrews McMeel Publishing.
- Yu, H., & Lowie, W. (2020). Dynamic paths of complexity and accuracy in second language

speech: A longitudinal case study of Chinese learners. *Applied Linguistics*, 41(6), 855-877.

# 日本語学習者の作文における 自動評価と教師評価の比較

Comparative Analysis between Automated Scoring and Human Rating  
for Japanese Essays Written by the Japanese Language Learners

小 森 和 子<sup>1</sup>

KOMORI, Kazuko

伊 集 院 郁 子<sup>2</sup>

IJUIN, Ikuko

李 在 鎬<sup>3</sup>

LEE, Jae-ho

## 1. はじめに

言語知識や言語能力は、第二言語習得において受容と産出の二つの側面から論じられることが多い。受容的な言語知識や言語能力とは、言語使用者が心内に有している言語知識や世界知識を参照し、利用可能な方略を用いながら、言語情報を視覚や聴覚で処理していく知識や能力のことであり、産出的な言語知識や言語能力とは、それらの知識や方略を用いたり、参照できる情報や資料を活用しながら、言語使用者が伝達したい概念を言語化する知識や能力のことである。言語教育の分野では、読む、聞く、書く、話す、の四技能から言語知識や能力が論じられることが多い。これらの言語知識や能力を測定する言語テストは、読む、聞く、などの受容的な側面では、大規模に、効率的に、しかも、高い精度で測れるテストが数多くあるものの、書く、話す、などの産出的な側面は、評価基準が一律でなく、評価者にも訓練が必要であり、さらに、個別評価が必要であることなどから、効率的に評価を行うことは容易でない。そのため、テストにおいては、受容的側面の問題項目が占める割合は多く、それに応じるように、テスト受験者も受容的側面の言語知識や言語能力の習得や学習に力を注ぐことが多い。

こうした中において、共通テストへの「記述」の導入の検討は注目に値するものであった。文科省は、当初、「記述」の導入の経緯を、①自らの力で考え出すことにより、より主体的な思考力・判断力の発揮が期待できる、②文章執筆を通じて、より論理的な思考力・表現力の発揮が期待で

---

1 明治大学国際日本学部

2 東京外国語大学国際日本学研究院

3 早稲田大学日本語教育研究科

きる、③目的に応じた適切な表現様式を用いるなどの表現力の発揮が期待できる、と説明しており、単一の教科の枠を超えた、これまでにない知識や能力の測定になると期待されていた。しかしながら、評価に関する様々な問題提起と議論が起こったことから、再考を余儀なくされ、「記述」の導入は見送られることとなった。

その一方で、英語の TOEFL や留学生向け日本語試験 (EJU) 等においては、すでに実施されている。TOEFL では e-Rater® という自動評価システムで第一次評価を行った後に、人間が評価するという仕組みが採用されている。e-Rater® は情報処理工学や自然言語処理の研究成果に基づいている。自動評価システムは、日本では大規模試験への導入はまだ行われていないが、研究レベルでは成果が発表されている (例えば, Jess, *jWriter*, GoodWriting Rater)。しかしながら、自動評価の研究や実践においては、システムの構築過程、複数のシステムによる評価結果の比較等の研究が主で、システムの評価が人間の評価とどの程度近いのか、どのような点で人間の評価と乖離があるのか等、機械による自動評価と教師による人間評価を丁寧に比較した研究がない。システム開発の過程で、システムに機械学習させる文章がすでに人間の評価得点を付与されたものであるため、システム開発後に人間評価を比較する必要はないのかもしれないが、現場で利用する言語教員や研究者からは、機械による自動評価と自分自身の評価とのずれを感じるのの声も聞かれる。自動評価と人間評価がどの程度整合しているかは、実践の場で自動評価を利用している言語学系の研究者が検討すべき課題なのであろう。

こうした問題意識から、本稿では、次世代の文章評価のあり方を考えるために、自動評価を実践の場にどのように取り入れれば良いかを検討するための第一段階として、自動評価の妥当性と課題について、検討したい。

## 2. 先行研究

本章では、ライティングの自動評価研究の動向を概観した後、日本語教育における自動評価研究の現状と実証的研究を紹介する。

ライティングの自動評価 (自動採点と呼ぶ先行研究も多いが、本稿では自動評価と称す) 研究は、半世紀ほど前の Page (1966) による Project Essay Grade (PEG) が端緒だとする見方が一般的である (例えば、石井・近藤, 2020; Klebanov & Madnani, 2020; 石岡 2007; Burstein, Chodorow, & Leacock, 2004)。Page (1966) にはパイロット的に行った実験結果が述べられているのみで、詳細な研究過程を知ることはできないが、いまなお新鮮で、示唆に富む論考である。Page (1966) は、その冒頭で、(1) 作文の自動評価は必要性が高い、(2) 自動評価は実現可能性が高く将来性に富む、(3) 自動評価が結果として授業の質に目覚ましい発展的影響を及ぼす可能性が見込める、という3点を英語教育者は知るべきだと述べている。このような展望を動機として、Page (1966) は、8年生から12年生の英語母語話者の高校生が書いた276編のエッセイを対象に、4名の英語教師による包括的評価の結果を機械学習させ、32の言語特性を用いた自動評価システムの開発を行った。Page (1966) は、教師評価と自動評価の重相関係数が0.65であっ

たことから、システムの一定程度の有用性を示しているが、その一方で、自動評価のゴールは教育者の評価と高い相関が出るというところまでであり、教育者に代わって評価を行うものではないと記している。Page (1966) が自動評価システムをあくまでも教育支援として位置づけていようとしていたことが窺える。その後、MacDonald, Frase, Gingrich & Keenan (1982) が、ライティングの自動フィードバックツール Writer's Workbench (<https://www.writersworkbench.com/#>) を公開して以降、自動評価やそれに基づく診断的フィードバックツールの研究は盛んになり、現在の Criterion® (<https://criterion.ets.org/criterion/default.aspx>), e-Rater® (<https://www.ets.org/erater/about>) の成果につながった (Burstein et al., 2004)。また、2003年には Criterion® の開発メンバーである ETS 研究員の Burstein らによる『Automated essay scoring: A cross-disciplinary perspective』が刊行され、自動評価研究は、現在では、ETS のようなテスト専門機関だけでなく、個人の研究者も取り組む研究課題となっている。

自動評価研究は 1990 年代以降に本格化した。その社会的背景について、石井・近藤 (2020) は、コンピュータが普及したこと、学生が作文をコンピュータで執筆するようになったこと、および言語テストで「話す」「書く」のパフォーマンス評価に焦点があてられるようになったこと、を挙げている。また、自動評価研究の知見は、これまでのパフォーマンス評価が直面していた問題点、(1) 試験の実施と評価にかかる人的、時間的コスト、(2) 評価の信頼性の担保、(3) 評定者のトレーニングにかかる人的、時間的コスト、(4) トレーニングを受けた評価者の長期的確保、(5) とりわけ「話す」の評価に求められる評定者の高い能力、に対して解決策の一つとなり得るとしている。ただし、同時に重要な問題点にも言及している。それは、コンピュータは評価はしているものの、学習者のパフォーマンスの中身を理解しているわけではないということである。これは、人間が評価する際の評価の観点、例えば、作文や会話の内容、構成、根拠の説得力、独自性、客観性のように、評価者の経験、専門、教育観などによって変わり得る主観的側面を、コンピュータの計算に入れることが困難であることを示唆するものである。また、信頼性の高いシステムの構築には大量のデータが必要だという開発過程に関わる問題点もある。開発を行うには、教師によって既に評価が付与されたデータが、タスクごとに用意されていなければならない。とりわけ、ハイ・ステイクスなテストであれば、公平性の観点からも多様なテストタスクが必要である。また、仮に同一タイプのタスクであったとしても、それを複数用意するためには、タスク間の等質性を保証するための内的一貫性の検証が必須である。こうした検証を経たデータで、かつ、教師評価が付与されているデータを大量にそろえることは簡単ではない。つまり、自動評価には大きな可能性はあるものの、テスト理論に則った精度の高いシステムの完成には課題が多いということである。

一方、日本語教育においては、自動評価研究はまだ緒に就いたばかりである。李 (2021) によると、日本語教育における自動システムの開発は、「リーディングチュウ太」(1999 年公開)、「あすなろ」(2000 年公開) などの読解支援から始まったという。これらのシステムの目的や用途は、学習者自身が理解できない語や句の意味を確認しながら読解が進められるように支援する

ことであった。その後、「*jReadability*」(2013年公開)、「やさしちチェッカー」(2015年公開)、「*jWriter*」(2017年公開)、「*Goodwriting Rater*」(2018年公開)などが開発された。これら2010年代に開発されたシステムは、多様な言語統計量を用いて文章の特徴を捉え、その難易度や熟達度を評価し、ライティング教育や評価の支援のために利用されている。これは、国内外の日本語学習者の多様化と、それに伴って求められる教育支援の在り方の変化と無関係ではなからう。ただし、日本語教育におけるライティングの自動評価システムに限定すると、「*jWriter*」(李ほか2017)、「*Goodwriting Rater*」(田中・坪根2019)しか見当たらない。こうしたわずかな成果について、李(2021)は「これらの自動採点システムは、採点精度が不十分であることや採点の仕組みが難解であることなどの課題はあるものの、言語教育分野におけるイノベーションにつながる」と、その可能性に触れ、「自動採点は、熟練した教師のみができると考えられてきたパフォーマンス評価をコンピュータが担えることを示すものであり、今後、利用可能なデータが増え、多様な計算モデルが競争的に提案されれば、確実に成長できる分野である」として、今後の可能性を示している。

以上のように、英語教育の分野でも、日本語教育の分野でも、ライティングの自動評価研究は、現下に課題は抱えつつも、人工知能研究や情報処理技術の発展と共に、今後の発展的展望が期待できる研究課題として取り組まれている。

その一方で、自動評価が、人が行う評価と同じような結果を表出しているか否かに関する実証的研究は少ない。その中から、本研究の基となった筆者らの成果である小森・李・長谷部・鈴木・伊集院・柳澤(2018)、および日本人学生への自動評価の可能性を検討した影山(2019)を紹介しておく。

小森ほか(2018)では、*jWriter*(詳細は後述)の評価と教師の評価の関係性を検討するべく、日本語の3レベルのクラス、中級( $N=11$ )、中上級( $N=12$ )、上級入門( $N=7$ )の日本語学習者が書いた、比較論証型の意見文30編に対して、教師3名が包括的評価(7段階評価)、および「内容」、「構成」、「言語」、「誤用」の項目別評価(5段階評価)を行った後、*jWriter*で解析し(5段階評価)、両者の評価を比較、分析した。その結果、*jWriter*の評価は、教師による包括的評価と項目別評価の「言語」との間で、正の強い相関が示された。しかし、その他の項目別評価「内容」、「構成」、「誤用」とは無相関であった。

*jWriter*が解析する言語統計量のうち、教師の包括的評価と正の強い相関が認められたのは、漢字、助詞、内容語、異なり語数、延べ語数、和語、中級後半語、総文字数であった。一方、教師評価の「内容」や「構成」と正の強い相関が認められた統計量は少なく、総文字数、内容語であった。また、*e-Rater*®のように、評定の公式に文法的な誤用、語の用法に関する誤用を含む自動評価システムが存在することから、「誤用」の評価についても言語統計量と何らかの関係性が認められると予想していたが、特定の言語統計量との関係性を見出すことはできなかった。

以上の結果から、*jWriter*による自動評価は、教師の包括的評価にも比較的近く、作文の言語的側面の評価にはある程度有用であること、また、内容や構成に対する評価については、自動評

価をどのように活用すべきかが今後の課題であること、が示された。

影山 (2019) は、もう一つのシステム Goodwriting Rater を用いた研究である。この研究の目的は、日本語学習者向けのシステムである Goodwriting Rater が、日本人学生が書いた意見文をどのように評価するか、その結果を日本人学生はどの程度納得するか、さらに、Goodwriting Rater が日本人学生のライティングの評価にも利用可能であるか、を検討することであった。分析の対象となったのは日本人学部生の書いた 107 編の比較論証型の意見文であった。Goodwriting Rater は、ホリスティック評価 (包括的评价) と、「目的・内容」、「構成・結束性」、「日本語」の三つのマルチプルトレイト評価 (項目別評価) を 4 段階で返す仕組みになっている。また、言語統計量などが表示されるテキスト情報機能と、使用したメタ言語部分が機能毎に色別に表示されるメタ言語ハイライト機能を搭載している。

分析の結果、ホリスティック評価においては 70% 以上が、マルチプルトレイト評価の「日本語」では 80% 以上が、最高レベルに評価が集中していることがわかった。ただし、「目的・内容」や「構成・結束性」では、評価がある程度分散しており、日本人学生のライティング能力の差がある程度捉えられることも示された。また、評価結果に対する納得度は、ライティングの評価結果が高い者ほど、納得度も高いことがわかった。ただし、テキスト情報やメタ言語情報などが与えられても、日本人大学生は有効に活用できず、Goodwriting Rater の利用可能性を高めるためには、教師による介入や指導が必要であることも示された。

### 3. 研究課題

自動評価システムの特性を考えても、また、筆者らの行った先般の研究においても、機械による作文評価が言語的側面の評価や判定に対して有用であることは明らかである。しかし、それを実証するためには、作文執筆者の言語知識を測定する言語テストの結果と機械による作文評価の判定結果との間に一定程度以上の相関性があることを確認する必要がある。

また、仮に言語テストの結果が同等にもかかわらず、機械による作文評価に差あったとしたら、その理由を検討しておく必要もある。利用した言語テストの精度が低く、日本語力を識別できていなかったからなのか、機械による作文評価では書き手の言語知識を正しく判定できないからなのか。前者であれば、言語テストの科目に作文を加えることで、より正確に学習者の日本語力を判定できるということが確認できる。一方、後者であれば、機械による作文評価がより正確に言語知識の評価や判定ができるようになるためには、どの点を今後改善すれば良いか、具体的な示唆が得られる。

そこで、本稿では、以下を研究設問として立てることとする。

1. 機械による作文評価の判定と言語テストの得点には、相関性が認められるか、
2. 機械による作文評価の判定は、言語テストのどのような項目の正誤と関連があるか、
3. 言語テストで測れる言語能力において同等と見なされる学習者の作文は、機械による作文評価においても同レベルに評価されるのか。言語テストが同じ得点であるにもかかわらず、機械

による作文評価に差があった場合、それは作文のどのような違いに起因するのか。

## 4. 調査概要

### 4.1. オンラインによる言語テスト

言語テストは、オンライン形式のテストを用いる。これは、明治大学国際日本学部で使用している留学生向け日本語授業用のプレースメント・テストとして開発されたものである。科目は、文字語彙、文法、読解、作文の4科目で、いずれも、オンラインで実施するように開発されたものである。文字語彙、文法、読解は4枝選択式であり、作文はパソコンでタイプする記述式である。システム構築の仕組みやその構築過程、また、テストの項目分析等については、小森（2019）、小森ほか（2018）、柳澤・小森・李・長谷部（2018）、鈴木・柳澤・小森・李・長谷部（2018）、小森・柳澤・安高（2017）にて報告済みであるため、ここでは各科目の概要のみ記す。

まず、文字語彙は、四つの大問から成る。大問1は漢字の読み（12問）、大問2は漢字の表記（12問）、大問3は文脈により規定される語の意味（20問）、大問4は語の用法（6問）で、合計50問である。難易度は大問1から大問3まではN5からN1相当、大問4はN3からN1相当である。

次に、文法は、三つの大問から成る。大問1から3まではほぼ同じ形式で、設問文において、文法的に、または、語用論的に適切な語の活用形、文法形式、表現に関する知識を問う。大問の違いは対象となるレジスターや文法形式の難易度で、大問1は、N5、N4レベル（13問）、大問2はN4からN2レベル（12問）、大問3はN3からN1レベル（13問）で、合計38問である。

読解は、中級以上のレベルの学習者を識別する目的で作成され、五つの大問から成る。いずれの文章も常体で書かれた論説文で、大問1のみ書き下ろしの文章、残りは新書などから引用した生の文章を利用している。それぞれの大問の難易度、文字数、設問数は、大問1はN3レベル（約800文字、設問3題）、大問2と大問3はN3からN2レベル（約900文字、約1,000文字、各設問3題）、大問4と大問5はN2からN1レベル（約1,200文字、約800文字、設問4題、設問3題）である。いずれの大問でも、下線部の箇所に関する局所的な理解問題と、筆者の意見や主張などを文章全体から読み取る統合的な理解を測る問題が含まれており、大問が進むにつれて、後者の統合的な理解力を測る問題の比率が高くなるように作成されている。

最後に、作文は比較論証型の意見文で、プロンプトは「教科書はデジタルブックにするべきだという人がいます。反対に、教科書は今のまま、印刷した本の方がいいという人もいます。あなたはどちらの意見に賛成ですか。理由や根拠を示しながら、自分の意見を書きなさい。」である。辞書等は参照せずに、400字以上を30分の時間制限の中で書くように指示してある。なお、作文科目を受験するのは、文字語彙、文法、読解の合計点で一定以上の者（目安としてはN3程度以上）のみである。

## 4.2. 調査対象者

分析に利用したデータは、2018年3月から2020年9月の間に言語テスト3科目と作文を受験した51名のデータ<sup>4</sup>である。これらの51名は、英語母語話者、あるいは、英語母語話者相当の英語力を有する者である。国籍は、中国 (N=10)、日本 (N=9)、韓国 (N=8)、アメリカ (N=4)、ベトナム (N=4)、イギリス (N=3)、オーストラリア (N=3)、台湾 (N=3)、フランス (N=2)、イタリア (N=1)、オランダ (N=1)、タイ (N=1)、ドイツ (N=1)、ルクセンブルク (N=1) である。このうち、日本国籍者は学校教育を海外や日本のインターナショナル・スクールにおいて英語で受けた者で、日本語力が日本語母語話者レベルに達していないと判定された者である。また、中国や韓国の中には、それぞれの国のインターナショナル・スクールやアメリカやオーストラリアで英語で教育を受けた者も含まれる。

なお、前述のとおり、作文科目の受験には、言語テストの3科目で一定得点 (N3相当) 以上が条件になっている。また、N1の合格証を有する者は本テストの対象外である。したがって、この51名の日本語レベルはN3以上N1未満と考えてよい。

## 4.3. 日本語学習者作文評価システム *jWriter* (<https://jreadability.net/jwriter/>)

作文の評価には、日本語学習者作文評価システム *jWriter* (<https://jreadability.net/jWriter/>) を用いる。*jWriter* は日本語学習者が書いた作文の文章の難易度を、計量テキスト分析の手法を用いて自動評価するシステムである (李・長谷部・村田 2019)。*jWriter* によって得られる言語統計量は、総文字数、文数、段落数、延べ語数、異なり語数、平均文長 (一文の平均延べ語数)、語種別語彙数、難易度別語彙数、タイプトークン比 (語の多様性を示す指標、以下、TTR)、接続表現数等の45種<sup>5</sup>である。そのうち、平均文長、総文字数、動詞数、難易度別語彙数、TTR、和語数、漢語数によって作文能力値が推定される。推定の公式は、「 $1.637 + \text{平均文長} \times 0.045 + \text{中級後半語} \times 0.021 + \text{TTR} \times -0.430 + \text{動詞} \times 0.015 + \text{中級前半語} \times 0.011 + \text{総文字数} \times -0.004 + \text{和語} \times 0.007 + \text{漢語} \times 0.007$ 」 ( $R^2=0.760$ ) である。この公式は、「多言語母語の日本語学習者横断コーパス (International Corpus of Japanese as a Second Language: I-JAS)」にある12言語を母語とする日本語学習者 (初級、中級、上級の3レベル) 373名<sup>6</sup>のエッセイを基準データとして用い、

4 データの使用については、テスト実施の際にオンライン上で示した同意書によって許諾を得ている。

5 *jWriter* で表示される統計量は全部で45種類あるが、名称は異なるが同じ数値になる統計量がある。具体的には、総形態素数 = 延べ語数、長文力 = 平均文長、普通名詞 = 名詞である。また、レベル付き語数 = 初級前半語 + 初級後半語 + 中級前半語 + 中級後半語 + 上級前半語 + 上級後半語、であり、接続表現数には接続詞の数も含まれる。また、難解語使用率、接続表現使用率、その他の品詞、漢語力、定型句は、*jWriter* の定義に基づいて独自に算出される統計量である。

そこで、本稿では、45種類のうち、総形態素数、長文力、普通名詞、レベル付き語数、接続詞、難解語使用率、接続表現使用率、その他の品詞、漢語力、定型句の10種を除く35種類の統計量を参照することとする。

6 373名の母語と人数の内訳は、中国語 (N=82)、韓国語 (N=64)、スペイン語 (N=33)、インドネシア語 (N=33)、ベトナム語 (N=31)、英語 (N=28)、ドイツ語 (N=24)、ロシア語 (N=22)、ハンガリー語 (N=21)、タイ語 (N=16)、フランス語 (N=10)、トルコ語 (N=9) である。

重回帰分析によって得られたモデルに基づいている (Lee&Hasebe 2020)。I-JAS に関する詳細はここでは割愛するが、エッセイは「私たちは日常生活で、ファースト・フードと家庭でゆっくり味わう手作りの料理を食べています。ファースト・フードと家庭料理を比較し、それぞれの良い点や悪い点などを説明して、「食生活」についての意見を 600 字程度で書いてください。」というプロンプトに対して執筆された意見文である。詳細は迫田・石川・李 (2020)、迫田・小西・佐々木・須賀・細井 (2016) を参照されたい。

また、*jWriter* では、作文能力の評定値のほか、評定値を基に付与される作文評価・ガイドラインという 5 段階評価 (入門, 初級, 中級, 上級, 超級), テキスト情報 (文字数, 文数, 段落数), 語彙情報 (語の難易度を色別に表示), 接続表現情報 (機能毎に色別に表示), 語彙リスト一覧 (レベル表示, 辞書表示等) などの情報も、視覚的に捉えやすいデザインで、学習者にフィードバックできるように設計されている。

なお、*jWriter* では、解析のための最低文字数が 300 に設定されている。これは、上記の基準データが最低 300 文字であったことによる (Lee&Hasebe 2020)。ただし、精度の高い解析として 600 字前後が推奨されている。

## 5. 調査結果

### 5.1. 言語テストと作文評定値の概要

言語テスト 3 科目を、1 問 1 点で採点したところ (表 1)、平均は文字語彙 34.93 点 (50 点満点)、文法 25.90 点 (38 点満点)、読解 8.86 点 (16 点満点)、3 科目合計 69.71 点 (104 点満点) であった。また、*jWriter* で作文の評定値を求めたところ、平均 2.37 であった。

表 1 3 科目の得点と作文の結果

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>N</i>
文字語彙	34.93	9.73	17	50	51
文法	25.90	7.82	15	38	51
読解	8.86	4.01	0	16	51
3科目合計	69.71	20.14	39	104	51
作文	2.37	0.51	1.08	3.47	51

次に、3 科目の得点と作文評価値との相関の有無を確認したところ、いずれも正の相関が認められ [文字語彙  $r=.619$ ,  $df=51$ ,  $p<.001$ , 文法  $r=.581$ ,  $df=51$ ,  $p<.001$ , 読解  $r=.468$ ,  $df=51$ ,  $p<.001$ , 合計  $r=.618$ ,  $df=51$ ,  $p<.001$ ]、特に、文字語彙との相関が強いことがわかった。これは、*jWriter* の判定が文字や語を中心とした統計量に基づいていることによるものと考えられる。

また、*jWriter* が判定した評価ガイドライン (以下、作文レベル) では、初級 4 名、中級 27 名、

上級 20 名であったため、初級の 4 名を除いた中級と上級の 3 科目の得点差を確認したところ (表 2)、いずれの科目においても、中級より上級の方が得点が高く、その差は統計的に有意であった。

表 2 中級と上級の 3 科目の得点

	中級 (N=27)	上級 (N=20)	合計 (N=47)	t 値, 効果量 Pearson's r
文字語彙	33.22 (8.15)	39.70 (9.80)	35.98 (9.36)	$t(45)=2.472^*$ , $r=.350$
文法	24.37 (6.83)	29.75 (7.80)	26.66 (7.66)	$t(45)=2.516^*$ , $r=.350$
読解	7.96 (3.74)	10.75 (3.89)	9.15 (4.01)	$t(45)=2.484^*$ , $r=.350$
合計	65.56 (16.32)	80.20 (20.65)	71.79 (19.50)	$t(47)=2.716^*$ , $r=.380$

注 1: 効果量 Pearson's r は水本・竹内 (2010) に基づき算出

注 2: \* $p<.05$

以上のことから、jWriter による自動評価は、文字語彙、文法、読解等の総合的な言語知識や能力の推定に有用であることが示唆される。

## 5.2. 言語テストに見られる作文レベル中級と上級の違い

本節では、課題 2 を検討するために、作文レベルの中級と上級の間で、どのような問題項目で差が出ているのか、具体的に見ていく。文字語彙、文法、読解の全 104 項目について、正答者数の差が有意か否かをカイ二乗検定で確認したところ、文字語彙で 9 問、文法で 9 問、読解 (2 種類の読解文章) で 4 問の計 22 問で有意であった (表 3)。

表 3 正答者数の差が有意な問題項目一覧

科目	測定知識	対象箇所	対象箇所の日本語レベル	中級 (N=27)		上級 (N=20)		$\chi^2$ 値 (df=1)	効果量 Cramer's V
				正答者数 (率)	誤答者数 (率)	正答者数 (率)	誤答者数 (率)		
文字語彙	漢字読み	図書館	初級後半	19 (70.37%)	8 (29.63%)	19 (95.00%)	1 (5.00%)	4.502*	0.31
	漢字読み	血液型	中級前半	19 (70.37%)	8 (29.63%)	19 (95.00%)	1 (5.00%)	4.502*	0.31
	漢字読み	現象	中級後半	12 (44.44%)	15 (55.56%)	16 (80.00%)	4 (20.00%)	6.031*	0.37
	漢字表記	調べる	初級後半	22 (81.48%)	5 (18.52%)	20 (100.00%)	0 (0.00%)	4.145*	0.30
	漢字表記	支える	中級後半	19 (70.37%)	8 (29.63%)	19 (95.00%)	1 (5.00%)	4.502*	0.31
	意味	みにくい(争い)	中級後半	8 (29.63%)	19 (70.37%)	13 (65.00%)	7 (35.00%)	5.815*	0.36
	意味	(合併を)見送る	中級前半	4 (14.81%)	23 (85.09%)	11 (55.00%)	9 (45.00%)	8.599**	0.43
	用法	禁物	上級前半	7 (25.93%)	20 (74.07%)	12 (60.00%)	8 (40.00%)	5.539*	0.34
	用法	思い当たる	上級前半	10 (37.04%)	17 (62.96%)	14 (70.00%)	6 (30.00%)	4.996*	0.33
	授受	(〜)もらいました	初級後半	27 (100.00%)	0 (0.00%)	17 (85.00%)	3 (15.00%)	4.326*	0.30
文法	仮定条件	行くら	初級後半	15 (55.56%)	12 (44.44%)	18 (90.00%)	2 (10.00%)	6.518*	0.37
	複合辞	として	中級前半	22 (81.48%)	5 (18.52%)	20 (100.00%)	0 (0.00%)	4.145*	0.30
	体言化	の	初級後半	16 (59.26%)	11 (40.74%)	18 (90.00%)	2 (10.00%)	5.426*	0.34
	複合辞	とは限らない	中級後半	15 (55.56%)	12 (44.44%)	19 (95.00%)	1 (5.00%)	8.934**	0.44
	複合辞	つつある	中級後半	14 (51.85%)	13 (48.15%)	16 (80.00%)	4 (20.00%)	3.943*	0.29
	複合辞	ようでは	上級前半	4 (14.81%)	23 (85.09%)	10 (50.00%)	10 (50.00%)	6.801**	0.38
	複合辞	限りでは	中級後半	10 (37.04%)	17 (62.92%)	16 (80.00%)	4 (20.00%)	8.580**	0.43
	複合辞	ざるを得ない	上級前半	11 (40.74%)	16 (59.26%)	15 (75.00%)	5 (25.00%)	5.456*	0.34
	1,000字読解文	統合理解	上級前半	15 (55.56%)	12 (44.44%)	17 (85.00%)	3 (15.00%)	4.584*	0.31
読解	1,000字読解文	局所理解	上級前半	16 (59.26%)	11 (40.74%)	19 (95.00%)	1 (5.00%)	7.719**	0.41
	1,000字読解文	統合理解	上級前半	11 (40.74%)	16 (59.26%)	15 (75.00%)	5 (25.00%)	5.456*	0.34
	1,200字読解文	統合理解	中級後半	9 (33.33%)	18 (66.67%)	15 (75.00%)	5 (25.00%)	7.982**	0.41

注 1: 対象箇所の日本語レベルは日本語教育語彙表、機能語用例文データベースはごろも、jReadability による

注 2: 効果量 Cramer's V は水本・竹内 (2010) に基づき算出

注 3: \* $p<.05$ , \*\* $p<.01$

このうち、文字語彙の9問と文法の9問に出題されている語や表現の難易度を、それぞれ日本語教育語彙表 (<https://jreadability.net/jev/>) と機能語用例データベースはごろも (<https://www.hagoromo-text.work/>) で確認したところ、初級後半レベルが3問(文字語彙2問, 文法1問), 中級前半レベルが4問(文字語彙2問, 文法2問), 中級後半レベル8問(文字語彙4問, 文法4問), 上級前半レベルが3問(文字語彙1問, 文法2問)であった。また、読解の2種類の読解文章の難易度を *jReadability* (<https://jreadability.net/sys/ja>) で確認したところ、一方が中級後半レベル、他方が上級前半レベルであることがわかった。以上のことから、中級前半レベルから上級前半レベルの難易度の問題項目に正答するか否かが、作文レベルの中級と上級を弁別する要因となっていると言えそうである。

これを確認するために、特に通過率(全受験者に占める正答者の割合)において、作文レベルの中級と上級の差が大きかった問題項目(文字語彙2問, 文法2問)を以下に例示する。下線付きの選択肢が正答であり、その語彙レベルは( )に示した通りである。

・文字語彙1 (中級後半レベル: 作文レベル中級通過率 29.63%, 同上級通過率 65.00%)

「親の財産をめぐって、兄弟の間で( )争いが続いている。」

- (1) にくい (2) みにくい (3) ずるい (4) ずるがしこい

・文字語彙2 (中級前半レベル: 作文レベル中級通過率 14.81%, 同上級通過率 55.00%)

「A社とB社は、交渉がうまくいなくて、予定していた合併を( )ことにした。」

- (1) 見送る (2) 振り返る (3) 思いきる (4) 聞き入れる

・文法1 (上級前半レベル: 作文レベル中級通過率 14.81%, 同上級通過率 50.00%)

「締め切り直前になってテーマを変える( ), いい論文は書けないだろう。」

- (1) ならでは (2) ままでは (3) ようでは (4) ものでは

・文法2 (中級後半レベル: 作文レベル中級通過率 37.04%, 同上級通過率 80.00%)

「わたしが調べた( ), 工場の機械に問題はなかった。」

- (1) 限りでは (2) 次第では (3) うえでは (4) ようでは

文字語彙の2問は、いずれも学習者が初出で学ぶ基本義ではなく、派生義の知識を問う問題である。デジタル大辞泉で意味を確認してみると、文字語彙1の「みにくい」の基本義は<①姿かたちがよくない>で、外見上の不具合を指す。しかし、ここで問われているのは、ある状況や様子について、<②見て不快な感じがする。嫌な気持ちがある。見苦しい>であり、基本的な意味

ではない。また、文字語彙2の「見送る」の基本義は、<①遠ざかる人が見えなくなるまで後方から目で追う>で、この意味は比較的早い段階で学習され、かつ、日常生活においても使用頻度の高い意味で、学習者にも定着が良い。しかし、ここで問われている意味は<⑦今、事を行うのは不利と考えて、そのままにしておく>で、意味範囲が限られた派生義である。このような派生義を問う問題項目で差が出ているということは、それが両者の語彙知識の質的な差であること、また、そのような差があることが作文の語の産出にも反映され、結果として作文評定値の差となっている可能性がある。

また、文法の2問は複合辞の問題で、論理関係や接続関係の理解や、表現形式が含意する書き手の意図の正しい解釈ができなければ、正答できないような問題である。例えば、「よう」は、前件によって後件のような悪い結果が起こることを予測する際に用いられる用法で、前件の命題内容に対する書き手（話し手）の非難や批判が含意されている。また、「限る」はリソースを限定することで、後件の事象の解釈が限定的であることを示す場合に用いられる。いずれも、初出時に学んだ意味用法ではなく、難易度が高い。

一方、読解においては、紙幅の都合で詳細な検討はできないが、二つの読解文章の問題項目において、作文レベルの中級と上級の差が認められた。一つ目の文章は、大問3の1,000字程度の論説文で、人間の認知と記憶に関する文章である。*jReadability*では上級前半レベルと判定された。設問は三つあるが、いずれもその差が有意であった。具体的には、問1が「筆者が認知について勉強して知ったこと」、問2が「下線部の説明として正しいもの」、問3が「筆者の考えと合っているもの」、である。問1と問3は文章全体を読んだ上で、文章の情報を統合しなければ、正答することが難しい。二つ目の文章は、大問4の1,200字程度の論説文で、登校拒否に関する内容である。*jReadability*では文章の難易度が中級後半レベルと判定された。設問は四つあるが、そのうち作文レベルの中級と上級に差が有意であったのは、「筆者は登校拒否をどのようなものだと考えているか」という問題で、やはり、統合的な理解を問う設問であった。大問4のその他の設問は、文章の各段落の中に解答につながる情報が明示的に記されており、段落単位で文章が理解できれば、正答にたどり着ける問題で、それらにおいては差がなかった。

つまり、読解については、作文レベルの中級と上級の差は、文章の難易度が中級後半レベルから上級前半レベルであり、設問が文章全体を統合的に理解していなければ正答できない問題に正答するか否かであると考えられる。

### 5.3. 作文レベル中級と上級における言語統計量の違い

次に、*jWriter*が算出する作文評定値とその他の言語統計量を全対象者51名について確認したところ、表4の通りとなった。これを見ると、30分という時間制限の中で、学習者は平均400文字、約250語、10文程度で作文を執筆していること、作文評価レベルは中級後半レベル（2.5以上で上級判定）であること、がわかる。

表4 作文の言語統計量

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>N</i>
総文字数	412.88	141.14	110	731	51
総段落数	3.65	2.234	1	14	51
総文数	11.2	3.677	3	19	51
延べ語数	253.8	88.10	73	476	51
異なり語数	104.06	25.80	49	153	51
タイプトークン比	0.14	0.35	0	1	51
平均文長	23	4.72	12	34	51
名詞	63.82	24.23	13	123	51
動詞	15.33	6.89	4	32	51
形容詞	5.49	2.64	1	11	51
形状詞	3.16	1.96	0	8	51
連体詞	1.9	1.63	0	7	51
副詞	4.49	2.59	0	11	51
感動詞	0.04	0.28	0	2	51
助詞	73.29	26.66	23	127	51
助動詞	25.51	10.06	7	52	51
代名詞	3.96	2.59	0	12	51
固有名詞	0.24	0.65	0	4	51
接続表現数	2.86	1.83	0	8	51
初級前半語	30.9	12.77	8	66	51
初級後半語	27.65	11.16	6	53	51
中級前半語	34.86	13.27	11	66	51
中級後半語	18.82	10.64	4	43	51
上級前半語	2.14	2.71	0	12	51
上級後半語	0.1	0.36	0	2	51
ひらがな	222.82	74.33	65	391	51
カタカナ	48.49	23.31	5	95	51
漢字	109.43	43.97	30	216	51
和語	166.65	58.88	49	307	51
漢語	47.71	18.93	17	99	51
外来語	12.73	6.20	2	25	51
混種語	1.14	1.28	0	6	51
内容語	84.65	30.43	20	150	51
リーダビリティレベル	3.14	0.63	2	4	51
作文評価・レベル	2.37	0.51	1.08	3.47	51

次に、初級判定者4名を除く作文レベル中級と上級における、言語統計量の値とその差を *t* 検定で確認したところ、14の統計量においてその差が有意であった(表5)。差があった統計量は、*jWriter* の作文評定値の公式にある総文字数、平均文長、中級前半語、中級後半語、和語、漢語を中心に、既知語彙数に関わる延べ語数と異なり語数、文の必須要素となる名詞、動詞、助詞、内容語、初級前半語等において差が認められた。

表5 中級・上級の作文の言語統計量

	中級 (N=27)		上級 (N=20)		t値, 効果量Person's $r$
	M	SD	M	SD	
総文字数	371.78	109.22	495.25	136.14	$t(45)=-3.45^{**}$ , $r=.46$
総段落数	4.19	2.66	3.15	1.50	$t(45)=1.564$ , $n.s.$ , $r=.23$
総文数	11.00	3.44	11.55	3.73	$t(45)=-0.522$ , $n.s.$ , $r=.08$
延べ語数	226.78	66.96	306.05	86.75	$t(45)=-3.538^{**}$ , $r=.47$
異なり語数	96.44	19.96	121.00	22.10	$t(45)=-3.984^{**}$ , $r=.51$
タイプトークン比	0.11	0.32	0.10	0.31	$t(45)=0.12$ , $n.s.$ , $r=.20$
平均文長	21.11	3.48	26.80	3.32	$t(45)=-5.652^{**}$ , $r=.64$
名詞	56.67	18.61	77.80	24.75	$t(45)=-3.345^{**}$ , $r=.45$
動詞	13.48	5.84	18.85	7.11	$t(45)=-2.84^{**}$ , $r=.39$
形容詞	5.37	2.54	6.10	2.69	$t(45)=-0.948$ , $n.s.$ , $r=.14$
形状詞	2.78	1.74	3.95	2.16	$t(45)=-2.059^*$ , $r=.29$
連体詞	1.56	1.25	2.50	2.01	$t(45)=-1.98$ , $n.s.$ , $r=.28$
副詞	4.30	2.57	5.00	2.68	$t(45)=-0.912$ , $n.s.$ , $r=.13$
感動詞	0.07	0.39	0.00	0.00	$t(45)=0.858$ , $n.s.$ , $r=.13$
助詞	64.07	20.22	90.60	25.43	$t(45)=-3.984^{**}$ , $r=.51$
助動詞	24.48	9.87	27.95	9.57	$t(45)=-1.206$ , $n.s.$ , $r=.18$
代名詞	3.48	1.78	4.30	2.77	$t(45)=-1.23$ , $n.s.$ , $r=.18$
固有名詞	0.07	0.27	0.50	0.95	$t(45)=-2.231^*$ , $r=.32$
接続表現数	2.67	1.80	3.25	1.92	$t(45)=-1.07$ , $n.s.$ , $r=.16$
初級前半語	28.19	10.66	36.10	13.86	$t(45)=-2.215^*$ , $r=.31$
初級後半語	26.19	11.13	31.10	10.70	$t(45)=-1.521$ , $n.s.$ , $r=.22$
中級前半語	31.33	10.58	42.55	12.92	$t(45)=-3.271^{**}$ , $r=.44$
中級後半語	13.63	6.53	28.05	9.10	$t(45)=-6.332^{**}$ , $r=.69$
上級前半語	1.37	1.98	3.20	3.27	$t(45)=-2.38^*$ , $r=.33$
上級後半語	0.11	0.42	0.05	0.22	$t(45)=0.586$ , $n.s.$ , $r=.09$
ひらがな	201.41	57.00	266.70	71.06	$t(45)=-3.495^{**}$ , $r=.46$
カタカナ	45.59	21.39	55.20	22.71	$t(45)=-1.483$ , $n.s.$ , $r=.22$
漢字	93.63	33.60	138.40	42.61	$t(45)=-4.028^{**}$ , $r=.51$
和語	148.93	46.17	201.80	56.59	$t(45)=-3.526^{**}$ , $r=.47$
漢語	41.19	13.82	59.95	19.11	$t(45)=-3.911^{**}$ , $r=.50$
外来語	11.63	5.62	14.95	6.10	$t(45)=-1.931$ , $n.s.$ , $r=.28$
混種語	1.04	1.06	1.35	1.60	$t(45)=-0.808$ , $n.s.$ , $r=.12$
内容語	75.52	23.77	102.75	30.00	$t(45)=-3.473^{**}$ , $r=.46$
リーダビリティレベル	3.41	0.50	2.70	0.47	$t(45)=4.913^{**}$ , $r=.59$
作文評価・レベル	2.15	0.22	2.87	0.24	$t(45)=-10.604^{**}$ , $r=.85$

注1: 効果量 Person's  $r$ は水本・竹内(2010)に基づき算出注2: \* $p<.05$ , \*\* $p<.01$ 

ただし、作文の良し悪しと無関係でない、接続表現（接続詞を含む）、形容詞、連体詞、副詞、代名詞、助動詞等においては差が認められなかった。接続表現は作文の段落間の関係を示すディスプレイコース・マーカーとして用いられるため、適切で効果的な使用が読み手の理解を助ける。また、形容詞、連体詞、副詞は、名詞や動詞のあり様を、詳細に、正確に、豊かに表す。さらに、代名詞は結束性を高める上で有効に用いられ得る品詞である。ヴォイスやモダリティに関わる助動詞

は、書き手の視点や態度が現れる重要な言語形式である。

しかしながら、今回の分析では、中級と上級でこれらの言語形式において差がなかった。この要因の一つには、*jWriter*の評定値推定の公式において、これらの言語統計量が含まれていないことが考えられる。しかし、それ以上に関わっている要因としては、今回は400字程度で比較的短く、また、人数も20名、27名と少なかったことから、分析にかけられる絶対的な言語産出量が限定的であったため、各品詞の出現数も少なく、結果として、差が認められなかったのではないかと考える。この結果だけから、自動評価では作文の構成や日本語の豊かさに直接的に関わる言語形式の差を反映しない、と軽々に結論付けられない。

さらに、量的には差が認められなかった言語形式は、作文の中でどのように用いられているかが、評価の高低の指標になる可能性がある。自動評価が示す量的な差だけでは認められない、言語使用における質的な側面の評価には、経験を積んだ教師による判定、特に、ルーブリック等を用いながら、評価の観点ごとに診断的評価を行う必要があろう。これについては、次節の事例研究で検討していく。

## 5.4. 個別分析

### 5.4.1. 言語テストの結果と作文評定値の関係

本節では、課題3を検討するために、言語テストで同程度の得点をとった学習者の作文が、*jWriter*でも同じ判定を付与されているか確認する。そのために、言語テストを5点刻みで分割し、作文レベル中級と上級の人数を数えたところ、表6の通りとなった。

表6 言語テストの得点幅別に見る中級・上級判定者数

	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99	100-104	合計
中級	4	0	4	3	3	2	1	5	1	1	2	1	0	27
上級	0	1	0	5	2	1	0	0	0	1	1	5	4	20
合計	4	1	4	8	5	3	1	5	1	2	3	6	4	47

一部例外はあるものの、大まかな傾向としては、90点未満には作文レベルで中級判定が多く、90点以上になると上級判定が多くなる。また、全くの同点で作文レベルが中級と上級に分かれたのは、58点、59点、63点、97点、であった。このうち、58点と59点は、中級(N=3)の総文字数がM=402.67、平均文長がM=16.75であったのに対して、上級(N=3)は総文字数がM=539.33、平均文長がM=24.35と、長さの違いが作文レベルの判定に特に影響していることが確認できた。また、63点では、平均文長においては、中級(N=1)が23.78、上級(N=1)が24.30と同程度であったが、総文字数が中級が361、上級が415で差があった。また、中級後半レベル以上の難易度の語の使用数を見ると、中級が7語、上級が21語で、難易度の高い語の産出数が大きく異なっていることがわかった。

一方、97点で、作文レベルが中級（N=1）と上級（N=1）に別れた点については、その理由を精査する必要がある。95点以上の10名の中で唯一中級に判定されたのが、この97点の1名である。104点満点の言語テストで97点であることから、言語知識は十分にあると言える。それにもかかわらず、作文レベルでは中級判定となったのはなぜか。

#### 5.4.2. 同点の二作文の比較

本節では、言語テストは同点の97点であったが、作文レベルでは、一方が中級、他方が上級に判定された2名（以下、中級判定者と上級判定者）について、それぞれの作文を見ながら、どのような点で作文レベルに差が生じたのか、検討を試みる。

##### 5.4.2.1. 言語統計量に基づく二作文の特徴比較

まず、取り上げるそれぞれの作文、および言語統計量（表7）を以下に示す。

#### 【中級判定者の作文】

教科書は今のまま、印刷した本を使うべきだ。なぜなら、子供やその保護者に負担がかかるからだ。

教科書をデジタルブックにするメリットは確かにある。印刷されている教科書は厚く、重量があるため、子供に持たせると重い。デジタルにしまえば何冊でも一つの端末に収めることができ、便利になる。だが、デジタル化してしまえば、教科書が重いという負担よりも、子供の視力に多大な負荷がかかると考える。デジタルブックにするということは、子供は液晶画面の端末を本の代わりに使用することになります。そうすると、長時間画面を見つめていた目は疲れ、視力も低下するだろう。

また、教科書をデジタルブックにした場合、端末を買うにはかなりのお金が必要になると思われる。お金がかかるということは、保護者の負担が増えるということだ。その上、端末等の機械は壊れやすいため、何度も購入を繰り返さなければならない可能性が高い。

これらのデメリットを踏まえて、教科書はデジタルブックではなく、印刷されたものを使うべきである。

#### 【上級判定者の作文】

私は教科書をデジタルブックにするべきだと思う。

その最大の理由は紙で作られた教科書を何冊も持ち運ぶことは小学生はもちろんのこと、中学生、高校生、大学生にとっても心身の疲れにつながり、勉学に支障をきたすからだ。特に通勤通学ラッシュで混みあう中、学校まで電車やバスなどの公共機関を利用している学生にとっては

なおさらである。実際私も小学生の時から電車通学であったが、最初のうちは教科書がたくさん入ったランドセルをしょって長時間、電車で通学するだけで疲れてしまい、授業中に眠くなってしまうことも多々あった。しかしながら、デジタルブックならば、そんな必要もない。なぜならば、デジタルブックは一台の電子機器の中に複数の教科書の情報を搭載することができる上に、軽量で、厚さも薄く、持ち運びにも適しているからだ。

もう一つの理由は紙媒体の教科書で学ぶのと同じように、デジタルブックでも学ぶことができるという点である。デジタルブックにはブックマークやマーカー機能、書き込むためのタッチペンをも利用することもできるので、紙で作られた教科書に怠らない機能を体験できるはずだ。

以上の二点の理由から、私は教科書はデジタルブックにするべきだと思う。

表7 中級判定者の作文と上級判定者の作文の言語統計量

	中級	上級	中級<上級	中級M	上級M
総文字数	439	511	○	371.78	495.25
総段落数	4	4	=	4.19	3.15
総文数	11	9	×	11.00	11.55
延べ語数	276	308	○	226.78	306.05
異なり語数	107	131	○	96.44	121.00
タイプトークン比	0.39	0.43	○	0.11	0.10
平均文長	25.09	34.22	○	21.11	26.80
名詞	67	88	○	56.67	77.80
動詞	17	19	○	13.48	18.85
形容詞	2	7	○	5.37	6.10
形状詞	4	1	×	2.78	3.95
連体詞	2	3	○	1.56	2.50
副詞	2	7	○	4.30	5.00
感動詞	0	0	=	0.07	0.00
助詞	75	93	○	64.07	90.60
助動詞	25	22	×	24.48	27.95
代名詞	2	3	○	3.48	4.30
固有名詞	0	0	=	0.07	0.50
接続表現数	5	5	=	2.67	3.25
初級前半語	36	36	=	28.19	36.10
初級後半語	17	27	○	26.19	31.10
中級前半語	44	38	×	31.33	42.55
中級後半語	23	38	○	13.63	28.05
上級前半語	9	5	×	1.37	3.20
上級後半語	0	0	=	0.11	0.05
ひらがな	239	260	○	201.41	266.70
カタカナ	45	65	○	45.59	55.20
漢字	122	152	○	93.63	138.40
和語	183	188	○	148.93	201.80
漢語	50	72	○	41.19	59.95
外来語	12	19	○	11.63	14.95
混種語	1	2	○	1.04	1.35
内容語	89	111	○	75.52	102.75
リーダビリティレベル	3.16	2.08	○	3.41	2.70
作文評価レベル	2.49	3.47	○	2.15	2.87

注1: リーダビリティレベルは値が小さいほど難易度が高い

注2: 中級&lt;上級は○, 中級=上級は=, 中級&gt;上級は×

言語統計量(表7)の網掛けの項目は作文評定値推定の公式に含まれる変数である。そのうち、下線を付した総文字数とTTRは公式の係数がマイナスの変数である。簡便な比較のために、表7の中に中級判定者の作文と上級判定者の作文の値の大小を、○、=、×で示す。さらに、中級全体(N=27)と上級全体(N=20)の平均値Mも併記しておく。

作文評定値は、中級判定者の作文は2.49、上級判定者は3.47で、その差はおおよそ1ポイント

トある。*jWriter*では、2.50以上が上級判定、3.50以上が超級判定になることから、この2例は境界値に近い作文だと言える。2例を直接比較すると、多くの言語統計量で上級判定者の方が値が大きいの。特に、総文字数や平均文長等の作文の長さに関わる指標、延べ語数と異なり語数等の語彙知識の量的側面に関わる指標、漢字や漢語等の漢字知識に関わる指標、名詞と動詞等の内容語数に関わる指標において、上級判定者の値が大きいの。この結果から、上級判定者は中級判定者に比べて、漢字をより多く用い、さまざまな種類の語を使って、より長い文章を記述している、ということが言えよう。

#### 5.4.2.2. 語彙的複雑性から見た二作文の傾向

次に、作文や会話で産出された文章の評価に影響する言語的特徴CALFの観点から考察を行う。CALFの要素は、統語的複雑性(Syntactic Complexity)、語彙的複雑性(Lexical Complexity)、正確性(Accuracy)、流暢性(Fluency)から成り(小島・金田2020:33-34)、言語学習者の習熟度やパフォーマンスの評価に有用な言語的側面であると言われている(Housen & Kuiken, 2009)。本稿では、*jWriter*で解析される統計量を用いて、語彙的複雑性について考察を試みる。

語彙的複雑性は、語彙の豊かさを示す指標で、その下位要素として、多様性、洗練性、および語彙密度がある(Wolfe-Quintero, Inagaki & Kim 1998; Bulté & Housen, 2012)。多様性には、*jWriter*でも判定されるTTRがよく用いられる。ただし、「TTRはテキストが長くなるほど異語数の増加が緩やかになることでスコアが減少するという性質がある」(小島・金田2020:37)という指摘もある。しかしながら、李(2020)が*jWriter*を用いて行ったIJASの作文データの分析結果を見ると、TTRは言語習熟度と正の相関が認められていることから、本研究ではTTRを多様性の指標として用いる。洗練性は、「基本語彙以外の語の割合に基づくlexical frequency profileが代表的で、大規模コーパスをもとにした語の頻度情報を利用して算出される」(ibid.)が、基本語の定義は研究者によって異なり、頻度情報はコーパスのレジスターによって異なるため、洗練性の定義は一概に定まらない。そこで、本研究では、レベル付き語彙数全体に占める中級後半以上の難易度の語彙を、洗練性の指標とすることとする。これは、レベル付き語彙は日本語学習者が学ぶべき語としての指標であるが、初級から中級前半レベルの語は汎用性の高い基本的な語で、使用頻度も高いが、中級後半以上のレベルの語は抽象度が上がり、学術的文書や文学作品に利用されるような語が中心になることに基づく。語彙密度は、内容語の総語数に対する割合で、「情報が密なテキストほど高くなると言われている」(小島・金田2020:38)が、機能語が十分に習得できていない学習者の方が内容語の占める割合が高くなることから、一定レベル以上にのみ適用するべきという指摘もある。今回の二例の作文執筆者の言語テストの得点は満点に近く、機能語の習得は十分であると言えることから、総語数に対する内容語の割合を語彙密度の指標として用いる。

語彙的複雑性に関するこれら三つの要素、多様性、洗練性、語彙密度について、言語統計量の

値を用いて算出したところ（表8）、いずれの値においても、上級判定者の値が高いことがわかった。

表8 中級判定者の作文と上級判定者の作文の語彙的複雑性

	多様性	洗練性		語彙密度
		異なり語数	延べ語数	
中級	0.39	0.18	0.12	0.32
上級	0.43	0.39	0.26	0.36

なお、特に差が大きい洗練性については、レベル付き語彙のそれぞれのレベルの語彙の数を示しておく（表9、ただし上級後半レベルの語彙は0だったため省略）。なお、*jWriter*が返す言語統計量（前掲表7）では、レベル付き語彙数は延べ語数でカウントされたものであり、それを見ると、中級判定者の作文の方が数が多いレベルの語彙もあるが、異なり語数で数えると、中級判定者の作文の方が多いのは、中級前半レベルの語彙のみである。レベル付き語彙数については、延べ語数だけでなく、異なり語数でも情報提供されるように、*jWriter*が改良されることが望ましい。

表9 レベル付き語彙の差

	初級前半		初級後半		中級前半		中級後半		上級前半	
	異なり語数	延べ語数								
中級	21	36	9	17	23	44	14	23	5	9
上級	23	36	16	27	19	38	27	38	5	5

以上の結果を総合すると、言語テストが同点であったこの作文執筆者2名は、多枝選択式の言語テストで測られる言語の受容的知識には差はないものの、産出的知識に量的な差があり、それが作文の*jWriter*による判定により可視化されていると言える。

#### 5.4.2.3. 二作文に対するルーブリック評価

次に、作文の質的側面について検討するために、ルーブリック評価を行うこととした。ルーブリックとは、「外国語学習者の学習成果を、大規模な客観的テストではなく、パフォーマンスを通して到達度を評価する際に用いられる評価規準」（野口・大隅 2012: 357）のことで、「評価の観点と到達度をマトリックスの形で組み合わせ、到達度に関しては順序尺度で表される段階と、文章で記述される具体的パフォーマンスが、一体になって表される点に特徴がある」（Ibid.）と言われている。作文のルーブリック評価はこれまでにいくつか公開されているが、今回は伊集院・高野（2020）のルーブリック（以下、伊集院ルーブリック）を用いることとする。伊集院ルーブリックは、教師の経験則によって作成されたものではなく、大学教員が実際に作文評価を行い、評価の過程で各作文に付した評価コメントを分析し、大学教員に内在している作文評価の観点を抽出した上で作成されたもので、実証的な裏付けに基づいた信頼性の高い評価指標の一つある（開発の経緯に関する詳細は、Ijuin 2020, 伊集院・李・小森・野口 2020, 伊集院・小森・奥切 2018

を参照のこと、指標の妥当性検証に関する詳細は、伊集院・李・小森・高野・野口2021を参照のこと。

伊集院ルーブリックの評価の観点、内容構成、言語表現、形式、その他、の四つで、内容構成と言語表現には下位観点がある。各観点とその説明は以下の通りである（表10）。各観点は5段階評価（すばらしい5点、よい4点、ふつう3点、もう一歩2点、がんばれ1点）である。

表10 伊集院ルーブリック

内容構成	主張／主張（課題に対する書き手の回答）が容易に伝わるか。
	根拠／主張を支える根拠に説得力があるか。
	論理構成／全体の論理構成が明確で、段落間、段落内の情報が整理されているか。
言語表現	正確さ／日本語表現（文型、文法、語彙、表記）に誤用がないか。
	適切さ／アカデミック・ライティングにふさわしい表現（「ふつう体」、「かたい表現」、多様な表現）が適切に用いられているか。
形式	書式／書式など（フォーマット、文字数、原稿用紙の使い方など）の指定が守られているか。
その他	加点対象／独創的な視点が盛り込まれていて読み手の興味を引くなど、書き手の意欲や工夫が伝わる。
	減点対象／タイトルが本文の内容と合っていない、手書きの場合に極端に読みづらいなど、読み手を混乱させる要因がある。

表11 作文例のルーブリック評価結果

	中級		上級	
	評価者1	評価者2	評価者1	評価者2
	内容構成・主張	4	4	4
内容構成・根拠	3	3	4	4
内容構成・論理構成	3	3	4	4
言語表現・正確さ	5	5	4	5
言語表現・適切さ	4	4	5	5
形式	5	5	5	5
加点対象	—	—	—	—
減点対象	—	—	—	—
合計	24	24	26	27

伊集院ルーブリックを使用して、筆者のうちの2名で評価を行ったところ（表11）、2名の評価はほぼ一致し、中級判定者より上級判定者の方が合計点が高かった。つまり、総合的な評価においては、教師評価は自動評価と一致していると言えよう。さらに、ルーブリックの観点ごとに見てみると、【内容構成・主張】と【形式】では中級判定者の作文と上級判定者作文に差はないが、【内容構成・根拠】、【内容構成・論理構成】、【言語表現・適切さ】で差があった。【内容構成・主張】と【形式】で差が出なかったのは、中級判定者も上級判定者も、言語テストが満点に近い上級レベルの日本語学習者で、主張を明確に示すという意見文執筆に関する最低限の作文技術を身につけていたためであろう。また、今回はパソコンで入力する形式であったため、形式には差が出なかったと考えられる。

次に、それぞれの作文の構成と内容を概念的に捉えやすくするために、伊集院・高橋（2012）や伊集院（2017）等を参照しながら各文の機能（主張、根拠等）をコーディングした。その上で、一文毎に、判定した文の機能、文末表現、および談話展開を表示する接続表現やメタ言語表現等を太字にして転載し、段落ごと□□で括ったものが以下の通りである。

#### 中級判定者の作文

主張（紙の本）べきだ。なぜなら、根拠の要約（デジタルブックの欠点の要約）。

譲歩（デジタルブックの利点）。だが、根拠1（デジタルブックの欠点；視力への負担）と考  
える。根拠1の背景情報。そうすると、根拠1の言い換え。

また、根拠2（デジタルブックの欠点；経済的負担）と思われる。根拠2の言い換え。その上、  
根拠2の補足。

これらのデメリットを踏まえて、主張（紙の本）べきである。

#### 上級判定者の作文

主張（デジタルブック）べきだと思う。

その最大の理由は根拠1（紙の本の欠点；持ち運びの不便性による勉学への負の影響）。特に  
根拠1の補足。実際根拠1の補足。しかしながら、根拠2の要点（デジタルブックの利点；持ち  
運びの利便性）。なぜならば、根拠2。

もう一つの理由は、根拠3（デジタルブックの利点；機能性）。根拠3の補足はずだ。

以上二点の理由から、主張（デジタルブック）べきだと思う。

まず、【内容構成・根拠】については、中級判定者の作文では、主張が紙の本であるにも関わらず、根拠として挙がっているのがデジタルブックの欠点のみで、紙の本を支持する積極的根拠が記述されていない。また、デジタルブックの欠点については、視力への負担という根拠1と、経済的負担という根拠2が挙げられているが、根拠の言い換えが多く、論が深まっていないため、根拠としての説得力にやや欠ける。

一方、上級判定者の作文では、主張がデジタルブックで、その根拠に紙の本の欠点を挙げた上で、紙の欠点が解消されるデジタルブックの利点と紙と同等の機能性、の二つを根拠として提示している。また、「特に」、「実際」、「なぜならば」等のメタ言語表現を上手に使うことで、根拠の情報を補足している。

このように、主張をサポートするのにふさわしい根拠が、適切な言語形式を用いて示されているか否か、という点で差が認められる。

次に、【内容構成・論理構成】については、中級判定者の作文は段落内の論理関係が整合的でない箇所がある。具体的には、「教科書は今のまま、印刷した本を使うべきだ。なぜなら、子供やその保護者に負担がかかるからだ。」と、冒頭で紙の本を主張しているが、次に「なぜなら」で根拠が続くような書きぶりになっているにもかかわらず、「負担がかかる」と何の負担か不明な文が続く。おそらく、デジタルブックの欠点を総合して一言で述べれば、「負担がかかる」ということなのだろう。しかし、人間の自然な思考では、紙の本を主張した後、「なぜなら」があれば、紙の本を支持する根拠の出現を予測する。それにもかかわらず、読み手の思考に逆らって、紙の本の利点ではなく、デジタルブックの欠点を挙げているという点で、整合性に欠ける。また、「なぜなら」の後には、理由の説明に関する記述の展開を期待するが、ここでは理由の要点のみがあいまいに示されており、前後関係を示すマーカーとして効果的に使われていない。さらに、次の段落の冒頭では、デジタルブックの利点を譲歩として提示しているが、後続するデジタルブックの欠点まで読まないで、ここが譲歩として機能していることが理解できず、唐突感を与えている。このように、中級作文は前後の結束性が低く、段落内、段落間の整合性がつかみにくい。

それに対して、上級判定者の作文の段落内は同一命題でまとめられており、整合的である。また、根拠1で紙の本の欠点を示した後、「しかしながら」と逆接表現を冒頭に置き、紙の欠点に対立するデジタルブックの利点が記述されるであろうことを、読み手に予想させた上で、実際にデジタルブックの利点の一つである持ち運びの利便性が記述されている。また、根拠2のデジタルブックの携帯性という利点に関する記述の箇所では、中級判定者の作文と同様に、「なぜなら(ば)」が使われているが、上級判定者の作文では、「なぜならば」以降にデジタルブックの携帯性の良さが具体的に説明されている。さらに、根拠3では、「もう一つの理由は」から始めて、持ち運びとは異なる視点の根拠が展開されることを読み手に予測させる。

このように、中級判定者の作文は段落内のまとまりと整合性に不備があるのに対して、上級判定者の作文は段落内の論点が一貫している。特に、接続表現やメタ言語表現を、読み手の論理的思考に即して、適切な箇所適切に使用することができており、それによって段落間や文間の論理関係を明確にして、論が展開できている。使用されている接続表現数が同じでも、その使い方という質的な側面で両者に差が出ている。この点については、*jWriter*のような自動評価が表出する言語統計量では評価できない点であり、教師によって判断、評価されるべき側面であろう。

三つ目の【言語表現・適切さ】については、文体の統一ができていないか否かが、評価を決定づけた。中級判定者の作文では、常体の中に敬体が含まれているが、上級判定者の作文では文体は

常体に統一されている。また、アカデミック・ライティングにふさわしくない表現がどのように含まれているかも、適切さの評価に影響を与えた。例えば、上級判定者の作文でも、「ランドセルを背負って」とすべきところ、「ランドセルをしょって」とあるが、それ以外には特に問題となる表現は認められない。一方、中級判定者の作文では、「お金がかかる」「お金が必要だ」とやや口語的な表現が複数ある。「負担」という語が数か所で使用されていることを考慮すると、「経済的な負担が大きい」のように表現するべきであっただろう。また、両者とも「デジタルブックの容量の大きさ」に関して言及している箇所があるが、上級判定者の作文では「デジタルブックは一台の電子機器の中に複数の教科書の情報を搭載することができる」と記述されているのに対して、中級判定者の作文では「印刷されている教科書は（中略）デジタルにしまえば何冊でも一つの端末に収めることができ」とある。「複数の教科書」と述べている上級判定者の作文に対して、中級判定者の作文では「教科書は何冊でも」と述べられていたり、上級判定者の作文では「一台の電子機器の中に搭載する」としているところを、中級判定者の作文では「一つの端末に収める」となっている等、比較してみると、表現の洗練性に差があると言える。このように、対応する命題ごとに二つの作文を比較することで、どこが、どのように、洗練されているか、適切であるか、を確認することができる。この点も、自動評価で計算される語彙的複雑性では説明できない、教師評価の強みであろう。

## 6. おわりに

本研究では、機械による作文評価の言語的側面の評価における有用性と現状の課題を探ることで、自動評価と人間評価のそれぞれの強みと適切な併用のあり方を検討することを目的に、三つの課題を設定し、検討を行った。

課題1「機械による作文評価の判定と言語テストの得点には、相関性が認められるか」については、言語テストの3科目、文字語彙、文法、読解のいずれとも相関が確認された。また、課題2「機械による作文評価の判定は、言語テストのどのような項目の正誤と関連があるか」については、*jWriter*による中級と上級の間で統計的に有意な差がある項目を確認したところ、言語形式の派生的意味や用法を問うものであることがわかった。以上の課題1と2の結果から、*jWriter*による作文評価が、言語知識の量的側面と質的側面のいずれにおいても、一定程度の推定力を有することが確認できたと言える。

課題3「言語テストで測れる言語能力において同等と見なされる学習者の作文は、自動評価においても同レベルに評価されるのか。言語テストが同じ得点であるにもかかわらず、自動評価に差があった場合、それは作文のどのような違いに起因するのか」については、*jWriter*による中級と上級の間で、言語統計量の多くに差があることが確認された。特に、作文の長さに関わる指標、語彙知識の量的側面に関わる指標、漢字に関わる指標、内容語数に関わる指標において、上級の方が値が大きいことがわかった。また、語彙的複雑性の指標を導入して分析した結果、その下位要素である、多様性、洗練性、および語彙密度のいずれにおいても、上級の方が値が大きかっ

た。jWriterの言語統計量を使って追加分析することで、言語的側面の差をより詳細に評価できると考えられる。さらに、言語テストが同点の2名の作文について、実例を見ながら考察した結果、同じ数の接続表現が用いられていても、論理的なつながりや整合性を高めるような適切な使い方がされているか否か、という点において、中級と上級には差があった。このように、数が同じであっても、どのような表現が、どこで、どのように使われているのか、それが適切であるか、有効であるか、という評価は、現状では自動評価では対応できない。この点にこそ、教師による評価が有用であり、教師評価において特化されるべき評価の観点だと言えよう。

以上の分析から、自動評価の出す言語統計量は、作文の書き手の言語知識の量的側面、質的側面を推定するのに有用である一方で、言語統計量が示す数の多寡では説明されない、使用の適切性や有用性については教師が評価する、という役割分担が示唆される。こうした役割分担の提言の裏付けを、実証的に示すことができたことが、本研究の成果である。

その一方で、さらに検討すべき点も明確になった。まず、作文の言語的側面においては、量的な評価では自動評価が、質的な評価では教師評価が有用であることは確認されたが、具体的に質的側面のどのような点を教師が評価すべきかについては、今回は、語彙的複雑性、接続表現と文章の機能や構成との関係、接続表現の使われ方、について検討するに留まった。今後は、教師でしか評価できない、言語使用の適切性や有用性について、どのような言語形式を取り上げるべきか、また、そのためには具体的にどのような方法で適切性や有用性を評価すれば良いか等、検討しなければならない。特に、作文の書き手の態度を示すモダリティ形式等の文末表現、結束性や一貫性の指標となる指示詞や代名詞、作文の構成と内容の関係を示して読み手の理解を助けるメタ言語表現は、作文評価に関わるため、これらについては、作文での実際の使われ方とそれに対して教師がどう評価するか、検討していきたい。さらに、今回分析した語彙的複雑性のように、自動評価の評定値推定公式に搭載が可能な要素もあることから、語彙的複雑性以外の、統語的複雑性、正確性、流暢性のCALF全体を取り上げて、作文の包括的評価に対してどの程度の説明力を有するのかを、検討を試みたい。その上で、自動評価の評定値推定公式に含めるべき要素を再考し、自動評価の精度向上に貢献できるような知見を提示していきたい。

今後はこれらの課題に取り組みながら、日本語の文章評価において機械に評価がゆだねられる点をさらに明確化し、機械の精度向上に示唆を示しながら、機械と人間がどのようにそれぞれの役割を担っていくべきか、それを今後の評価実践に活かしていけるか、検討を続けていきたい。

## 付記

本研究は、科研費基盤研究(C)『評価者・学習者・意見文の分析に基づく評価の検証：作文の自律学習支援を目指して』(研究代表者:伊集院郁子, 課題番号:18K00680, 2017～2021年度)、科研費基盤研究(B)『論理的文章作成を支援するウェブ型自動処理システムの開発と普及に関する研究』(研究代表者:李在鎬, 課題番号:19H01273, 2019～2023年度)、および科研費挑戦的研究(萌芽)『リーダービリティシステムを活用した日本語テスト問題の自動生成に関する研

究』(研究代表者:李在鎬, 課題番号:19K21637, 2019~2022年度)の助成を得て行われたものです。

また,本研究のデータ入手やデータ整理においては,本学部の柳澤絵美先生,およびテストシステムの開発と管理を担当する(株)ピコラボの鈴木泰山氏にご尽力いただきました。ここに記して感謝申し上げます。

#### 参考文献

- Blute, Brian & Housen, Alex (2012) Defining and Operationalising L2 Complexity. In Housen, Alex, Kuiken, Folkert, & Vedder, Ineke (eds.) *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*, 21-46. Philadelphia: John Benjamins.
- Burstein, Jill & Chodorow, Martin (2004) Automated Essay Evaluation: The *Criterion* Online Writing Service, *AI Magazine*, 25(3), 27-36.
- Housen, Alex & Kuiken, Folkert (2009) Complexity, Accuracy and Fluency in Second Language Acquisition. *Applied Linguistics* 30(4), 461-473.
- Ijuin, Ikuko (2020) Developing an Essay-Writing Rubric for Learners of Japanese: Based on the Analysis of Writing Assessment by Japanese University Faculty, *Learner Corpus Studies in Asia and the World*, 5, 79-940. (<http://www.lib.kobe-u.ac.jp/repository/81012491.pdf>)
- 伊集院郁子 (2017) 「作文と評価:日本語教育的観点から見たよい文章」李在鎬(編)『文章を科学する』38-57. 東京:ひつじ書房.
- 伊集院郁子・小森和子・奥切恵 (2018) 「大学教員によるライティング評価の観点を探る」『Learner corpus studies in Asia and the world』Vol. 3, 159-176.  
([http://www.lib.kobe-u.ac.jp/infolib/meta\\_pub/G0000003kernel\\_81010125](http://www.lib.kobe-u.ac.jp/infolib/meta_pub/G0000003kernel_81010125))
- 伊集院郁子・高橋圭子 (2012) 「日本・韓国・台湾の大学生による日本語意見文の構造的特徴-『主張』に着目して-」『東京外国語大学国際日本研究センター日本語・日本学研究』2, 1-16.
- 伊集院郁子・李在鎬・小森和子・野口裕之. 2020. 「評価コメントに見られる意見文評価の様相-共起ネットワーク及びコレスポネンス分析に基づく考察-」『第二言語としての日本語の習得研究』23, 26-43. 東京:凡人社.
- 伊集院郁子・高野愛子. 2020. 『日本語を学ぶ人のためのアカデミック・ライティング講座』東京:アスク出版.
- 石井雄隆・近藤悠介 (2020) 「自動採点研究とは?」石井雄隆・近藤悠介(編)『英語教育における自動採点-現状と課題』, 1-15. 東京:ひつじ書房.
- 石岡恒憲 (2007) 「コンピュータによるエッセイ, 小論文の自動採点について」『教育テスト研究センター第3回研究会報告書』1-8.
- 影山陽子 (2019) 「作文の自動評価システムの日本人学部大学生への活用可能性-評価への納得度と推敲への動機に着目して-」『アカデミック・ジャパニーズ・ジャーナル』11, 28-36.
- Klebanov, B. Beigman & Madnani, Nitin (2020) Automated Evaluation of Writing - 50 Years

and Counting, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7796-7810.

- 小島ますみ・金田拓 (2020) 「ライティング評価と言語的指標の関係—メタ分析による研究成果の統合」  
石井雄隆・近藤悠介 (編) 『英語教育における自動採点—現状と課題』, 33-72. 東京: ひつじ書房.
- 小森和子 (2019) 「日本語オンライン・プレースメント・テストの分析—紙版テストとの比較を通して—」『明治大学国際日本学研究』10(1), 1-28.
- 小森和子・李在鎬・長谷部陽一郎・鈴木泰山・伊集院郁子・柳澤絵美 (2018) 「教師による評価とコンピュータによる自動評価はどの程度一致するのか—中上級日本語学習者の意見文の評価を対象に—」『2018年度日本語教育学会秋季大会予稿集』278-283.
- 小森和子・柳澤絵美・安高紀子 (2017) 「日本語プレースメント・テストの開発と問題項目の分析—国際日本学部のET日本語科目における試み—」『明治大学国際日本学研究』9(1), 31-62.
- 李在鎬 (2021) 「書くことを支援する自動評価システム『jWriter』(特集 AIやICTが変える言語教育)」『日本語学 (2021 冬号)』40(4), 42-51.
- Lee, Jae-Ho, & Hasebe, Yoichiro (2020) Quantitative Analysis of JFL Learners' Writing Abilities and the Development of a Computational System to Estimate Writing Proficiency. *Learner Corpus Studies in Asia and the World*, 5, 105-120. (<http://www.lib.kobe-u.ac.jp/repository/81012493.pdf>)
- 李在鎬・長谷部陽一郎・村田裕美子 (2019) 「学習者作文の習熟度に関する自動判定と Web システムの開発について」, 李在鎬 (編) 『ICT×日本語教育』38-53. 東京: くろしお出版.
- 李在鎬・長谷部陽一郎・迫田久美子 (2017) 「人工知能の仕組みを利用した学習者作文評価システム『jWriter』—I-JASを利用した試み—」『2017年度日本語教育学会秋季大会予稿集』289-291.
- 水本篤・竹内理 (2010) 「効果量と検定力分析入門—統計的検定を正しく使うために—」『外国語教育メディア学会 (LET) 関西支部 メソドロジー研究部会 2010 年度報告論集: より良い外国語教育研究のための方法』47-73.
- 野口裕之・大隅敦子 (2014) 『テストングの基礎理論』東京: 研究社.
- Page, B. Ellis (1966) The Imminence of ... Grading Essays by Computer, *The Phi Delta Kappa*, 47(5), 238-243.
- 迫田久美子・石川慎一郎・李在鎬 (編) (2020) 『日本語学習者コーパス I-JAS 入門 研究・教育にどう使うか』東京: くろしお出版.
- 迫田久美子・小西円・佐々木藍子・須賀和香子・細井陽子 (2016) 「多言語母語の日本語学習者横断コーパス International Corpus of Japanese as a Second Language」『国語研プロジェクトレビュー』6(3), 93-110.
- Shermis, D. Mark & Burstein, C. Jill (2003) *Automated Essay Scoring: A Cross-disciplinary Perspective*, NY: Routledge.
- Wolfe-Quintero, Kate, Inagaki, Shunji, & Kim, Hae-Young (1998) *Second Language Development*

*in Writing: Measures of Fluency, Accuracy, & Complexity*. University of Hawai'i Press.

#### インターネット

Criterion® < <https://criterion.ets.org/criterion/default.aspx> > (最終閲覧日: 2021 年 10 月 12 日)

e-Rater® < <https://www.ets.org/erater/about> > (最終閲覧日: 2021 年 10 月 12 日)

GoodWriting Rater 読み手と構成を意識した日本語ライティング < <https://goodwriting.jp/wp/> > (最終閲覧日: 2021 年 10 月 21 日)

Jess® 日本語小論文 評価採点システム < <http://tk2-203-11024.vs.sakura.ne.jp/jess/> > (最終閲覧日: 2021 年 10 月 12 日)

*jReadability* 日本語文章難易度判別システム < [https://jreadability.net/sys/terms\\_of\\_use?lang=ja](https://jreadability.net/sys/terms_of_use?lang=ja) > (最終閲覧日: 2021 年 10 月 21 日)

*jWriter* 学習者作文評価システム < <https://jreadability.net/jWriter/> > (最終閲覧日: 2021 年 10 月 12 日)

Writer's Workbench < <https://www.writersworkbench.com/#> > (最終閲覧日: 2021 年 10 月 12 日)



## 【研究論文】

## まちづくり政策の現状と課題

— 一川越市の事例を中心に —

Current Status and Issues of Urban Regeneration Policies:  
on the Case of Kawagoe-City白戸 伸 一  
SHIRATO, Shin-ichi

## はじめに

この半世紀の間に我が国の小売業は顕著な構造変化を遂げている。それに対応して、流通政策にも著しい変化が生じている。小稿では、このような変化を数量的に確認しつつ、その変化に対してどのような流通政策が採られてきたのか、そして政策対応がまちづくりという包括的社会経済政策の中に位置づけられながら、今後どのような方向に向かうべきなのかを検討したい。

21世紀初頭までの流通政策については、通産・石原（2011）で詳細な検討がされている。また、渡辺（2014）・（2016）、畠山（2017）・（2020）などでは2014年以降のまちづくり政策と地域商業政策の関連についても言及されており、小稿では地域経済の再生に対する商業の役割をあらためて検討することとする。

## 1. まちづくり政策の開始

流通分野では図1にみられるように、1980年代から21世紀初頭にかけて大きな構造変化が進

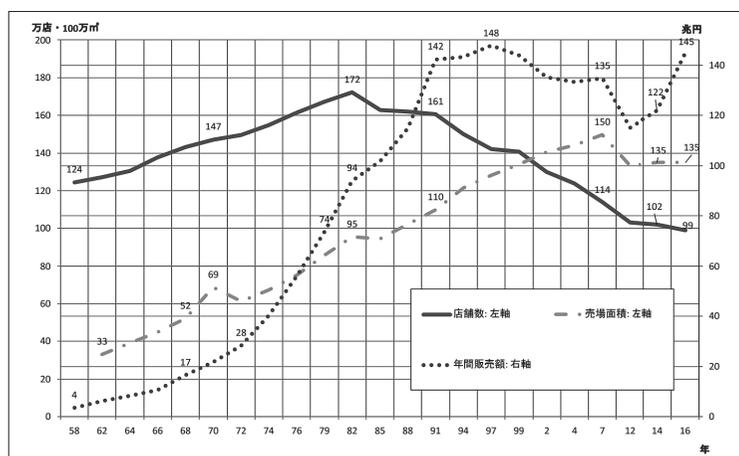


図1 小売業の構造変化

(各年「商業統計表」・2016年「経済センサス」の数値を使用)

行した。高度経済成長期以降右肩上りで小売店舗数も増加し1982年には172万店に達したが、それ以降は減少に転じ2016年には99万店へと減少している。これは60年代初頭には9割前後を占めていた従業員数5人未満の零細経営が、1982年には145万店（84%）へ、

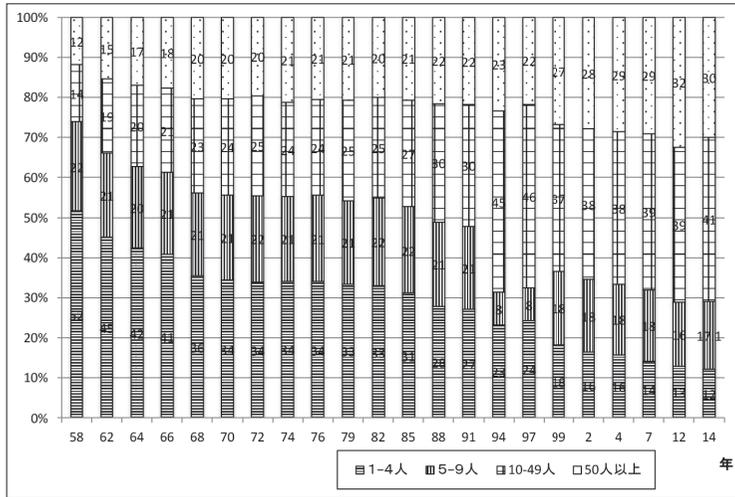


図2 従業員数別販売額構成の推移  
(各年「商業統計表」の数値より算出)

そして2007年には76万店へと半減したためであり、構成比も66%にまで低下している。地方の商店街が活気をなくし、中にはシャッター通りに化したところもある。

しかし店舗面積は21世紀初頭まで伸び続けており、東日本大震災や集計方法の変更で若干低下が見られるものの、さほど減少していない。また販

売額も、バブル崩壊後の伸び悩みやリーマンショック、東日本大震災等による低下も見られるが、その後の販売額の増加はやはり店舗数の減少とは異なる動向を示している。そして、今世紀初頭にやっと事業所構成比を1%台まで高めた従業員数50人以上の大規模経営が、販売額では30%を占めるようになり(図2)、全国にチェーン展開している大規模小売企業が小売市場への影響力を強めている。

このような大規模経営・大型店が拡大傾向を辿る要因のひとつは、1989年日米構造協議以降小売業の分野でも規制緩和が進められ、大型店出店を「調整」していた大規模小売店舗法が廃止されまちづくり3法(大規模小売店舗立地法・中心市街地活性化法・改正都市計画法)による環境やまちづくりを重視した政策へと転換したためである。そのため環境規制でさほど出店が難しい郊外へ大型店の出店が続き、地方都市中心部の衰退傾向が問題視されるようになった。本来、中心市街地活性化法(以下中活法と略記)は中心市街地の活性化を目的に制定されていたが、中核となるべきTMO(まちづくり組織)が十分に機能しなかったことや大型店の郊外出店が横行されたため、政策の修正がおこなわれた<sup>1</sup>。具体的には、2006年の改正都市計画法で延床面積1万㎡超の大規模集客施設の立地を近隣商業、商業、準工業地域のみ限定し、同年の改正中活法で、都市機能を中心部に集約し、まちなか居住を促進することによりコンパクトシティを実現することを目指した。小売業は、都市住民や近接住民の生活を支える重要な都市機能のひとつとして位置づけられた。

このような変更の背後には、それまでのまちづくり3法に基づく政策があまり成果を上げてないという厳しい評価があった。それに対し、第一に旧中活法では商業振興に重心があり「生活空間」再生施策が乏しいという問題があったので、都市機能集積による居住者の利便性を図る、第二に中小小売商業高度化事業策定など商業振興を重視したTMOではなく、まちづくりという点

1 TMOの問題点については、拙稿(白戸(2020))でも触れておいた。

で商業者以外の事業者や地権者、住民も参加して協議・推進するためにまちづくり協議会を設置し、これを協議・推進機関とした、第三に「選択と集中」を掲げて内閣府に中活本部が置かれ、「中心市街地活性化基本計画」（以下では中活計画と略記）を内閣総理大臣認定に格上げした、第四に内閣府中心市街地活性化担当室が「中心市街地活性化基本計画認定申請マニュアル」を作成し、認定計画のフォローアップを求めるとともに、通行量、居住人口、施設入込数、販売額、公共交通利用者数、空き店舗数等をカテゴリ別に数値目標として設定し、達成度の評価に客観性を持たせるといった商業重視からの政策変更と厳格さが求められるようになっていく。

さらに、流通政策を管轄する行政機関の変化も関係している。小売業の振興や商業集積に関する政策については、もっぱら通産省の管轄であったが、まちづくり3法の制定以降、小売商業を都市機能の一端を担うものとして「まちづくり」という視点から扱うことになり、例えば中心市街地活性化法（以下中活法と略記）が13省庁共管であったように省庁横断的な政策対象となった<sup>2</sup>。2001年中央省庁再編により内閣府が設置され、2002～2006年の都市再生特別措置法、地域再生法、改正都市計画法、改正中活法の制定において内閣府の関与が積極化してくると、内閣府、国交省、経産省がまちづくりや地方都市の再生に連携して取り組むようになった。そして、「市街地の整備改善」関連は国交省、「都市福祉施設の整備」関連は国交省・厚労省、「街なか居住の推進」関連は国交省、「経済活力の向上」関連は経産省、「公共交通機関、特定事業等」関連は国交省がそれぞれ主管している<sup>3</sup>。

これらに加えて、21世紀初頭に明らかになってきた人口減少による経済的停滞と地方の衰退傾向が強く認識されるようになったことも重要である。図3に示されているように、2008年をピークに人口減少が始まり2050年には1億人を割り込む見通しが示された。2014年に日本創成会議は、2040歳女性人口が5割以上減少する「消滅可能性都市」が896自治体（地方公共団体）

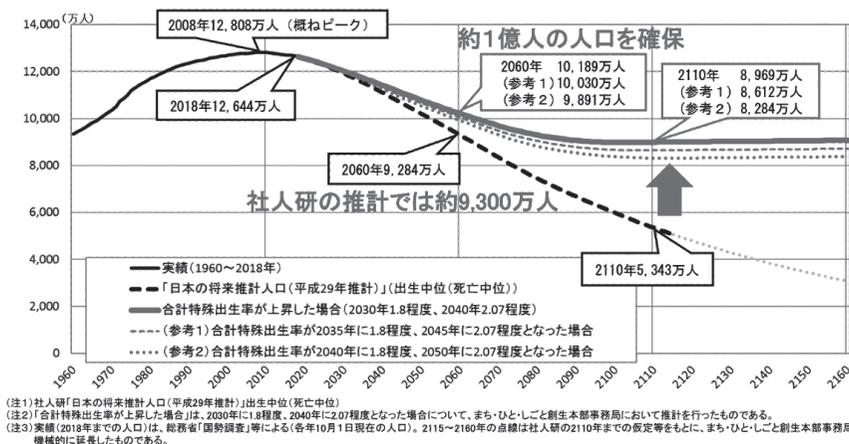


図3 我が国の人口の推移と長期的見通し  
 (「まち・ひと・しごと創生長期ビジョン(令和元年改訂版)」及び「まち・ひと・しごと創生総合戦略」より転載。  
<https://www.chisou.go.jp/sousei/info/pdf/r1-12-20-gaiyou.pdf> 参照)

2 通産・石原(2011) 265頁

3 中活協議会支援センター「中心市街地活性化支援策ハンドブック」、2015.12 参照

あり、これを食い止めるために「ストップ少子化・地方元気戦略」を提起した<sup>4</sup>。すなわち、30年後地方自治体の半分が人口減少で消滅の危機にあると警告したのである。政府もこれに呼応するかのよう同年内閣府に「まち・ひと・しごと創生本部」を設置し、「まち・ひと・しごと創生法」（本法については後述。以下「創生法」と略記）に基づく地域再生のための総合戦略と長期ビジョンの作成を各自治体に要請し、各自治体は自らが設定した戦略に従って地方創生を進めることになった。少子高齢化、人口減少、そして首都圏への一極集中を方向転換し、地方経済を立て直すことが各自治体にとっても重要課題となってきたのである。

## 2.都市再生・地域再生・まちづくりに関する法制の動向

この10年間のまちづくりに関する諸政策の展開状況を見ると、2002-06年期の改正中活法で示されたコンパクトシティ構築を中心とした政策展開と、2014-18年期の改正都市再生特措法や創生法にみられるコンパクト+ネットワーク構築をめざす政策展開が重要と思われる。

### (1) コンパクトシティ構築をめざす政策展開

この時期には2002年に都市再生特別措置法、2005年に地域再生法、2006年に改正都市計画法、改正中活法が公布・施行されている。都市再生特別措置法は、国際化、少子高齢化に対応して都市機能高度化・居住環境向上・防災機能確保のための基本方針を定め、政令で指定した都市再生特別地区や、認定された都市再生事業計画への金融支援を規定しており、後の法改正でまちづくり交付金制度創設や、まちづくり交付金対象エリアの民間都市再生整備計画への金融支援・税制特例などが創設されている。

地域再生法では、急速な少子高齢化、産業構造の変化等に対応して地域経済活性化・雇用創出等の地域再生を推進するため、内閣に地域再生本部を設置し、自治体による地域再生計画の総理大臣による認定、認定計画推進組織に関する税制上の特例措置や地域再生基盤強化交付金（道・污水处理・港の整備等）の交付等が規定されていた。

改正都市計画法では、床面積1万㎡超大規模集客施設は商業地域・近隣商業地域と制限付きの準工業地域のみ限定され、これまで郊外への大型集客施設が比較的容易に進出できた点に規制がかけられた。また、病院のような公共公益施設についても開発許可の対象とされ、周辺部への分散を制約する要因となった。

2006年公布の改正中活法では、改正都市計画法による郊外への大規模集客施設設置規制と相まって、医療、福祉、商業等の都市機能を中心部へ集積させ、さらにまちなか居住を促進することによってコンパクトシティの実現をめざしていた。旧中活法では、商業振興策が中心で都市機能集積・まちなか居住による「生活空間」再生施策が乏しかった点が改められ、協議・推進機関としてTMOに代えて中活協議会が位置づけられ、商業者以外の事業者、地権者、住民にも参加機会が広がられた。このようなコンパクトシティの構想は、自治体が策定・申請する中活計画に

表1 まちづくり3法以降のまちづくり政策の概要

年	月	事 項
1962		第一次国土総合開発計画
1979		大店法改正、500m <sup>2</sup> 以上規制
1983	12	産構審流通部会・中政審流通小委員会『80年代の流通産業ビジョン』
1989	8	産構審流通部会・中政審流通小委員会『90年代流通ビジョン』
	9	日米構造協議開始
1995	6	産構審流通部会・中政審流通小委員会「21世紀に向けた流通ビジョン」
1997	8	産構審流通部会・中政審流通小委員会「中心市街地における商業の振興について」中間答申
1998	3	第5次全国総合開発計画を閣議決定
	5・6	まちづくり3法公布
2001		中央省庁再編、総理府、経企庁等を統合し内閣府設置
2002	4	都市再生特別措置法（内閣に都市再生本部設置）
2005	4	地域再生法（法律第24号）
2006	5	改正都市計画法公布（2007.11施行）
	8	改正中活法（6月公布）
2007	5	地域公共交通活性化再生法公布
2008	12	「定住自立圏構想推進要綱」を総務事務次官通知
2009	8	地域商店街活性化法施行
2014	5	地方自治法一部改正法公布（「連携協約」制度創設）
	5	日本創生会議、「ストップ少子化・地方元気戦略」を発表
	5	改正地域公共交通活性化再生法（ネットワーク型まちづくりと連携）
	5	改正都市再生特別措置法公布、8月施行
	7	国交省、『グランドデザイン2050』公表。改正中心市街地活性化法施行
	9	内閣府にまち・ひと・しごと創生本部設置
	12	第一期「まち・ひと・しごと創生総合戦略」「まち・ひと・しごと創生長期ビジョン」閣議決定
2015	1	日本経団連、『豊かで活力のある日本』の再生』（人口1億人維持）
	3	内閣、地方創生特区として仙北市、仙台市、愛知県を設定
	8	改正地域再生法施行（小さな拠点形成を支援）
	8	第2次国土形成計画を閣議決定（2014「国土のグランドデザイン2050」がベース）
	11	安倍内閣、「地方創生」から「一億総活躍社会」へキャッチフレーズ変更
	12	まち・ひと・しごと創生総合戦略改訂版
	12	政府、通知で2015年度中にビジョン・総合戦略策定を希望→ほとんどの自治体が策定
	12	「中心市街地活性化支援策ハンドブック」作成
2016	3	第31次地方制度調査会答申：複眼的連携中枢都市圏の創設
	4	内閣府地方創生推進事務局設置（初代局長就任より）
	4	改正地域再生法施行（地方創生推進交付金、中高年齢者の「まちなか」居住支援など）
	6	まち・ひと・しごと創生基本方針2016を閣議決定
2017	1	国交省都市計画課「コンパクト・プラス・ネットワークの推進について」
	12	閣議決定、まち・ひと・しごと総合戦略改訂
2018	6	改正地域再生法施行（地域再生エリアマネジメント負担金制度創設、商店街活性化促進事業の創設）
	12	閣議決定、まち・ひと・しごと総合戦略改訂
2019	12	「まち・ひと・しごと創生長期ビジョン」「第二期まち・ひと・しごと創生総合戦略」策定
	12	地域再生法改正（「多世代共生型まち」に向け住宅団地再生、空き家の活用移住促進）
2020	6	改正都市再生特別措置法
	7	まち・ひと・しごと創生基本方針2020を閣議決定
	11	改正地域公共交通活性化法施行
	12	新型コロナ蔓延に対処するため「第二期まち・ひと・しごと創生総合戦略」改訂

・各法令等より作成。

まとめられるが、2兆円の国費投入にもかかわらずこれまでの基本計画の実現がきわめて不十分だったことを踏まえ、「選択と集中」を加味しながらトレーサビリティを担保したものとするため、数値目標の達成を求める形式が導入され、内閣に中活本部を設置（本部長は総理大臣）し基本計画も総理大臣認定に格上げされた。また、中心市街地への大型店立地促進のため認定基本計画に含まれている中心市街地中の特例区域では、大型店新設届出や説明会を不要とする便宜も提供されていた。さらに、同年9月には「中心市街地の活性化を図るための基本的な方針」を閣議決定し、「中心市街地活性化基本計画認定申請マニュアル」（内閣府中心市街地活性化担当室策定取りまとめ。以後何度か改定されている）を公表し、カテゴリー別に目標として通行量、居住人口、施設入込数、販売額、公共交通機関利用者数、空き店舗数等を設定し、5年後の検証に対応したものとしている。

これら3計画の認定・採択は、総務省行政評価局による2016年7月公表の「地域活性化に関する行政評価・監視結果報告」によると、a.地域再生計画認定は2005～2014年度で1870件・実績額1兆2638億円、b.都市再生整備計画採択は2004～2014年度で2563件・実績額7兆円、c.中活基本計画認定は2006～2014年度で177件・実績額5751億円であった。抽出した計画の効果の発現状況については、すべての指標の評価値が目標値に達した計画の比率は、a.37.1%、b.35.8%、c.0%、目標値に達した計画を除いてすべての評価値が目標達成度7割以上の計画の比率は、a.11.3%、b.7.4%、c.0%、前の2項目を除いて複数の指標のうちひとつ以上の指標の評価値が目標達成度7割以上の計画の比率は、a.35.5%、b.45.1%、c.59.1%、前の3項目を除いてすべての指標の評価値が目標達成度7割未満の計画の比率は、a.16.1%、b.11.7%、c.40.9%となっていた。これらの評価値の達成度を比較検討して総務省行政評価局は、認定された地域再生計画と採択された都市再生整備計画は「一定の効果が発現している」と評価されるものの、認定された中活基本計画については「所期の効果が発現しているとみることは困難」と結論している。

確かにこの時期は、リーマンショックや東日本大震災により経済的なダメージがあり、それぞれの計画に影響が及んだことが考えられるが、中活基本計画が他の計画の達成度と比較して低い評価となっている。このような違いの原因を究明するには、それぞれの計画の目標設定や評価方法の違い等をさらに検討する必要があるが、総務省評価局も指摘しているように、中心市街地の事業所数や事業所従業者数の減少・低下が一要因となっていることは明らかである<sup>5</sup>。すでに図1で示された小売店舗数の減少傾向は継続しており、他方では大手のネット販売の成長も著しい。その意味では、コンパクトシティ形成による中心部への都市機能集積で活性化が実現されると楽観視できるものではない。

5 総務省行政局(2016)「地域活性化に関する行政評価・監視結果報告書」150～153頁参照。なお、そこでは「市町村の規模に応じた成功例の提示」がされていないという指摘もあるが、時代に対応した小売業態開発も重要だろう。

## (2) コンパクト+ネットワーク構築をめざす政策展開

この時期には2013年に交通政策基本法が公布・施行され、少子高齢社会や国際化に対応しつつ交通政策としてまちづくりと一体となった公共交通ネットワークの維持・発展が方向づけられた。さらに2014年には改正都市再生特別措置法が公布・施行され、コンパクトなまちづくり支援のための立地適正化計画制度（都市機能誘導区域・居住誘導区域の設定）を創設し、「多極ネットワーク型コンパクトシティ」の形成を推進することとした。立地適正化計画制度には医療・福祉・商業等の生活サービス機能＝都市機能を有する施設を計画的に配置・誘導する区域＝都市機能誘導区域と、まとまった居住の推進を図る居住誘導区域が含まれており、一般的には居住誘導区域内もしくは隣接して都市機能誘導区域が配置される。このような区域設定は、地方都市の拡散した市街地での人口減少や商業の衰退に対し、従来の都市計画では市街区区域内部で空き地・空き家・空き店舗等が増加して「スポンジ化」しても対処できなかったので、このような立地適正化で計画的に対応しようとしたのである。また、まとまった居住が推進され、それらが中心部内外に点在している場合、都市機能が集積されている中心部とそれらを公共交通機関等で繋ぎ合わせる必要が生じる。2014年の改正地域公共交通活性化・再生法は、コンパクト+ネットワーク型のコンパクトシティを形成するために、地域の公共交通ネットワークを再構築することに適合した法改正であったといえよう。なお、この法の2020年の再改正は、人口減少本格化・運転者不足深刻化に対し「輸送資源の総動員」（例えば自家用有償旅客運送）等により地域における移動手段確保・地域公共交通充足が企図されている。また、都市再生整備計画に基づく事業には交付金が、立地適正化計画に基づく住宅や都市機能を高める施設については特例措置等も講じられることになっていた。

改正都市再生特別措置法ではさらに2015年に改正され、人口約20万人以上の市を中心に「連携中枢都市圏」を新たに形成し、一定の人口を確保しつつ活力ある社会経済圏を維持する方向が提起された。これに先立ち国交省や経産省、総務省などで異なる名称で同様のことが進められていたが、この改正時にこの名称に統一されている。人口減少・少子高齢社会で持続可能な経済・生活圏を築くために、中核となる都市が中枢都市宣言をし、周辺市町村と連携協約を締結することによって成立する。中枢都市が高次の都市機能を集積し協約締結市町村がそれらを共同利用することになる。2021年4月段階で連携中枢都市圏は34、宣言連携中枢都市36市、圏域構成市町村数328市町村になっている<sup>6</sup>。さらに同様の市町村連携の試みは、2009年より人口5万以上の市を中心とした市町村連携で圏域の生活関連機能を維持する「定住自立圏」形成の試みが始まっていたが、連携中枢都市圏形成とともに推進される。

同じく2014年には改正中活法が公布・施行され、「日本再興戦略」（2013年6月閣議決定）に含まれていたコンパクトシティ実現のため、認定中活基本計画への重点的支援、民間の投資喚起

6 既存の34圏域の人口は、最小の鳥取市中心圏域で26万人、最大で札幌市中心圏域で260万人と相当の格差があるが、多くみられるのは40～70万人規模である。[https://www.soumu.go.jp/main\\_content/000757545.pdf](https://www.soumu.go.jp/main_content/000757545.pdf) 参照。

(「特定民間中心市街地経済活力向上事業」を経産大臣が認定)、中心部への大型店立地手続きの簡素化、来訪者・就業者・小売販売額等の増員・増額への重点支援、小売経営の効率化やオープンカフェ等設置の推進、特例通訳案内士創設など盛り沢山の補強策が講じられた。

さらに同年、創生法が公布・施行された。本法の目的は、少子高齢化に的確に対応し人口減少に歯止めをかけるとともに、東京圏への過度の集中を是正することであり、そのために地方都市を環境面や経済面等で改善し将来にわたって活力ある社会を維持することである。「長期ビジョン」では2060年に人口1億人確保、2050年代にGDP1.5～2%の成長維持を目標としつつ、2015～19年5ヶ年の「総合戦略」の策定をおこなっている。さらに、3つの視点として①若い世代の就労・結婚・子育ての希望実現、②東京一極集中の歯止め、③地域の特性に即した地域課題の解決が提示されており、それらの点でこの創生法で課題とされたことが地方都市におけるまちづくり政策の課題に取り込まれ、自治体は総合戦略のフォローアップを求められる。

また、2015年改正地域再生法では、自治体が作成した地域再生計画を総理大臣が認定し、認定計画に基づく措置により経済基盤を強化し、生活環境整備を通じて地域活力の再生を支援するとされた。地域再生施策は就業機会の創出、経済基盤の強化、生活環境の整備を3本柱としている。特に2014年からの地方創生の流れに呼応して支援措置が強化され、「地方創生全体の方向を定める」創生法と、個別地域の具体的支援措置を提供する地域再生法が「両輪となって地方創生を推進」するものとされている<sup>7</sup>。本法ではこのような事例として「小さな拠点」形成支援が進められている。具体的には、中山間地に位置する「道の駅」等に生産や生活・福祉サービス施設を集約化し、交通ネットワークで周辺集落を結ぶものとなっている。

こうしてみると、中活法のみならず創生法、地域再生法、都市再生特措法、さらには地域公共交通活性化再生法等が全体として地方都市や地域のまちづくりを推進する体制づくりに向けられているといえよう。

ところで、これらの事業に対する政府予算は地方創生予算として計上されているが、2014年度は補正で地方創生先行型予算として1700億円計上され、2015年度以降通常予算として内閣官房が取りまとめ補正と合わせて2018年度まで単年度で2兆円程度の規模となっていた。内訳をみると1000億円が地方創生推進交付金として、各自治体が地方版総合戦略に基づいて自主的・主体的に取り組む先導的な事業を支援するために組まれている。さらに1兆円が「まち・ひと・しごと創生事業費」として、自治体が自主的・主体的に地方創生事業に取り組むことができるように、地方財政向け歳出に計上されており、同じく1兆円程度が「社会保障の充実」という名目で子ども・子育て支援等に当てられている。2019、2020年度は補正が急増しており、特に2020年度は「新型コロナウイルス感染症対応地方創生臨時交付金」に4.5兆円が組み込まれており、従来と異次元の予算規模となっている。2015年度以降の政策評価を見る必要もあるが、予算項目を概観するかぎりでは、中心市街地の商業対策にどの程度の予算措置が取られているのか判然としな

7 内閣府地方創生推進事務局「地域再生制度の概要」[https://www.chisou.go.jp/tiiki/tiikisaisei/pdf/seido\\_gaiyo.pdf](https://www.chisou.go.jp/tiiki/tiikisaisei/pdf/seido_gaiyo.pdf) 参照)

い。中活基本計画の数値目標には通行量や空き店舗、販売額等が含まれているので、地域商業や商店街に対する取り組みと成果については、中活計画に関するフォローアップを参照する必要があるだろう。

### 3.川越市におけるまちづくりの検討

ここでは中心市街地活性化事業として展開されたまちづくりの現状から、上述の諸計画が実際のまちづくりにどのように反映されているのか、そして実績としてどのような成果に結びついているのかを検討しておこう。1998年公布の旧中活法に基づく中活計画は、2005年度末までの中活計画策定市区町村数が676、中活計画数が687件あった。2006年公布の新中活法により認定された中活計画は、2007年2月から2020年10月までで148市3町より申請された257計画とかなり減少している。これは、それまでの計画策定とそれらに対する支援策が、資金の「バラマキ」的で成果が乏しいものであったと批判され、中活計画認定も数値目標の設定や毎年のフォローアップを求めるなど厳格化されたことも影響している。初年度の2006年度認定は青森市、富山市のみであったが、2007年度より2020年度までの一期目の中活計画が155件、2012年以降には二期目の中活計画77件、2017年以降には三期目の中活計画が25件認定されている。

そこで、このような転換が反映された2006年・2014年改正中活法や地方創生諸事業が適用されている川越市の事例から、まちづくり事業の展開状況を確認しておこう。川越市の中活計画は、旧中活法による一期計画が1999年に策定され、2009年6月に二期計画が、2015年3月に三期計画が認定されている。従って2014年の政府による諸制度改革を挟んで二期目と三期目の中活計画が実施されている。

川越市は人口35万人で「小江戸」、「商都」とも呼ばれる歴史のある埼玉県内の有力な経済拠点都市である。蔵造りの町並みが有名となり2018年の川越市入込観光客数は734万人、外国人観光客数も同年には27.9万人となっている<sup>8</sup>。JR川越駅と西武新宿線本川越駅を最寄駅とし、都心部からのアクセスの利便性がある。川越駅北側に広がる中心市街地は、川越城跡・蔵造りの町並みなどのある北部の歴史的・文化的地域と駅に近い南部の商業集積地域、そしてその中間で新たな賑わいをもたらした川越市産業観光館「小江戸蔵里」が立地する結節地点で構成されている。

このような賑わいの創出は、主として1980年代以降の市民団体・川越市・商業者等の協力によって実現されている。1983年には商業活性化による景観保全を標榜した市民団体「川越蔵の会」(2002年NPO法人化)が組織され、1985年には中小企業庁の「コミュニティマート事業」構想にエントリーし、1987年に一番街商店街に町並み委員会が結成され、1988年には建物や看板のあり方も規定した「町づくり規範」を策定して個店改装のルールとした。同年には「新富町まちづくり協定」、1992年には同地域の電線類地中化を実現、94年には一番街に隣接の銀座通り商店街に大正浪漫委員会が設置されている。1998年には伝統的建造物群保存地区条例が制定され、

8 株式会社まちづくり川越作成資料(2019)「まちづくり会社の活動と川越観光の現状」参照(以下の川越市中心市街地の概況については同資料参照)

観光市街地形成事業を展開して2000年には蔵造りの街並みが埼玉県下初の重要伝統的建造物群保存地区に選定されている。まちづくりへの取り組みはさらに輪を広げ、1988年には「新富町まちづくり協定」、1991年にはサンロード商店街にも同様の協定が結ばれるなど、JR川越駅北側の市街地一帯で新たなモールの形成が進められた。

このように川越市のまちづくりは約40年間の取り組みにより大きく前進したが、国の交付金を得て地域再生事業が展開され始めたのは、旧中活法に基づく中活基本計画（1999年）と「まちづくり交付金」による第一期都市再生整備計画（2004年）による事業展開以降である。2009年度以降は改正中活法（2006年公布）に基づく認定中心市街地活性化基本計画による事業展開となっているので、それ以前と比較しながら検討する。

### （1）旧中活法での中活計画（一期）

1999年には旧中活法（1998年公布）に基づく中活計画が策定され、市街地整備改善と商業等の活性化のため98事業が計画された。そのうち58事業が完了または着手済みとなった。2004年度より中心市街地地区を対象に「まちづくり交付金」による都市再生整備計画第一期事業（2004～09年度）により電線類地中化や酒造跡地再生事業等が実施され、2010～14年度には第二期事業が実施されている（後述）。2008年3月には「良好な市街地を形成するためまちづくりの推進を図るための事業活動を行う会社」として、「株式会社まちづくり川越」（中活法第15条規定、資本金3500万円、主要株主に川越市、川越市商工会議所等）が設立された。さらに同年6月には川越市中心市街地活性化協議会（参加団体：川越市、川越市商工会議所、株式会社まちづくり川越、川越商店街連合会、小江戸川越観光協会、NPO法人川越蔵の会、埼玉りそな銀行、西武鉄道、東武鉄道、市内大学等）が設置されている。そして前述のように2009年、2015年には川越市中活基本計画が認定されている。

1999年中活基本計画は期間設定がなかったこともあり98事業中40事業が未着手もしくは廃止となっている。「川越市中心市街地活性化基本計画」（2009年6月）に記載されたこれまでの取り組みによると<sup>9</sup>、実施されたのは市街地整備改善事業が71件中46件完了又は着手、商業等の活性化事業が27件中12件完了又は着手となっており、市街地整備改善事業ではクラッセ川越（2001年竣工、図書館・児童館等複合文化施設）、川越まつり会館（2003年9月開館、観光客誘致の核施設）、クレアモール（2商店街の電線類地中化を含む街路整備、1999年完成）、歴史的環境整備街路事業、道路・基盤整備事業等が実施されて、商業等活性化事業では空き店舗対策やチャレンジ・ショップ事業等が実施されている。残された課題として、交通渋滞解消への取り組みや駅への利便性向上、商業活性化については地元商店街との連携やまちづくり会社設立があった。

9 <https://www.city.kawagoe.saitama.jp/shisei/seisakushisaku/hoshinkeikaku/sangyo-kanko/shigaichikasseika.files/henkou6zenbun.pdf> 参照。

2004年度からの「まちづくり交付金」<sup>10</sup>を得た事業展開は、2002年公布の都市再生特別措置法(国の活力の源泉である都市再生をめざし、歴史・文化・自然環境等の特性を生かした個性あふれるまちづくりの支援を目的)に基づき2004年にこの交付金制度が設置されたことから始まっている。交付対象は市町村等が作成した都市再生整備計画に基づく事業であり、川越市は3つの目標(1.観光客・買物客が集う「核」施設整備、2.市域の主要3駅へのアクセス向上、3.中心商業地区・歴史的町並み地区の街路整備)を掲げて都市再生整備計画(中心市街地地区第1期計画)を作成した<sup>11</sup>。5年間で交付対象事業費約86億円、およそ半額程度が交付されており、当該地域の道路整備、電線類地中化、路面の石張り整備、旧酒造跡地の商業・観光施設としての再生(2010年10月川越市産業観光館(小江戸蔵里)として開業)、本川越駅西口広場予定用地の取得等が実現し、休日歩行者通行量や商業・観光施設利用者数を大幅に引き上げることに成功した。

この第1期計画に続く第2期計画は、2010年度から「都市再生整備計画事業」(～2014年度)として実施され、交付対象事業費は約45億円だった。事業としては当該地域の道路整備、本川越西口駅前広場の整備、旧川越織物市場保存活用事業、川越まちなかコミュニティ・サイクル事業等が実施され、休日の歩行者・自転車通行量や自動車の走行時間短縮などの数値目標が達成されている<sup>12</sup>。この二期計画事業は川越市中活基本計画(2009-15年度)に取り込まれており、基本計画全体の中で評価する必要がある。

## (2) 2006年改正中活法に基づく中活計画(第二期)

すでに触れたように2008年にはまちづくり会社川越と川越市中活協議会が設置され、コンパクトシティをめざす2006年改正(=新)中活法に基づく川越市の中活計画の「基本コンセプト」を「川越らしさを未来につなぐまちづくり」として、歴史的・文化的ストックのある中心市街地北部と駅周辺の商業・業務集積のある南部の連結を重視しつつ、回遊性と賑わいのあるまちづくりを基本方針とした。活性化の目標としては①回遊性の向上、②商業・サービス業の充実を掲げ、それらに関する数値目標も設定した。表2はこの第二期中活計画と次の第三期中活計画(2015-2020年度)に関するフォローアップ資料を集約したものである。改正中活法がコンパクトシティを強調し、街なか居住を推進しようとしていたことを反映して、川越市の中活計画でも上記の2目標を掲げたほか都市福利施設整備、中心市街地での共同住宅供給事業や居住環境向上、商業の活性化が掲げられていたが、数値目標にみられるようにまず回遊性が強調されているのは、多くの観光客が北部の歴史・文化地区をめざして来訪しているので、南部への回遊を働きかけること、

10 この交付金は2010年度に社会資本整備総合交付金に統合され、「立地適正化計画」に基づく事業支援「都市構造再編集中支援事業(個別支援事業)」(2019年度創設)とともに都市再生整備計画事業の支援に充当されている。

11 主として「まちづくり交付金(都市再生整備計画)川越市中心市街地地区【第1期計画】」、「事後評価シート」を参照。[https://www.city.kawagoe.saitama.jp/shisei/toshi\\_machizukuri/machizukuri/toshikeikaku/kouhukin/machizukurikofukin.html](https://www.city.kawagoe.saitama.jp/shisei/toshi_machizukuri/machizukuri/toshikeikaku/kouhukin/machizukurikofukin.html)

12 川越市(2015.3)「都市再生整備計画 事後評価シート」参照。<https://www.city.kawagoe.saitama.jp/shisei/oshirase/chikutoshisaisei.html>

表2 川越市の中心市街地活性化基本計画フォローアップ にみられる達成度

評価項目	第二期全体総括 (2009.6-2015.3)	第三期全体総括 (2015.4-2020.3)
1. 計画終了後市街地概況	新たな観光資源へ回遊、事業所数は微減	新たな賑わい・回遊性向上、街なか居住増
2. 計画事業の進捗・完了	概ね順調に進捗・完了 / 若干の活性化	概ね順調に進捗・完了 / 若干の活性化
3. 進捗・活性化状況詳細	計画事業の 66/82 事業を実行	計画事業の 80/84 事業が完了・実施中
4. 中活協議会の意見	若干の活性化	若干の活性化
5. 市民意識の変化	若干の活性化	若干の活性化
6. 今後の取組	空き店舗対策や商店街の連続性改善	観光地点広域化とイベント等の継続
個別目標		
回遊性の向上		
平日 (歩行者・自転車通行量)	超過達成 (A)	基準年値 < 実績 < 目標値 (B2)
休日 (歩行者・自転車通行量)	超過達成 (A)	
半日以上滞在観光客割合	基準年値 < 実績 < 目標値 (B)	* 基準年値 < 実績 < 目標値 (B2)
商業・サービス業の充実 (事業所数)	実績 < 基準年値 < 目標値 (C)	** 実績 < 目標値 < 基準年値 (A)

\* は観光客立ち寄り観光地点数、\*\* は空き店舗数。

・ (A) は計画事業の予定通りの進捗・完了と目標値越えの実績、(B) は計画事業は (A) と同じだが、実績が基準値超だが目標値未達、(B2) は計画事業は (A) と同じだが、実績が目標値の 80% に未達。

・ 出典: 各期の最終フォローアップ報告書より作成した (<https://www.chisou.go.jp/tiiki/chukatu/list.html> 2021年9月30日参照)。

商業・サービス業の充実によって日帰り観光客の滞在時間をできるだけ延ばすことや居住者への利便性を高めるという独自性を反映したものといえよう。数値目標の達成度をみると、歩行者・自転車通行量は実績で 9.3 万人 / 平日 (目標値 8.1 万人)・15.7 万人 / 休日 (目標値 13 万人) となっており、好成績だった。滞在時間半日以上以上の観光客の割合は、2014 年の実績で 55.6% で、基準値であった 2005 年の 51.6% を超えたものの目標値 61.6% には達しなかった。さらに商業・飲食業・サービス業の事業所数は、2014 年実績で 2,092 で、基準値であった 2006 年の 2,268、目標値の 2,310 には届かなかった。

この実績の背後どのようなことが実現されていたのか。まず第一に市街地再開発事業や道路等の公共施設整備の点で、2010 年度に回遊増進につながる歩行環境改善事業が戦略的市街地商業等活性化支援事業費補助金で実施されたほか、道路整備や本川越駅西口開設・駅前線整備のため社会資本総合交付金 (都市再生整備計画事業) を活用した整備事業が 2016 年実現めざし実施されている。第二に都市福利施設整備事業として「ウエスタ川越」(大ホール、生涯学習施設、県施設等で構成) が社会資本整備総合交付金を活用して 2015 年に開業した。第三に商業活性化事業として川越産業観光館 (飲食施設、物販・情報・休憩施設、ギャラリー等で構成) がまちづくり交付金・中小商業活性化支援補助金を活用して竣工・オープンした。この施設はまちづくり川越の管理施設でもあり、その事務所も置かれていた。第四に空き店舗対策事業やチャレンジショップ事業が、2009 年より中心市街地活性化ソフト事業としてまちづくり川越により実施され始めた。第五にコミュニティサイクル施設が都市再生整備計画事業として設置され、社会実験を経て 2013 年から稼働している。それ以外にも川越城本丸御殿保存整備事業 (埼玉県文化財保存事業補助金活用)、川越城中ノ門堀跡整理事業 (まちづくり交付金・埋蔵文化財国庫補助活用)、旧川越織物市場保存活用事業 (社会資本整備総合交付金活用の都市再生整備計画事業) 等が取り

組まれており、それらの相乗効果が上記の成果に結びついている。ただし、市街地居住の促進については、マンション建設が継続して居住人口が増加傾向にあったためか、顕著な事業はみられない。

ところで、事業所数に関する評価がCランクとなっている点で、フォローアップ報告では産業観光館やウェスタ川越で目標値に届かなかった理由を、活用用途変更や当初予定の小規模店ではなく大型店出店による減少としてさほど問題視してない。ただし市域及び当該市街地の小売店舗数は減少傾向にあり、空き店舗対策がさらに重要となってくることを勘案すると、「商店街等空き店舗情報登録制度」の実施は、まちづくり川越が担うチャレンジショップ事業の推進にも有益だったはずである。

### (3) 2014年改正中活法に基づく中活計画（第三期）

この計画の基本コンセプトは、「川越らしさを活かした交流とにぎわいのあるまち」であり、方針として第一に魅力あるまちなみづくり（そのために歩行環境改善、市街地の南から北までの連続性確保のため結節地域の活性化、新たなにぎわいスポット創出）、第二ににぎわいの創出（そのためにウェスタ川越の活用、結節地域の産業観光館活用、未活用の歴史的・文化的資産整備、空き店舗対策、まちづくりルールの策定）の2つを掲げていた。具体的事業としては、①市街地の整備改善事業（都市計画道路整備、本川越駅西口開設工事、3駅周辺地区の整備、川越城富士見槽跡整備）、②都市福祉施設整備事業（地域子育て支援拠点事業、文化芸術振興、市民活動拠点施設運営管理）、③街なか居住の推進事業（建築物耐震化促進、住環境改善、共同住宅低層階への商業施設等入居啓発）、④経済活力向上事業（旧鶴川座等歴史的建造物の保存活用、チャレンジ・ショップ事業、小江戸川越春まつり等の実行、空き店舗対策事業、商店街一店逸品運動推進等）、⑤公共交通機関の利便性増進事業（交通円滑化、自転車シェアリング等）が掲げられていた。

この計画に関するフォローアップも表2に示したとおりである。計画事業をほとんど完了もしくは実施した点で積極性が窺える。国の支援があった事業として、旧鶴川座再生・利活用事業（経産大臣認定、改正中活法の特定民間中心市街地経済活力向上事業、経産省）、旧川越織物市場保存整備事業（社会資本整備総合交付金、国交省）、伝統的建造物群保存地区保存整備事業（国宝重要文化財等保存・活用事業補助金、文科省）、オリンピック・パラリンピック推進（地域活性化・地域住民生活等緊急支援交付金、内閣府）、チャレンジ・ショップ事業（中活ソフト事業、総務省）等があった。管轄省庁が多岐にわたっており、中活事業は省庁横断的に推進されている。前回の計画対象期間と異なる点として、川越駅西口の文化芸術振興・市民活動拠点施設のウェスタ川越が稼働（2015年3月開業）していたことと、本川越駅西口が利用開始（2016年2月）となり、川越市駅との連絡道路も整備されたので移動時間が半減している。また北部に位置する旧山崎家別邸母屋が2019年に国の重要文化財指定を受けた。今期の基本コンセプトが「交流とにぎわい」となっているように川越市は観光を重視しており、入込観光客数は2015年665万人が

2019年には776万人に増加している<sup>13</sup>。しかしほとんどが日帰り客なので、北部の歴史・文化エリア以外の南部や結節地域への回遊を課題としており、今回は「半日以上滞在観光客割合」を「観光客の立ち寄り観光地点数」に置き換えて調査された。調査結果は、基準値の4.27か所（2013年）<4.38か所（2019年実数）<4.50か所（2019年目標値）ということでB2という評価にとどまっている。伝統的建造物群保存地区内の蔵造り資料館が閉館中という不測の事態もあったが、未活用の歴史的・文化的資産の整備や、結節地域でのイベント等による誘引力引き上げによる広域回遊に課題が残されたといえる。歩行者・自転車通行量も期間中に8.5万人/時間から10.5万人/時間へと増加し、11万人/時間の目標値に接近したが、想定していた2つの整備事業の遅れが影響したようである。

個別目標のうち商業・サービス業の充実・活性化のみがA評価となっている。具体的には空き店舗数が2015年の83店舗から2019年には62店舗に減少しており、基準値（2012年74店舗）、目標値（2019年64店舗）をクリアしていた。まちづくり会社である「(株)まちづくり川越」は空き店舗情報登録制度により関係情報を収集し、空き店舗を借り上げて新規開業希望者等に貸し出し、商工会議所と連携しながら経営指導と独立開業を支援するチャレンジ・ショップ事業やテナントミックス事業を担っている。小売店舗等の減少傾向が続くことを勘案すると、この組織の果たす役割はさらに重要となってくるだろう。

なお、旧川越織物市場保存整備事業は、中活計画でも「文化創造のためのインキュベーション機能と来街者も利用できる交流休憩機能を有する拠点」と位置づけられているが、「川越市まち・ひと・しごと創生総合戦略」（2016年1月策定）においても「文化創造インキュベーション施設」、「現代のアーティストやクリエイターの活動拠点」として位置づけられている<sup>14</sup>。中活計画では、計画期間内に整備工事が完了しなかったために、インキュベーション機能を発揮できていなかった。

#### （4）立地適正化計画における中心市街地の位置

2014年に改正都市再生特別措置法が公布され、コンパクトなまちづくりを支援するため立地適正化計画制度が導入されたが、川越市でも2018年に「川越市立地適正化計画」が策定された。すでに指摘したように、この計画は市町村が都市全体の観点から「都市機能誘導区域」と「居住誘導区域」を定め、これらの指定区域以外で開発・建築行為をおこなう場合は市町村への届出が必要になるというものであり、人口減少と少子高齢社会に対応した自治体のあるべき姿を方向づけたものといっても過言ではない。そこでは中心市街地のみに都市機能を集中させるのではなく、多極ネットワーク型都市構造を形成する（図4）。川越市の場合、市内6ヶ所に都市機能誘導区域を設け、中心市街地を「都心核」（北部の歴史的町並みを包摂した区域と南部の3駅を中

13 川越市の統計による。<https://www.city.kawagoe.saitama.jp/welcome/kankobenrijoho/kankotokeishiryoy/irikomi.html>

14 川越市政策財政部政策企画課（2016）「川越市まち・ひと・しごと創生総合戦略」参照。

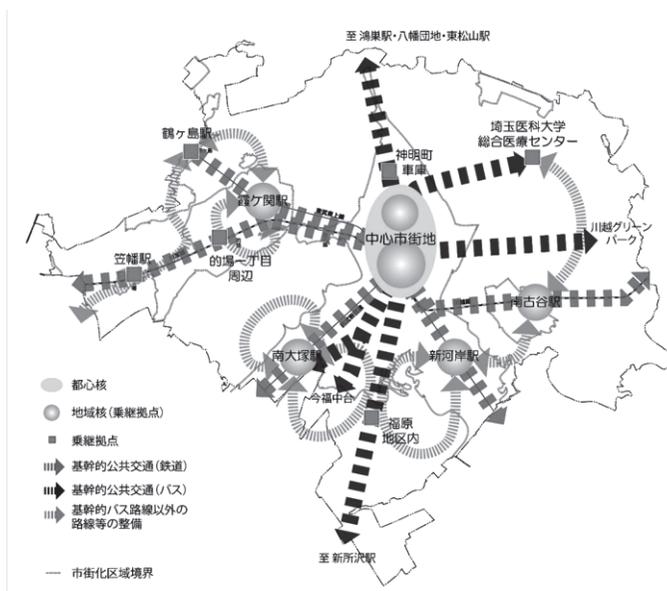


図4 立地適正化計画における将来都市構造(川越市(2018)『立地適正化計画』41頁

[https://www.city.kawagoe.saitama.jp/shisei/toshi\\_machizukuri/machizukuri/toshikeikaku/masterplan/toshikeikaku20330.files/ritittekiseikazenntai.pdf](https://www.city.kawagoe.saitama.jp/shisei/toshi_machizukuri/machizukuri/toshikeikaku/masterplan/toshikeikaku20330.files/ritittekiseikazenntai.pdf)

心とした区域の2区域で構成)とし、周辺の4つの駅をそれぞれの中心とした4「地域核」で構成し、都心核と地域核を基幹的公共交通で結ぶというプランを作成している。そしてこの都心核及び地域核を中心に居住誘導区域をを設け、医療・福祉・商業施設などの生活サービス機能を備えた都市機能誘導施設を配置する。さらにこれらの都心核と各地域核は、機関的公共交通網(鉄道またはバス)で結合されている。そしてこれらの居住誘導区域は公共交通の利便性のほかに、適度な人口密度を維持

しながら、災害の危険性の高くない良好な住宅市街地となるエリアを設定することになる。

このように川越市の中心市街地が都心核を構成することになるが、南部の3駅中心エリアには県南西部地域の拠点にふさわしい総合病院や大型商業施設を配置するほか、高齢者支援のための相談窓口施設、障害者総合相談支援施設、認定こども園、認可保育所、生涯学習施設、地域交流施設等が配置され、北部の歴史的町並み周辺エリアには病院、高齢者支援のための相談窓口施設、認定こども園、認可保育所、地域交流施設等が配置されるという。従って、中心市街地内でも若干異なる都市機能誘導施設を配置したエリアに分かれることになり、より大型の商業施設は3駅周辺エリアに位置づけられている。さらに将来にわたって人口30万人以上の都市を維持するためには、地域核にも大型スーパーなどが存続する必要があるだろうが、すでに触れたように利便性ととも個性豊かな魅力的な都市であるためには、中心市街地における歴史的な町並みや地域産品、独自商品・サービスを提供する商店などで構成される商店街がまちの顔として維持・再生される必要がある。

## 小括

小稿では、主として中心市街地活性化政策に注目しながらまちづくり政策における商業政策のあり方を検討しようとした。2010年代以降、人口減少と少子高齢社会、首都圏一極集中と地方都市の衰退が現実味を帯びてきくようになり、多極ネットワーク型コンパクトシティ形成を目標に、人口20万人以上の地方都市を中心に連携中枢都市圏の形成、5万人以上都市を中心に小規模な定住自立圏形成、そして中山間地での「小さな拠点」形成と、さまざまなレベルでの都市

再編構想や地方の生き残り戦略が探求されてきた。その中で生産と消費を接合する流通の役割が改めて見直されはじめている。小売業者の生き残りではなく、地域での豊かな生活にとって不可欠な都市機能の担い手としての役割が見直されているのである。この都市機能の役割と配置については、地域再生、地方創生、都市再生、中心市街地活性化に関する政策等でも取り上げられているが、これらの政策のフォローアップでは事業所数や売上高、従業者数、空き店舗数等が必ず評価の基準に含まれており、それらを通して地域の社会経済的基盤としての小売業集積の維持・発展が評価されている。その点では、当該地域のニーズを満たしうる小売業の集積が求められているのであって、必ずしも小規模零細小売企業の維持・発展が志向されているのではない。しかし、それぞれの地域の個性や豊かさを継承・形成する上で、全国展開しているチェーン店や大規模商業資本にはない、地域に根ざした小売業の発展は必要不可欠と思われる。また、商店街を形成する地元業者はしばしば地域の文化や祭礼などの担い手でもある場合が多く、空き店舗対策やチャレンジ・ショップの支援は次代を担う人材育成にも繋がるだろう。

錯綜する政策展開を地方の実態で追跡することによりその妥当性・有効性が確認できるだろう。川越市での中心市街地活性化を検討することで政策の東がどのような結果をもたらすのかを解明しようとした。対象地が首都圏にあり、多くの観光客が来訪するという比較的良好な環境ではあったが、小売店舗数の減少は着実に進行しており、それを巻き返す試みが歴史的・文化的遺産を再生整備しつつ、まちづくり会社のチャレンジ・ショップ事業として展開されている。そもそもこれらの遺産も川越商人が築いたものであったし、それらを活かす積極的試みも商店街や住民、商工会議所、自治体が続けて今日に至っている。

ところで、中活計画のフォローアップでは、「中心市街地の状況に関する基礎的なデータ」として計画期間中の人口動態表が挿入されているものの、まちなか居住に関する言及がほとんどみられない。まちなか居住を重視している改正中活法の趣旨から考えると疑問に思われるが、川越市の中心市街地では全国的にみられる少子高齢化や人口減少がみられないためであろう。中心部にマンションが建設され転入者や子供たちが増えている。そのためまちなか居住について当面は問題視しなくても良いと考えられている。それならまちなか人口にもっと注目している事例はないのか。表3は富山市のフォローアップの事例である。富山市は青森市とともに中活計画を最初に提出した自治体であり、公共交通機関のLRT（Light Rail Transit）を導入して従来の路面電車と接合することによりコンパクトシティ化を推進し、世界的にも注目される成果を上げている。3期にわたる中活計画の当初より「まちなか居住」を意識していることがよくわかる。第1期では高齢化率の高い地域であったため人口の自然減が転入人口を上回る状態だったが、第2期、第3期になると「まちなか居住」や共同住宅建設の支援等により転入人口が大幅に増え目標を達成している<sup>15</sup>。このような富山市の事例は、人口が減少傾向にある首都圏以外のほとんどの

15 「まちなか居住」増加にもかかわらず中心商業地歩行者通行量が目標に届いてないが、富山市は中心市街地での空き店舗数の増加や店舗構成の偏りに問題があるとみている。ただし、2014年以降は通行量は増加傾向にあり、2017年以降は目標を達成したものの、2019年以降は新型コロナの影響で減少している。

表3 富山市中活基本計画最終フォローアップにみられる達成度

評価項目	第1期全体総括 (2007.2-2012.3)	第2期全体総括 (2012.4-2017.3)	第3期全体総括 (2017.4-2022.3)
1. 計画終了後市街地概況	交通インフラ整備で回遊性向上、転入増	再開発事業が進展、富山駅周辺の回遊性弱	目標達成したが新型コロナで後退
2. 計画事業の進捗・完了	計画事業 10/27 完了、16/27 着手済み	概ね順調に進捗・完了 / 若干の活性化	路面電車南北接続、商業施設新設、居住人口増
3. 活性化如何の要因	まちづくり理念浸透、民間投資活性化	計画事業の 9/66 完了、51/66 着手	ワクチン接種につれ通行量・賑わい回復傾向
4. 中活協議会の意見	乗車数・通行料で着実な成果、居住人口改善	かなり活性化、乗車数・居住人口は目標値達成	すべての目標値達成、ワクチン接種後回復期待
5. 市民意識の変化	概ね好評だが都心ライフの認知度が低い	若干の活性化	
6. 今後の取組	市街地整備等の公共投資で民間投資促進	富山駅での分断状況解消、回遊促進、中心での QOL 向上	
個別目標			
回遊性の向上			
路面電車市内線乗車数	基準値 < 実績 < 目標値 (B)	基準値 < 目標値 < 実績 (A)	最新値 < 基準値 < 目標値 (C)
中心商業地歩行者通行量	基準値 < 実績 < 目標値 (B)	実績 < 基準値 < 目標値 (C)	最新値 < 基準値 < 目標値 (C)
まちなか居住人口	実績 < 基準値 < 目標値 (C)	基準値 < 目標値 < 実績 (A)	基準値 < 目標値 < 最新値 (A)

- ・第3期全体総括は2020年5月のフォローアップ報告なので最終報告ではない。
- ・第3評価項目は、第2期は進捗・活性化状況詳細を、第3期は2021.5「目標達成見通しの理由」を記入。
- ・「最新値」は2020年度の数値。
- ・出典：各期の最終フォローアップ報告書等より作成した (<https://www.chisou.go.jp/tiiki/chukatu/list.html> 2021年9月30日参照)。

地方都市における中心市街地活性化計画の典型的事例といえよう。

最後に、最近、中心市街地活性化事業展開で計画立案と実行に重要な役割を果たしてきた中活協議会の解散が散見されることに留意しておく<sup>16</sup>。川越市でも2020年12月に中活協議会を解散しているが、川越市産業振興課によると第三期中活計画の終了により業務満了となり解散したとのことであった。今後の中活事業については(株)まちづくり川越に委嘱しながら進める予定とのことで、まちづくり業務の管理と実行で実績を上げてきたこの法人が中活協議会の機能のある程度継承しようとしている。この会社が、観光産業館の管理運営、チャレンジショップ事業、自転車シェアリング事業、情報センター事業、観光案内所事業等を担ってきた実績を勘案すると、実働部隊としてさらなる活躍が期待できる反面、中活協議会が市内の多様な機関や組織の協議機関であったことを勘案すると、「新しい公共」の担い手となるためには相応の財源と人材が必要になるのではないだろうか。

#### 参考文献

- みずほ総合研究所編 (2018) 『キーワードで読み解く地方創生』 岩波書店
- 田中道雄 (2021) 『まちづくり大全』 中央経済グループパブリッシング
- 岡崎昌之 (2020) 『まちづくり再考—現場から学ぶ地域自立への道しるべ—』 ぎょうせい
- 畠山直 (2020) 「都市施設としての小売商業—まちづくり3法から立地適正化計画へ—」 (日本流通学会『流通』No.47)

16 大津市の「まちづくり大津」による調査を参照。 <http://www.machidukuri-otsu.jp/pdf/33-11.pdf>

- 白戸伸一 (2020) 「イギリスにおけるまちづくり推進組織 (BIDs) に関する検討—まちづくり組織の日英比較—」(『明治大学社会科学研究所紀要』第58巻第2号)
- 山口幹幸 (2019) 『コンパクトシティを問う』株式会社プロGRESS
- 佐藤道彦・佐野修久編 (2019) 『まちづくりイノベーション—公民連携・パークマネジメント・エリアマネジメント—』日本評論社
- 馬奈木俊介・中村寛樹・松永千晶 (2019) 『持続可能なまちづくり—データで見る豊かさ—』中央経済社
- 山出保 (2018) 『まちづくり都市金沢』岩波書店
- 月尾嘉男 (2017) 『転換日本 地域創生の展望』東京大学出版会
- 畠山直 (2017) 「転機を迎えた商業まちづくり政策」(日本流通学会『流通』No.40)
- 参照URL
- 中山徹 (2016) 『人口減少と地域の再編—地方創生・連携中枢都市圏・コンパクトシティ—』自治体研究社
- 渡辺達朗 (2016) 『流通政策入門 (第4版)』中央経済社
- 渡辺達朗 (2014) 『商業まちづくり政策—日本における展開と政策評価—』有斐閣
- 佐々木保幸・番場博之編著 (2013) 『地域の再生と流通・まちづくり』白桃書房
- 通商産業政策史編纂委員会編・石原武政編著 (2011) 『通商産業政策史 1980-2000 第4巻 商務流通政策』経済産業調査会
- 石原武政・西村幸夫編著 (2010) 『まちづくりを学ぶ—地域再生の見取り図—』有斐閣
- 角谷嘉則 (2009) 『株式会社黒壁の起源とまちづくりの精神』創成社
- 横森豊雄・久場清弘・長坂泰之 (2008) 『失敗に学ぶ中心市街地活性化—英国のコンパクトなまちづくりと日本の先進事例—』学芸出版社
- 野木村忠度 (2008) 「地域活性化とマーケティング—事例研究：埼玉県川越市の商店街の取り組み—」(『明大商学論叢』第90巻(特別号))
- ・ [https://machi.smrj.go.jp/document/i7q3ae00000031xc-att/handbook\\_h27.pdf](https://machi.smrj.go.jp/document/i7q3ae00000031xc-att/handbook_h27.pdf)
- 中活協議会支援センター (2015) 「中心市街地活性化支援策ハンドブック」
- ・ [https://www.mlit.go.jp/crd/index/handbook/2020/2020tyukatuhandbook\\_1.pdf](https://www.mlit.go.jp/crd/index/handbook/2020/2020tyukatuhandbook_1.pdf)
- 中心市街地活性化法の改正の経緯、国土交通省都市局まちづくり推進課(2020)「2020(令和2年度)中心市街地活性化ハンドブック」所収
- ・ [https://www.soumu.go.jp/main\\_content/000635353.pdf](https://www.soumu.go.jp/main_content/000635353.pdf)
- 内閣官房まち・ひと・しごと創生本部事務局・内閣府地方創生推進事務局参事官島田勝則 (2019) 「地方創生の現状と今後の展開」
- ・ <https://www.chisou.go.jp/sousei/index.html>
- 内閣官房まち・ひと・しごと創生本部事務局・内閣府地方創生推進事務局 (2021) 「内閣官房・内閣府総合サイト 地方創生」

・ <https://www.chisou.go.jp/tiiki/tiikisaisei/pdf/tebiki.pdf>

内閣府地域活性化推進室（2014）「地域再生計画作成の手引き～地域再生基盤強化交付金活用のために～」

・ <https://www.chisou.go.jp/sousei/info/index.html#an6>

地方創生 関係法令・閣議決定等

・ [https://www.chisou.go.jp/sousei/info/pdf/panf\\_vision-sogo.pdf](https://www.chisou.go.jp/sousei/info/pdf/panf_vision-sogo.pdf)

内閣官房 まち・ひと・しごと創生本部事務局 まち・ひと・しごと創生「長期ビジョン」「総合戦略」

・ <https://www.chisou.go.jp/sousei/info/pdf/r1-12-20-gaiyou.pdf>

まち・ひと・しごと創生長期ビジョン（令和元年改訂版）及び 第2期「まち・ひと・しごと創生総合戦略」（概要）

・ [https://www.soumu.go.jp/main\\_content/000431927.pdf](https://www.soumu.go.jp/main_content/000431927.pdf)

総務省行政評価局（2016）「地域活性化に関する行政評価・監視結果報告書」

・ <https://www.chisou.go.jp/sousei/about/yosan/index.html>

内閣官房・内閣府総合サイト 地方創生（地方創生予算）



【研究論文】

# 英語スピーキング授業に対する学習者の評価と 個人差要因の関連

－対面授業とオンライン授業の比較を通じて－

The Relationship between Learners' Evaluation of English Speaking Class  
and Individual Difference Factors

－A Comparison between Face-to-Face and Online Classes－

廣 森 友 人

HIROMORI, Tomohito

金 田 慎 吾

KANEDA, Shingo

新 井 智 大

ARAI, Tomohiro

大 石 有 紗

OISHI, Arisa

大 谷 温 理

OTANI, Onri

柴 田 恒

SHIBATA, Wataru

## Abstract

Many universities have introduced “hybrid-type” classes in recent years, and students have more opportunities to choose between face-to-face and online classes. On the other hand, the learning effects of both types of classes can vary greatly depending on individual differences among learners. In this study, we focused on university students' English speaking classes and investigated (1) whether learners' evaluations of face-to-face and online classes differed and (2) if so, how they were related to their proficiency, gender, and individual difference factors. The participants were 100 first-year university students who were taking a required English speaking class. A two-part questionnaire measuring (1) learners' evaluation of face-to-face and online classes (e.g., ease of class participation, ease of speaking English) and (2) individual difference factors (e.g., personality, learning style) was developed for the survey, which was

administered during part of the class time. After a preliminary analysis of the questionnaire, *t*-tests were used to determine whether there were differences in learner evaluations between face-to-face and online speaking classes and whether proficiency and gender affected these evaluations. For individual difference factors among learners, cluster analysis was employed to classify learners into groups with similar individual characteristics, and then ANOVAs were used to examine whether there were differences in learner evaluations among the groups. The results showed that while differences in language proficiency and gender had no significant relationships with class evaluation, individual difference factors significantly impacted its evaluation. Based on the results, pedagogical implications for learners and teachers are discussed concerning the use of face-to-face and online classes in the university English speaking course.

**Keywords:** face-to-face class, online class, learners' evaluation, individual difference factors

## I はじめに

近年、新型コロナウイルス感染症（以下：COVID-19）の蔓延に伴い、多くの大学がオンライン授業を実施している。文部科学省（2020a）は、2020年6月1日時点において、963校（90.3%）の高等教育機関が遠隔授業を一部、あるいは全面的に実施していると報告している。

オンライン授業についてのアンケートは数多く実施されてきたが、その評価は割れている。文部科学省が実施した調査（2021）によると、56.9%の学生はオンライン授業に対して「満足」「ある程度満足」であった。その理由として、「自分の選んだ場所で授業を受けられた」「自分のペースで学修できた」という回答が得られた。一方で、「あまり満足していない」「満足していない」と回答した学生は20.0%であった。その理由としては、「友人などと一緒に授業を受けられず、寂しい」「レポート等の課題が多かった」「身体的疲労を感じた」「質問等、相互のやりとりの機会がない」という意見であった。

昨今、COVID-19対策として、ハイブリッド型授業が導入され、学生は対面とオンライン授業の選択を迫られる機会が増大した。感染収束後も、発達段階に応じて対面とオンラインによる指導を並行して使いこなすことが教育現場に求められることが予想される（文部科学省、2020b）。

これらのことから、学習者にとって対面授業とオンライン授業を比較し、自身に合う授業形態を選択することは、今後の大学教育において必要不可欠となる可能性がある。しかし、その選択の基準は明らかにされていない。したがって、学習者の個人差要因（Individual Difference Factors）を考慮した授業形態の選択に資する指標が必要である。

そこで、本研究では、大学生の英語スピーキング授業を例に、対面とオンラインの比較を行い、それぞれの授業形態に適した個人差要因を検討することにより、授業選択の一助とすることを目的とする。

## II 先行研究

### 1 オンライン授業の評価

コロナ禍以前から、オンライン授業は、主に対面授業と比較されることによって評価されてきた。その評価基準は、授業を受ける環境の快適さ (Spencer & Temple, 2021), コミュニケーションの取りやすさ (Forture et al., 2011), 学習成果 (Johnson et al., 2000) などがある。例えば, Spencer and Temple (2021) は, 大学生に対して対面とオンライン授業に対する認識を調査した。その結果, 学習成果に関しては対面授業のほうがオンライン授業よりも評価が高かった一方, オンライン授業は学習のしやすさ, デバイスの使いやすさという点で対面よりも評価が高かった。さらに, フィードバックを受け取りやすいこと, 課題解決能力が向上することなどがオンライン授業の利点として挙げられている。また, Forture et al. (2011) による同様の調査では, オンライン授業は対面に比べて, コミュニケーションが快適で, 質問がしやすく, 発言への抵抗感が少ないという結果が得られた。

日本国内においてもオンライン授業についてのアンケートは多くの大学で実施されているが, その評価は割れている。慶應義塾大学が教員と学生を対象に行った調査によると, オンライン授業に対し, 教員の約半数, 学生の約7割がコロナ危機後も大学がオンライン授業を継続することを希望していることが示されている (植原, 2020)。また, 立教大学 (2020) が発表した「オンライン授業に関する学生の意識調査の結果」では, 双方向オンライン授業の満足度が前年度の対面授業よりも高いことがわかった。一方で, 東洋大学の調査 (2020) によると, 従来の対面授業と比較して, 現在のオンライン授業において, 肯定的な意見が37%, 否定的な意見が28%であった。いずれの調査においても, オンライン授業に対してやや肯定的な意見が多いが, その差は僅かであることがわかる。

このことから, 対面授業もオンライン授業もメリットとデメリットの両方があり, それらの感じ方も学習者によって多種多様であることが伺える。

### 2 オンライン授業と日本における英語スピーキング授業

COVID-19 以前から, 大学における ICT 活用と英語学習の有用性に関しては, 多くの実践がなされてきた。日本においては, 1990 年代後半より CALL (Computer Assisted Language Learning) とインターネットサービスの普及に伴い, 大学でも ICT を活用した授業実践が急速に広まり, 従来の環境では想像し得なかった規模のコミュニケーションが現実となった。COVID-19 以降は, 英語学習における同時双方向型オンライン授業の実践研究が少しずつ進められてきた。

また, オンラインでのコミュニケーションは対面と比較して「対人圧力」の弱い状態である (木村・都築, 1998) という特性を利用し, 海外では同時双方向型オンライン授業におけるスピーキ

ングに着目した研究が進んでいる (e.g., Martin, 2020; Nurwahyuni, 2020)。さらに、コミュニケーション能力がより重視される現代社会で注目されてきた「コミュニケーション・ランゲージ・ティーチング (CLT)」が同授業形態でも可能であることを示す研究もある (Ng, 2020)。しかし、日本では COVID-19 以降、同授業形態による英語スピーキング授業に特化した研究は管見の限りあまり行われていない。数少ない事例の一つである大嶋 (2021) は、スピーキングにおいて同時双方向型オンライン授業を重ねるごとに、言語面、内容面、パフォーマンス面のそれぞれで質的な向上の様子が確認できたことを報告している。大谷 (2021) や前川 (2020) も同様に、オンライン授業における学生の発言について触れているが、いずれも対面授業との比較はされておらず、その検討が必要であるとしている。コロナ禍以前だと、吉田 (1999) による大学生を対象にした研究では、対面授業とオンライン授業での発話時間を比較したところ、オンライン授業での発話時間の方が多かったということが報告されている。

日本人のスピーキング能力は他のアジア諸国と比較して低い。2019年の TOEFL iBT® のスピーキングスコアを見ると、日本はアジアにおいて最下位であった (ETS, 2019)。また、スピーキングに自信が持てない日本人英語学習者は非常に多く (古家・櫻井, 2014), (株)アルク (2016) では同社が実施する英語スピーキングテストを受験した 25,559 人において、9 割近くが英語を話して業務や学業を行うことが難しいと報告した。この結果を鑑みると、英語スピーキング授業という枠組みの中で、対面とオンラインのどちらの授業形態の活用がより日本人学習者のスピーキング能力に寄与するかについて検討することは価値があると考えられる。

### 3 オンライン授業と個人差要因

上述の通り、全ての学習者がどちらか一方の授業形態で学習することは合理的とは言えないことを考慮すると、学習者の個人差要因に目を向ける必要がある。個人差要因とは個々の学習者特有の要因であり、学習者の学習効率や達成レベルの違いの要因になる可能性があるものと定義されている (Richards & Schmidt, 2010)。第二言語学習における個人差要因は様々なものが存在するが、先行研究では主に動機づけ、学習ストラテジー、学習スタイル、パーソナリティ、言語適性などが研究対象となってきた (Dörnyei, 2005; Skehan, 1998)。また、個人差要因は生得的で変化しないものではなく、学習の過程で絶えず変化する動的なものとして捉えられてきた (Gregersen & MacIntyre, 2014)。

一方、オンライン授業の成否と関係する個人差要因は、先行研究でいくつか挙げられている。Borg et al. (2021) は、大学生における双方向同期型オンライン授業への認識とパーソナリティの関連について調査している。ここでは、外向性と調和性が高い学生はオンラインで実施されるグループワークに対して肯定的なビリーフを持つことが明らかにされ、対面授業を実施できない際は、オンラインによるグループワーク等のスピーキング活動が奨励された。Weiser et al. (2018) は同授業形態における大学生のインタラクションへの参加度とパーソナリティの関連を調査した。彼らは外向性-内向性、および情緒安定性-情緒不安定性の尺度を用い、その結果

として、外向的な参加者はどの授業形態においても積極的に対話を図る一方で、不安などの情緒の安定性は授業の参加度とはあまり相関がなかったことを報告している。

オンライン授業と学習スタイルの関係性に着目した研究も存在する。例えば、Goodridge et al. (2017) は双方向オンライン授業と対面授業の成績の比較を Kolb (1984) の学習スタイル調査を用いて実施した。その結果、収束型（まず考え、それが正しいかどうかを実行して確かめるタイプ）の学習者と同化型（頭の中で論理を構築し、考えることを好むタイプ）の学習者が双方向オンライン型の授業形態においてほかの学習者よりも好成績であった。Dag and Gecer (2009) はオンライン英語学習における学習スタイルと学習効果の関連を調査した。ここでは、英語学習の効果を高めるためには学習スタイルだけではなく、学習者の動機づけ、教師の指導法も大きな要因になっていることが示された。

また、第二言語学習では一般的に、発話に対する抵抗感が高いと考えられる。磯田 (2008) は授業内での英語を話すことに対する抵抗感を測定する尺度を WTC (Willingness To Communicate; ある状況で、第二言語を用いて、自発的にコミュニケーションをしようとする意思) の理論に位置づけて開発した。その尺度をもとに、磯田 (2009) は SPM (Sentences Per Minute) というスピーキングの流暢さを高める指導法を用いて、それによって発話に対する抵抗感が低下するという仮説を立てた。結果、その仮説は立証される形となった。彼による一連の研究結果は、オンライン授業との直接の関連はないものの、授業形態に応じたスピーキング授業への評価を調査する本研究においては、学習者の発話抵抗感は1つの大きな個人差要因となり得ると考える。

個人差要因には含まれないものの、対面とオンラインのスピーキング授業評価を大きく分ける要因として、学習者の「習熟度」や「性別」も考えられる。Chew and Ng (2016) の研究によると、オンライン授業の方が対面授業に比べ、低い習熟度の学習者がより発言し、より積極的な授業参加を行ったという結果が報告されている。また、遠山ら (2017) は、英語習熟度の低い学生を対象に、短時間のオンライン会話グループ学習を実施した。その結果、短時間で少ない回数であっても、習熟度の低い学生のスピーキング力向上と不安軽減に効果があり、場合によっては英語学習に対する自己効力感と動機づけの向上も期待できることを示唆している。習熟度別クラス編成の意義と問題点を解明する目的で、田原ら (2001) が実施した大学生を対象にした英語授業満足度アンケートでは、上級クラスにおいて肯定的な意見が多く、普通クラスでは満足度が全体的に低いという結果が得られた。彼らはその原因として、普通クラスでは学生の能力差が大きいことを挙げ、普通クラスの細分化の可能性を示唆している。

性別と第二言語に関するパフォーマンスの関係性はかねてより議論がなされてきた (e.g., Andreou et al., 2005) が、対面とオンラインの授業評価を性別に基づいて調査したものは稀有である。その中で、Caspi et al. (2008) の対面授業とオンライン授業の比較研究では、男性は女性よりも対面授業において積極的に発言し、女性はオンラインによるチャット等を使用して積極的にコミュニケーションをとる傾向にあることが報告された。この結果の原因の一つとしては、外

国語不安 (Foreign Language Anxiety) が考えられる。例えば Mesri (2012) や Çağatay (2015) は、性別と外国語教室での不安の関係を調査し、女性のほうが男性よりも不安を強く感じていることを報告している。

### Ⅲ 目的

先行研究の整理により、コロナ禍において双方向同期型オンライン授業が注目され、学習者がオンライン授業か対面授業の選択を迫られる状況になっていることを指摘した。しかし、双方向オンライン授業において、英語のスピーキングと個人差要因の関連に特化した研究はほぼなく、学習者が自らの適性に依拠して適切な授業形態を選択することは容易ではない。

そこで、本研究の目的は対面授業とオンライン授業の比較を通じて、英語スピーキング授業に対する学習者の評価と個人差要因の関連を明らかにすることである。また、個人差要因には含まれないが授業評価の差と関連がありそうな要因として、学習者の英語習熟度と性別についても同様に調査する。

リサーチ・クエスチョン (RQs) は以下の通りである。

RQ1. 英語スピーキング授業において、オンラインと対面の形式の差によって授業評価に差はあるのか。

RQ2. RQ1 で差があった場合、それは学生の習熟度や性別とどのような関連があるか。

RQ3. RQ1 で差があった場合、それは学生の個人差要因とどのような関連があるか。

## Ⅳ 方法

### 1 調査対象者

本研究の調査対象者は、外国語 (English as a Foreign Language) として英語を学ぶ明治大学国際日本学部の1年生100名 (男性61名、女性37名、無回答2名) である。国際日本学部では英語のスピーキングの授業が必修科目であり、生徒は、入学時に実施された TOEFL ITP® の結果によって3つのレベルに分けられている。本研究では、習熟度の差を明確にするために、上位レベルと下位レベルの生徒 (上位56名、下位44名) を調査対象者とした。また、実際にスピーキング能力に差があるかどうかを確認するために、『CEFR-J リソースブック』の付録1「CEFR-J ディスクリプタ」を参考に、調査対象者に自己評価をさせた。結果は図1、2に示すとおりである。上位群はB2が15%、B1が62%、A2が23%であり、下位群はB2が2%、B1が46%、A2が52%だったことから、レベル分けの妥当性は一定程度認められるものと考えられる。

図1 上位群のスピーキング能力における自己評価 (n = 56)

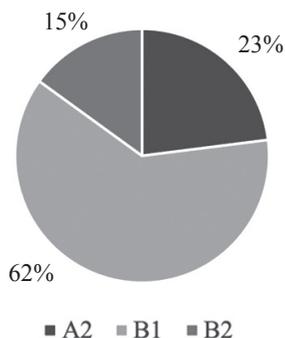
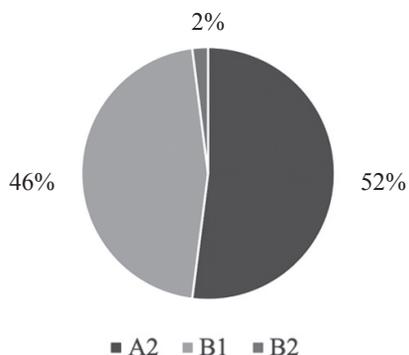


図2 下位群のスピーキング能力における自己評価 (n = 44)



## 2 調査方法

調査は2021年6月下旬から7月上旬の授業時間の一部を用いて行った。本研究ではGoogle Formで作成した質問紙を用いた（質問紙の詳細は後述）。授業中、前方スクリーンにQRコードを投影し、調査対象者自身の電子デバイスで読み取り、回答させた。質問紙は4部構成となっており、各部の調査時間を管理するために、各部で1度回答を止め、指示を出した。調査への参加は任意であり、回答は調査対象者の成績等に一切影響しない旨が伝えられた。回答時間は、合計で20分程度であった。

## 3 質問紙票

本研究では、スピーキングの授業において対面とオンラインの授業評価を、個人差要因を考慮に入れ、検討することを目的とする。したがって、質問紙は大きく「対面・オンライン授業に対する学習者評価」と「個人差要因」を測る2つのパートに分かれ、さらに、個人差要因は、測る内容によって、抵抗感、パーソナリティ、学習スタイルの3つに分かれる。したがって、①対面・オンライン授業に対する学習者評価、②抵抗感、③パーソナリティ、④学習スタイルの全4部構

成となっている。

### (1) 対面・オンラインによるスピーキング授業に対する学習者評価

スピーキング授業における「授業への参加のしやすさ」「英語の話しやすさ」について、計19項目からなる独自の質問項目を作成し、対面とオンラインそれぞれに対して学習者評価アンケートを実施した。

#### ① 授業への参加のしやすさ

「授業への参加のしやすさ」に関する測定には、星野・牟田(2003)の授業評価モデルを一部修正して使用した。この授業評価モデルでは、学生の授業に対する満足度を「教授努力」「理解度」「コミュニケーション」及び「学生の努力」の4つの要因(計14項目)で測定している。しかし、本研究の調査対象となった学生はクラスごとに異なる教員の授業を受講していたため、クラスによって授業の進め方、課題の内容も異なる。そのため、本研究では「教授努力」及び「学生の努力」に関する6項目を除いた計8項目を用いて、スピーキング授業における参加のしやすさを測定した。

#### ② 英語の話しやすさ

英語教育の分野においては、これまで「スピーキング授業における英語の話しやすさ」を測る尺度の開発が試みられたことはほほない。そこで、当該要因の測定にあたっては、遠山ら(2017)を参考にし、本研究独自の質問項目を作成した。具体的には、遠山ら(2017)では英語学習や英語使用の不安、英語学習に対する自己効力感、英語学習に関する動機づけの3つの要因に関して、計63項目を用いて測定している。本研究では、授業内のスピーキングに焦点を当てているため、英語学習や英語使用の不安、とりわけ「教室内での発話活動における不安」に関する項目の一部を、文言を変えて使用した。例えば、「教室で、英語を使って口頭発表するとき、緊張します」という項目については、口頭発表という言葉だけでは、どのような活動かイメージしにくい回答者も出てくる可能性があることを鑑み、口頭発表の部分を発表スライドの有無、また即興か事前準備有りか、といった情報を追加した。また、「教室で、英語を使ってディスカッションするとき、緊張します」という項目については、対象クラスが授業内でディスカッションを実施する頻度が少ないことから、「ペアワーク」及び「3人以上でグループワーク」とし、ペアやグループでの英語使用の状況について尋ねる質問項目に変更した。本研究では上記の手順を経て開発した計11項目を用いて、スピーキング授業における英語の話しやすさを測定した。

### (2) 個人差要因

#### ① 抵抗感

スピーキングに対する抵抗感の測定には、磯田(2008)の抵抗感尺度の項目を一部削除して、使用した。磯田は話すことへの抵抗感をWTCの概念に位置付けており、抵抗感を能力認知、不安、WTCが低いことで現れる話すことの回避の3つのパラメーターで構成している(磯田, 2008:

42)。しかし、本研究の調査環境は、授業自体すべて英語で行われている、なおかつ、話すことが授業の課題・アクティビティとなるため、話すことを回避することは困難である。したがって、磯田が作成した15項目から「話すことの回避」を除いた10項目を用いて、スピーキングに対する抵抗感を測ることとした。

## ② パーソナリティ

本研究では参加者のパーソナリティの測定に、並川ら（2012）の Big Five 尺度短縮版を使用した。パーソナリティの測定において、Big Five 尺度と同様に、多く用いられている尺度として Myers-Briggs Type Indicator (MBTI) があるが、青木（2005）では MBTI は実際の行動傾向を測ることから、MBTI の領域は学習スタイルに踏み込んでいるとしている。本研究ではパーソナリティと学習スタイルを別の構成概念として区別しているため、Big Five 尺度を採用した。また、日本で多く用いられている和田（1996）の Big Five 尺度があるが、並川ら（2012）は質問数が多いことによる回答者の負担を懸念点として挙げている。並川らの尺度は、外向性と情緒不安定性が各5項目、開放性と調和性が各6項目、誠実性が7項目の計29項目で作成されており、回答者の負担軽減によるデータの質向上に寄与していると考えられる。本研究は、全体の質問項目数が多いことから、回答者負担を考慮した並川ら（2012）の Big Five 尺度短縮版を採用した。

## ③ 学習スタイル

学習スタイルを診断するための尺度もこれまで数多く開発されている。代表的な例としては、Reid（1995, 1998）などによる Perceptual Learning Style Preference Questionnaire (PLSPQ) が挙げられる。この尺度では学習スタイルを6つのタイプ（視覚型・聴覚型・運動型・触覚型・グループ型・個人型）に分類し、各タイプ5項目（計30項目）で学習方法の好みを測定している。一方、本研究では先述したように全体の質問項目数を抑えるために、より簡易な質問紙尺度を選定する必要があった。そこで、本研究では、学習スタイルを3つのタイプ（視覚／聴覚／身体動作型、知識先行・熟考／実践先行・能動型、個別学習／グループ学習型）に分類し、計16項目でその診断を可能としている加藤・山下（2014）の「学習スタイル診断」を利用した。本尺度はウェブサイト上でも公開されており（2021年8月時点でのURLは次の通り。<http://www.kandagaigo.com/elp/style/index.html>）、スタイル別の効果的な学習方法が紹介されている点なども、本尺度を選定した理由である。

## 4 分析方法

スピーキング授業の評価に関するアンケートについては、本研究用に独自に開発した。そのため、予備分析として、質問項目（ $n = 19$ ）を整理、統合する目的で、SPSS Statistics 26を使用して主成分分析（プロマックス回転）による探索的因子分析を行った。その後、探索的因子分析で得られたモデルの適合度を分析するために Amos 26 により確認的因子分析を行った。

次に、分析の手順について説明する。①対面とオンラインの授業形式に対する評価の差を明らかにするために、対応のある  $t$  検定を行った。②その後、明らかになった授業評価の平均の差と習熟度（上位群・下位群）や性別との関連を調べるために、2元配置の分散分析（混合計画）を行った。③最後に、個人差要因と授業評価の関連を明らかにするために、まずは各個人差要因と各授業評価の相関を、ピアソンの相関分析によって検証した。その後、個人差要因の質問紙調査の結果をもとに対象者をクラスター分析（平方ユークリッド距離、ウォード法）によってグループ分けし、グループごとに詳細な検討も行った。具体的には、グループ間における授業評価に差があるのかを明らかにするために、1元配置の分散分析を実施した。

## V 結果

### 1 因子分析の結果

まず、主成分分析、プロマックス回転による探索的因子分析の過程で、因子負荷量 .40 以下の項目は削除し、かつ対面とオンラインの両方に共通した因子構造をとるように項目の選定を行った。その結果、19項目のうち7項目が削除された。残りの12項目についても、同様に探索的因子分析を行った。その結果、固有値の下限を1としたスクリープロットを基準に、因子数は4つに定められた。表1には最終的に得られたパターン行列と  $\alpha$  係数を示す。

表1 対面・オンライン形式によるスピーキング授業評価尺度の因子分析結果

質問項目	因子1	因子2	因子3	因子4
<b>因子1：授業の受けやすさ (<math>\alpha=.739/.770</math>)</b>				
Q6 スピーキングの授業に対するモチベーションがある。	.914/.757			
Q7 スピーキングの授業では、先生の声が明瞭で指示が伝わりやすい。	.708/.895			
Q3 スピーキングの授業に積極的に参加しやすい。	.662/.658			
Q1 スピーキングの授業に集中して取り組みやすい。	.593/.609			
<b>因子2：授業内での緊張・不安のなさ (<math>\alpha=.780/.644</math>)</b>				
Q11 スピーキングの授業で話す時、緊張する。		.803/.635		
Q4 スピーキングの授業中にミスをしたら、恥ずかしい。		.779/.816		
Q5 スピーキングの授業で英語を話している時、他者の目が気になる。		.756/.755		
<b>因子3：1人での発表のしやすさ (<math>\alpha=.814/.842</math>)</b>				
Q18 スピーキングの授業中、発表スライドを用いて1人で発表しやすい。			.897/.904	
Q17 スピーキングの授業中、発表スライドを用いずに1人で発言・発表しやすい。			.867/.887	
Q19 スピーキングの授業中、事前準備せずに1人で発言・発表しやすい。			.669/.774	
<b>因子4：ペアワーク・グループワークのしやすさ (<math>\alpha=.781/.816</math>)</b>				
Q15 スピーキングの授業中、3人以上でグループワークをしやすい。				.928/.888
Q14 スピーキングの授業中、ペアワークをしやすい。				.866/.893

(注)：数値は、対面・オンラインの順。Q 4, 5, 11 は逆転項目（主成分分析、プロマックス回転）。

第1因子は4項目で構成され、「先生の声が明瞭で指示が伝わりやすい」や「集中して取り組みやすい」など、授業形式によるスピーキング授業の受けやすさを示す内容の項目が高い負荷量を示していた。そこで、第1因子を「授業の受けやすさ」と命名した。第2因子は3項目で、「英

語で話すとき緊張する」や「ミスをしたら恥ずかしい」など、スピーキング授業内で生じうる緊張や不安を表す項目が高い負荷量を示した。これらの3項目は逆転項目であることを考慮し、第2因子を「授業内での緊張・不安のなさ」と命名した。第3因子は3項目で、「発表スライドを用いて1人で発言・発表しやすい」や「事前準備せずに1人で発言・発表しやすい」など、1人で大勢の前で話す内容の項目が高い負荷量を示した。そこで、第3因子を「1人での発表のしやすさ」と命名した。第4因子は2項目で、ペアワークやグループワークのしやすさに関する内容の項目が高い負荷量を示した。そこで、第4因子を「ペアワーク・グループワークのしやすさ」と命名した。

さらに、上記で得られた4因子計12項目の尺度に関して、対面、オンラインのそれぞれで確認的因子分析を行い、適合度 (goodness of fit) に基づいてモデルを評価した。適合度とは、分析で得られた因子構造が実際のデータとどれほど適合しているかを示すものである。その結果は表2に示す通りであり、West et al. (2012) の基準をもとにすると、モデルの適合度は概ね許容できる値であった。以上の結果から、のちに続く分析では、これら12項目から構成される4因子を使用することとする。

表2 対面・オンラインの各モデルにおける適合度指標

モデル	CMIN/DF	CFI	RMSEA	GFI	AGFI
対面	1.312	.962	.056	.908	.850
オンライン	1.249	.969	.050	.909	.852

## 2 RQ1：対象者全体における対面とオンラインの授業評価の差

表3のとおり、スピーキング授業における対面形式とオンライン形式の授業評価の差を対応のある  $t$  検定により検討した。有意確率については、 $p = .05$  に設定した。その結果、「授業の受けやすさ」「緊張・不安のなさ」「1人での発表のしやすさ」の3変数において統計的に有意な差が見られた。平均の差が示す通り、「授業の受けやすさ」に関しては対面形式が、「緊張・不安のなさ」「1人での発表のしやすさ」に関してはオンライン形式が好まれるという結果になった。

表3 対面・オンライン形式のスピーキング授業評価における  $t$  検定の結果 ( $n = 100$ )

変数	対面		オンライン		$t$	平均の差 (対面-オンライン)
	平均	標準偏差	平均	標準偏差		
授業の受けやすさ	4.20	0.65	3.75	0.75	5.81 ***	0.45
緊張・不安のなさ	2.59	1.11	2.85	0.97	-2.85 **	-0.26
1人での発表のしやすさ	2.77	0.92	3.06	0.97	-4.77 ***	-0.29
ペアワーク・グループワークのしやすさ	4.30	0.90	4.09	0.92	1.54	0.21

(注) : \*\*  $p < .01$ , \*\*\*  $p < .001$

### 3 RQ2：習熟度・性別と授業評価の関連

スピーキング授業における対面とオンラインの授業評価が参加者の習熟度（下位群・上位群）によって異なるかどうかを、2元配置の分散分析（混合計画）により検討した。表4には、両者のクロス集計表を示す。分散分析の結果、授業評価と習熟度の交互作用はいずれも有意ではなく（ $p > .05$ ）、習熟度の主効果についても同様だった（ $p > .05$ ）。したがって、参加者の習熟度によって、どのような授業形式を好むかは、対面、オンラインのいずれにおいても違いはないことが示唆された。このような結果が得られた原因については、のちにより詳しく考察する。

表4 学習者の授業評価と習熟度のクロス集計表

変数		下位群 (n = 44)		上位群 (n = 56)	
		平均	標準偏差	平均	標準偏差
授業の受けやすさ	対面	4.20	0.59	4.21	0.70
	オンライン	3.73	0.77	3.76	0.75
緊張・不安のなさ	対面	2.66	1.11	2.53	1.12
	オンライン	2.89	0.99	2.82	0.97
1人での発表のしやすさ	対面	2.89	0.91	2.68	0.93
	オンライン	3.20	0.94	2.95	0.99
ペアワーク・グループワークのしやすさ	対面	4.23	0.85	4.36	0.94
	オンライン	4.02	1.02	4.14	0.85

つぎに、上記と同じ要領で、スピーキング授業における対面とオンラインの授業評価が性別によって異なるかどうかを、2元配置の分散分析（混合計画）により検討した。その際、性別を「回答しない」と答えた2名の参加者は除いた。表5には、両者のクロス集計表を示す。分散分析の結果、授業評価と性別の交互作用はいずれも有意ではなく（ $p > .05$ ）、性別の主効果は「1人での発表のしやすさ」のみが有意であった（ $F [1, 96] = 9.99, p = .00$ ）。この結果を図示すると、図3の通りであった。これらの結果から明らかのように、ペアやグループにおける発表のしやすさについては対面、オンラインとも男女差が見られなかった一方、1人での発表については対面、オンラインともに男性の方が女性に比べて、より肯定的に捉えていること、つまり1人の方が発言をしたり発表がしやすいと認識していることが示唆された。

表5 学習者の授業評価と性別のクロス集計表

変数		男性 (n = 37)		女性 (n = 61)	
		平均	標準偏差	平均	標準偏差
授業の受けやすさ	対面	4.12	0.80	4.28	0.52
	オンライン	3.73	0.84	3.76	0.72
緊張・不安のなさ	対面	2.77	1.15	2.48	1.07
	オンライン	3.12	0.98	2.71	0.95
1人での発表のしやすさ	対面	3.17	0.86	2.53	0.90
	オンライン	3.37	0.96	2.87	0.95
ペアワーク・グループワークのしやすさ	対面	4.22	0.92	4.40	0.85
	オンライン	4.07	0.99	4.11	0.90

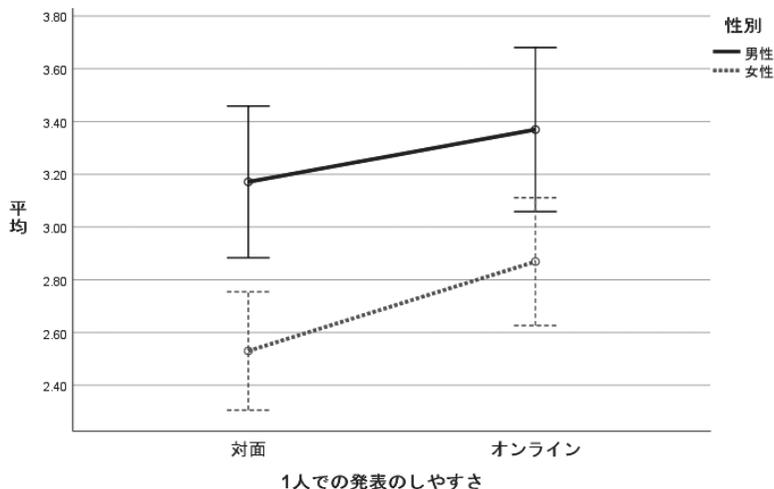


図3 学習者の授業評価（1人での発表のしやすさ）と性別の関係

(注)：エラーバーは95%信頼区間

#### 4 RQ3：個人差要因と授業評価の関連

##### (1) 相関分析

個人差要因（例：抵抗感・パーソナリティ・学習スタイル）と対面とオンラインの授業評価との関連を、相関分析によって検証した。表6はその結果を示している。まず特筆すべきは英語スピーキングに対する抵抗感と授業評価との関係である。抵抗感と「授業の受けやすさ（対面）」「緊張・不安のなさ（対面・オンライン）」「1人での発表のしやすさ（対面・オンライン）」において、弱～中程度の負の相関（ $r = -.29 \sim -.64$ ）が見られた。言いかえると、スピーキングに対して抵抗が強いほど授業は受けやすく、心地よく、話しやすいという結果になった。

次に、パーソナリティに焦点を当てると、外向性と情緒不安定性において授業評価との間に有意な相関があった。まず、外向性と「授業の受けやすさ（対面）」「緊張・不安のなさ（対面）」「1人での発表のしやすさ（対面・オンライン）」との間に、弱～中程度の正の相関（ $r = .21 \sim .27$ ）が見られた。また、情緒不安定性と「授業の受けやすさ（対面）」「緊張・不安のなさ（対面）」「1人での発表のしやすさ（対面）」「ペアワーク・グループワークのしやすさ（対面）」との間に、弱～中程度の負の相関（ $r = -.25 \sim -.43$ ）が見られた。つまり、外向的な性格の学習者は対面でのスピーキング授業に対して評価が高く、神経質な性格の学習者はオンラインの形式を好むという結果となった。その他のパーソナリティの項目では、調和性と「授業の受けやすさ（対面）」の間に弱い正の相関（ $r = .28$ ）があった。

最後に、学習スタイルと授業評価の相関について見ていくことにする。これらの間には無相関が多かったものの、授業評価の項目である「緊張・不安のなさ（対面）」と「聴覚型」「身体動作型」「実践先行型」の間に弱い正の相関（ $r = .22 \sim .30$ ）、同項目と「知識先行型」の間には弱い負の相関（ $r = -.29$ ）が確認された。

表6 授業評価と個人差要因の相関

変数	抵抗感	パーソナリティ					学習スタイル						
		外向性	誠実性	情緒不安定性	開放性	調和性	視覚型	聴覚型	身体動作型	知識先行型	実践先行型	個別学習型	グループ学習型
<b>対面</b>													
授業の受けやすさ	<b>-.29**</b>	<b>.23*</b>	.02	<b>-.27**</b>	.12	<b>.28**</b>	-.14	.03	.01	-.17	.18	.01	.06
緊張・不安のなさ	<b>-.64**</b>	<b>.21*</b>	.05	<b>-.43**</b>	.05	.07	-.12	<b>.25*</b>	<b>.22*</b>	<b>-.29**</b>	<b>.30**</b>	.07	.00
1人での発表のしやすさ	<b>-.40**</b>	<b>.27**</b>	.15	<b>-.25*</b>	.09	.19	-.11	.13	.06	-.18	.17	-.13	.07
ペアワーク・グループワークのしやすさ	-.19	.14	-.12	<b>-.27**</b>	-.05	.19	-.07	.03	.01	-.08	.05	-.02	.02
<b>オンライン</b>													
授業の受けやすさ	-.19	.05	.18	-.18	.07	.13	.07	-.08	.11	-.04	.03	-.06	.09
緊張・不安のなさ	<b>-.43**</b>	.09	.01	-.16	.03	.12	-.14	.15	.07	-.13	.12	-.06	.08
1人での発表のしやすさ	<b>-.38**</b>	<b>.24*</b>	.17	-.17	-.01	.14	-.06	.09	.06	-.14	.12	-.14	.10
ペアワーク・グループワークのしやすさ	.01	.15	.04	-.07	.17	.16	.12	-.08	-.06	.12	-.13	.03	-.06

(注)：有意な相関係数は太字。\*  $p < .05$ , \*\*  $p < .01$

(2) クラスタ分析

ここまで、個人差要因と授業評価の関係を見てきた。しかし、個人差要因の変数が多いため、結果をまとめて一般化した主張をするには全体的な相関を検討するだけでは不十分だと考えられる。そこで、対象者を個人差要因に応じてグループ分けをする目的で、クラスタ分析（平方ユークリッド距離，ウォード法）を行った。これにより、個人差要因の傾向や特性がどのように授業評価と関係しているかを明らかにすることが可能になる。ウォード法を用いた理由は、この方法によるクラスタ分析は、比較的まとまったクラスタが得られやすくパターン分類に有用であると考えたためである。

その結果、図4のような結果を得た。結合距離の変化、及び得られるクラスタの特徴を吟味して、図4の点線で示されるところにカットングポイントを定め、3つのクラスタに分類するのが妥当と判断した。各クラスタの所属人数は、第1クラスタが37人、第2クラスタが31人、第3クラスタが32人であった。クラスタ分析によって得られた結果の妥当性を、

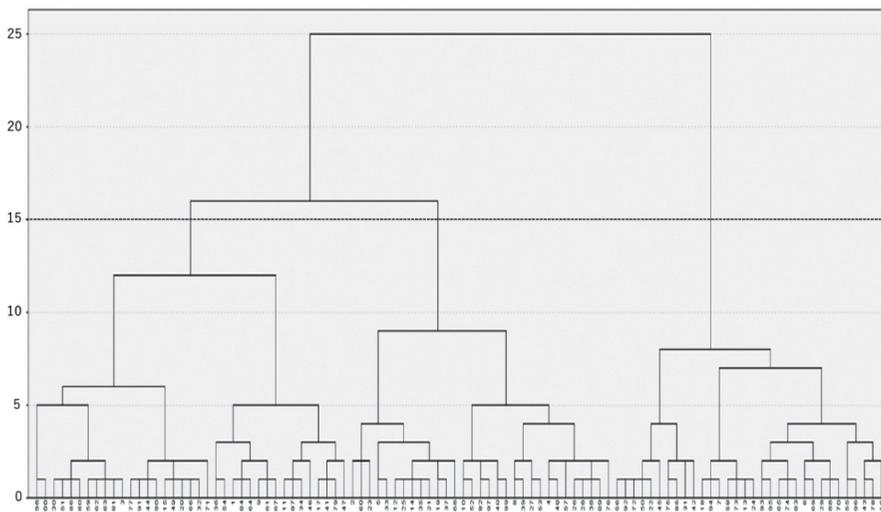


図4 クラスタ分析に基づくデンドログラム

個人差要因における各下位尺度において一元配置の分散分析を行った。その結果、個別学習型 ( $p = .15$ ) を除いて、各クラスター間に有意な差があることが確認された ( $p < .05$ )。このようにして得られたそれぞれのクラスターの個人差要因を図5に示す。

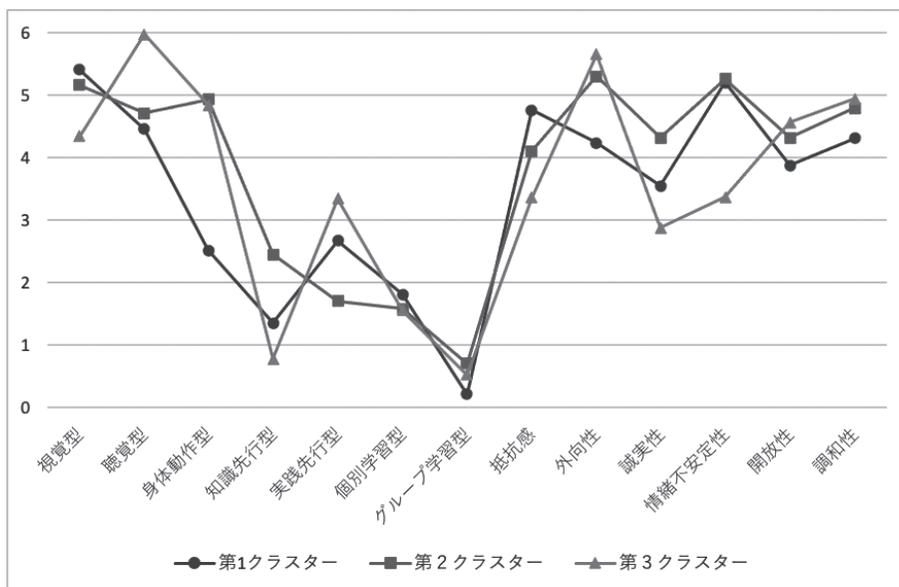


図5 各クラスターの個人差要因

以下では、各クラスターの特徴について説明する。まず、第1クラスター ( $n = 37$ ) では学習スタイルにおいて視覚型のインプット傾向が強く、身体動作型の傾向が低く、他のクラスターより、個別学習を好む傾向がある。また、抵抗感が高く、パーソナリティの項目においても、外向性、開放性、調和性が低く、情緒不安定性が高いことから、ペアワークやグループワークといった他者との交流を通じた学習を好まない傾向にある。従って、第1クラスターは1人で視覚型のインプットをする学習を好み、神経質であり、アウトプットであるスピーキングに対して抵抗感があるため、他者と交わる学習を好まない傾向がある「個別インプット群」と呼ぶことにした。

第2クラスター ( $n = 31$ ) は学習スタイルにおいて視覚型、聴覚型、身体動作型の平均値が高いことから、バランスの取れたインプット傾向で、知識先行型が高いため、理論を学び知識を身につけてから実践に移す学習スタイルを好む傾向にある。パーソナリティの項目では、外向性、誠実性、調和性、情緒不安定性が高い。従って、神経質ではあるが、勤勉であり、他者との交流を通して、知識をアウトプットする傾向にある。よって、第2クラスターは勤勉で、バランスが取れたインプットでアウトプットも行う「バランス群」と呼ぶこととする。

第3クラスター ( $n = 32$ ) では学習スタイルにおいて聴覚型の平均値が高く、また知識先行型よりも実践先行型の平均値が顕著に高い。パーソナリティの項目を見ると、外向性が高い一方で、誠実性、情緒不安定性が低い。また、3クラスターの中で最も話すことに対する抵抗感が低い。

従って、スピーキングに抵抗感がなく、個人で勤勉に学習するよりも、ペアやグループのアクティビティを行いながら学ぶ傾向が強いことから、「グループアウトプット群」と呼ぶこととする。

つぎに、それぞれのクラスターが授業評価においてどのような傾向にあるかを示し、クラスター間の対面・オンラインそれぞれの授業評価の平均の差を検証するために、1元配置分散分析を行った。その結果を表7に示す。

表7 クラスターごとの平均値、標準偏差と分散分析の結果

変数	対面			オンライン		
	平均	標準偏差	F	平均	標準偏差	F
授業の受けやすさ						
第1クラスター (n=37)	4.05	0.61		3.65	0.69	
第2クラスター (n=31)	4.10	0.51	2.89	3.90	0.58	0.92
第3クラスター (n=32)	4.34	0.43		3.73	0.95	
緊張・不安のなさ						
第1クラスター (n=37)	2.20	0.83		2.66	0.86	
第2クラスター (n=31)	2.41	1.05	9.25***	2.82	0.95	1.97
第3クラスター (n=32)	3.22	1.19		3.11	1.08	
1人での発表のしやすさ						
第1クラスター (n=37)	2.51	0.82		2.83	0.89	
第2クラスター (n=31)	2.75	0.84	3.58*	3.10	0.84	2.11
第3クラスター (n=32)	3.09	1.03		3.30	1.13	
ペアワーク・グループワークのしやすさ						
第1クラスター (n=37)	4.08	1.08		4.19	0.72	
第2クラスター (n=31)	4.18	0.81	4.41*	4.16	0.83	0.94
第3クラスター (n=32)	4.67	0.60		3.91	1.19	

(注) : \*  $p < .05$ , \*\*\*  $p < .001$

対面授業においては、「緊張・不安のなさ」「1人での発表のしやすさ」「ペアワーク・グループワークのしやすさ」でクラスター間に有意な差があった。その後、どのクラスター間に差があったのかを検討するために、Tukeyの多重比較を行った。その結果、「緊張・不安のなさ」では個別インプット群（第1）＜グループアウトプット群（第3）、バランス群（第2）＜グループアウトプット群（第3）、「1人での発表のしやすさ」では個別インプット群（第1）＜グループアウトプット群（第3）、「ペアワーク・グループワークのしやすさ」では個別インプット群（第1）＜グループアウトプット群（第3）であった（ $p < .05$ ）。このことから、対面授業においては、グループアウトプット群のような個人差要因を持つ学習者が高い評価を示したと言える。一方、オンライン授業においては、すべての変数においてクラスター間に有意な差は確認されなかった。

最後に、クラスターごとに、対面とオンラインの授業形式の違いによって授業評価に差があるのかどうかを検証した。具体的には、各クラスターの対面・オンラインの授業評価における平均値の差を対応のある  $t$  検定により検討した。

個別インプット群では、対面とオンラインの授業評価に差が見られたのは、「授業の受けやすさ」( $t(36) = 3.66, p = .00$ ), 「緊張・不安のなさ」( $t(36) = -3.43, p = .00$ ), 「1人での発表のしやすさ」( $t(36) = -4.02, p = .00$ ) の3変数であった。それらの中では、「授業の受けやすさ」のみ対面の評価が高く、他の2つの変数はオンラインの評価が高いという結果となった。バランス群で対面とオンラインの評価に差が見られたのは、「1人での発表のしやすさ」( $t(30) = -3.30, p = .00$ ) であり、オンラインの評価が高かった。グループアウトプット群で対面とオンラインの評価に差が見られたのは、「授業の受けやすさ」( $t(31) = 3.73, p = .00$ ), 「1人での発表のしやすさ」( $t(31) = -1.55, p = .00$ ) の2変数であった。平均値を比較すると、「授業の受けやすさ」は対面、「1人での発表のしやすさ」はオンラインの評価が高かった。

## V 考察

本論文の目的は、対面とオンラインの比較をもとに、英語スピーキング授業に対する学習者の評価と個人差要因の関連を明らかにすることであった。そのため、日本人大学生を対象に、独自に開発した授業評価アンケートと、妥当性と信頼性の担保された個人差要因（抵抗感・パーソナリティ・学習スタイル）アンケートを実施し、統計的に分析した。その結果、対面・オンラインの英語スピーキング授業の評価は学習者によって様々であり、それらは個人差要因と大きく関連していることが明らかになった。一方、それらは学習者の習熟度や性別とは密接な関連があるとは言えない結果となった。

次に、得られた研究結果を、先行研究との比較を交えて RQ ごとに考察する。

RQ1：英語スピーキング授業において、オンラインと対面の形式の差によって授業評価に差はあるのか。

「授業の受けやすさ」に関して、参加者全体を通じて対面形式のスピーキング授業の方が好まれるという結果になった。この結果が得られた要因の1つとして、オンライン授業に関して教師、学生ともにまだノウハウが蓄積されていないことが挙げられる。COVID-19 の感染拡大防止に伴い、大学では早急なオンライン授業への移行が行われた。教師はオンライン授業を運営するにあたっての技量や経験が少ないことが想定されるため、対面授業と比較してオンライン授業の実施には困難が生じていると考えられる。加えて、学生側は教室で受講する対面授業形式とは異なる在宅での受講になる傾向があるため、オンライン授業では動機づけの維持や集中力の持続が難しいことも考慮すべきだと言える。

一方、「緊張・不安のなさ」「1人での発表のしやすさ」に関しては、オンライン形式が好まれるという結果になった。これは、オンラインでのコミュニケーションは対面に比べて、気軽に緊張感の少ない「対人圧力」の弱い状態であると主張する木村・都築（1998）を支持するものである。また、吉田（1999）はオンライン授業における発言回数が対面に比べて増加したこと、遠山ら（2017）はオンライン授業でのスピーキングにおいて、不安が軽減したことについて言及している。これらは双方向のコミュニケーションに焦点を当てた研究であったが、本研究により、1

人で発表することに焦点を当てた一方コミュニケーションにおいても、対面授業と比較してオンライン授業では緊張・不安が軽減する傾向にあることが明らかになった。

RQ2：RQ1で差があった場合、それは学生の習熟度や性別とどのような関連があるか。

対面・オンラインのスピーキング授業評価において、習熟度が評価の差に影響を与えているとは言い難い結果が得られた。これは英会話の授業において、習熟度がより高く、学習内容の難易度の高いクラスに所属する学生の授業評価が高くなることを明らかにした田原ら（2001）と一致しない結果となった。この理由の1つとしては、習熟度編成の妥当性が低いことが挙げられる。本研究の対象者は私立大学の国際系学部所属の学生であったため、元々外国語に対して好意的で、特に英語を得意とする学生の割合が高いと考えられる。また、習熟度を分ける際にはリスニングとリーディングのスキルを測定するTOEFL ITP®の結果を基にしていたため、本研究で焦点を当てたスピーキング力が習熟度に加味されていない。この点は今後、さらに検討を重ねるべきだと考えられる。

また、性別に関しては、「1人での発表のしやすさ」において、対面とオンラインの両方で男性の方が女性より授業評価が高かった。表5からも明らかのように、対面、オンラインに関わらず、男性は1人での発表（対面：3.17；オンライン：3.37）とペア・グループでの発表（対面：4.22；オンライン：4.07）のいずれにも好意的で、どちらかと言えば後者の方がその傾向は顕著だった。一方で、女性の場合は1人での発表（対面：2.53；オンライン：2.87）とペア・グループでの発表（対面：4.40；オンライン：4.11）の評価差はより大きく、前者に対しては必ずしも前向きに捉えているわけではないことがわかる。こうした点は、実際にスピーキングの授業をデザインする際に留意するべきであろう。

RQ3：RQ1で差があった場合、それは学生の個人差要因とどのような関連があるか。

RQ3に関して、本研究の結果をそれぞれの個人差要因の観点から考察する。まず、学習者の発話抵抗感を対面・オンラインの授業評価と関連付けた研究はほぼない。本研究では、抵抗感が高いグループ（具体的には、個別インプット群）は低いグループ（グループアウトプット群）と比較して、対面よりもオンラインの方が授業を受けやすいと認識していることが明らかとなった。磯田（2008, 2009）では対面授業において、SPMを用いて学習者の発話に対する抵抗感を軽減させるという取り組みについて報告している。本研究の結果から、教師は抵抗感の高い学習者に対してはあらかじめオンライン授業の中でSPMを実施することで、その軽減を促すことができる可能性が示唆される。また、学習者の目線に立つと、発話に対して抵抗を強く感じている学習者は、オンライン授業の方が全体的に授業評価は高くなることも予想される。

パーソナリティに関して、本研究の結果からは大きく次の2点を指摘することができる。第一に、情緒の安定性が低い学習者は、対面よりもオンラインの授業の方を相対的に受けやすく感じることである。第二に、外向的で調和的な学習者は、オンラインよりも対面の授業を高く評価する傾向にあるということである。パーソナリティと対面・オンラインの両方の授業評価の関連を

調べた研究は稀有であり、本研究の結果は今後の研究に対して具体的な示唆を与えるものと考えられる。一方で、先行研究との比較でいえば、外向性と調和性が高い学生（本研究でいえば、グループアウトプット群）はオンラインで実施されるグループワークに対して肯定的なレビューを持つと報告した Borg et al. (2021) とは必ずしも一致しない結果となった。この理由として、本研究は対象者に対面・オンラインの両方の授業形態における評価を求めたため、彼らは必然的にそれぞれの形態を相対的に比較する形となったことが考えられる。したがって、今後は対面・オンラインをそれぞれ別々に取り上げ、調査・比較するなど別のアプローチについても検討する必要があるだろう。

最後に、学習スタイルについて、これまでの多くの研究は学業成績との関連を検討したものがほとんどであり、学習スタイルと授業評価との関連を調査した研究はほぼない。そのような現状において、本研究は学習スタイルを含め、類似した個人差要因を有する学習者をグループ分けし、対面・オンラインの授業評価をそのグループごとに比較・検討した数少ない研究である。結果として、表7からも明らかなように、学習スタイルを含めた個人差要因は、オンラインよりも対面においてその違いが顕著に確認された。この違いは両授業形態のいわゆる特性に起因するものなのか、あるいは先述したように学生自身がオンラインでの授業に慣れていなかったことによるものなのかは、今回の調査結果からは定かではない。今後、類似した研究の成果を積み重ねることで、どのような学習スタイルを有する学習者にはどういった授業形態が有効なのか、より具体的な示唆が可能になるものとする。

## VI 結論と課題

本研究では、個人差要因が対面・オンラインの授業評価の差に影響を及ぼしており、学習者が自分自身の特性を正しく理解することで、授業形態を選択する際の手助けになる可能性があることを示した。しかし、本研究には以下のような課題も存在する。第一に、授業評価アンケートにおける理論的裏付けの欠如である。つまり、本研究では、対面とオンライン授業を比較する際に差が出るであろうと予測した項目を先行研究より選択し、尺度を開発したため、因子分析を実施したとはいえ更なる改良が必要であろう。第二に、本研究では習熟度が対面・オンライン授業の評価に影響を及ぼすとは言えない結果となったが、習熟度の分類自体にも課題があると言える。なぜなら、本研究の参加者は1つの大学の100名の学部生のみを対象としており、英語力に応じてクラス分けがなされているとは言え、彼らの基礎力はクラス間で有意な差があるとは言いきれないからである。そして第三に、今回は質問紙による量的な調査のみを実施したが、データのより深い考察のために質的なものと統合すべきとの見方もある（American Psychological Association, 2020）ことを考慮すると、どのような個人差要因を有する学習者がどのような理由で対面あるいはオンライン授業を選択するのかは完全に明らかになったとは言えない。

したがって、今後の研究では、理論的枠組みに基づいた質問紙を用い、様々な大学レベルからより多くの被験者を対象に同様の調査を実施する必要がある。また、状況に応じてグループ・

フォーカス・インタビューなどの手法を用い、特定の個人差要因を持つ学習者が特定の授業形態を選択する理由をさらに理解することも視野に入れるべきだと考える。加えて、本研究は個人差要因を3つに限定したが、ほかの個人差要因でも比較・検討したり、スピーキング授業ではなく他のスキルを対象にしたりしても有意義な結果が得られるであろう。

以上のような課題を残すものの、大学生の個人差要因を複合的に捉え、対面・オンラインの両方の形式に対する授業評価を比較・検討した本研究は、COVID-19収束後も学習者の授業選択の一基準として大きな意味を持つだろう。比較的新しい分野の研究であることを考えると、その実態についてさらに理解を深めることは、大学生の授業選択の一助のみならず教師側の指導実践においても有益な示唆を与えることになり、英語教育の分野において重要な知見を提供できるものと考えられる。

#### 注

本論は、明治大学大学院国際日本学研究科における講義科目「英語教育学研究（心理言語学）」（担当教員：廣森）を受講する院生が、講義内容に基づき共同で行った調査の結果をまとめ、報告したものである。廣森はプロジェクト全体の統括、金田は受講者全体の取りまとめ役を担い、両者を含めた全員ですべてのセクションを共同執筆するとともに、推敲作業を行った。その意味で、本論は担当教員と受講者とのコラボレーションの成果である。なお、本論における不備、遺漏はすべて筆頭著者にその責があることを申し添えます。

## 参考文献

- 青木久美子 (2005). 「学習スタイルの概念と理論 - 欧米の研究から学ぶ」 *メディア教育研究*, 2(1), 197-212.
- 磯田貴道 (2008). 「英語スピーキング抵抗感尺度の作成」 *広島外国語教育研究*, 11, 41-49. <https://doi.org/10.15027/24578>
- 磯田貴道 (2009). 「英語でのスピーキングに対する抵抗感の軽減」 *JACET Journal*, 48, 53-66. <https://doi.org/10.15027/24807>
- 植原啓介 (2020). 「慶應 SFC における遠隔授業とアンケート調査結果」 [https://www.nii.ac.jp/event/upload/20200605-5\\_Uehara.pdf](https://www.nii.ac.jp/event/upload/20200605-5_Uehara.pdf) (参照日：2021.09.05)
- 大嶋秀樹 (2021). 「新型コロナウイルス感染拡大の状況下での大学でのリアルタイム・オンライン英語授業の試み」 *滋賀大学教育実践研究論集*, 3, 63-70.
- 大谷杏 (2021). 「新型コロナウイルスの影響による大学の英語オンライン授業」 *関西英語教育学会紀要*, 44, 21-39. [https://doi.org/10.18989/selt.44.0\\_21](https://doi.org/10.18989/selt.44.0_21)
- 加藤聡子・山下尚子 (2014). 「学習スタイル診断」 <http://www.kandagaigo.com/elp/style/index.html> (参照日：2021.08.10)
- 株式会社アルク (2016). 「アルク英語教育実態レポート Vol.7」 [https://cdn2.alc.co.jp/sa/www/report/alc\\_report\\_20160627.pdf](https://cdn2.alc.co.jp/sa/www/report/alc_report_20160627.pdf) (参照日：2021.09.05)
- 木村泰之・都築誉史 (1998). 「集団意思決定とコミュニケーション・モード：コンピュータ・コミュニケーション条件と対面コミュニケーション条件の差異に関する実験社会心理学的検討」 *実験心理学研究*, 38(2), 183-192. <https://doi.org/10.2130/jjesp.38.183>
- 田原良子・堀江美智代・竹内光悦 (2001). 「習熟度別クラス編成に関する考察 (1)」 *鹿児島純心女子短期大学研究紀要*, 31, 215-244.
- 遠山道子・森一将・新谷真由 (2017). 「オンライン英会話グループ学習を用いたスピーキング技能と心理的要因の改善：英語リメディアル教育への適用に向けて」 *外国語教育メディア学会関東支部研究紀要*, 1, 31-59. [https://doi.org/10.24781/letkj.1.0\\_37](https://doi.org/10.24781/letkj.1.0_37)
- 東洋大学 (2020). 「コロナ禍対応のオンライン講義に関する学生意識調査」 <https://www.toyo.ac.jp/-/media/Images/Toyo/research/labo-center/gensha/research/52395/1questionnaire.ashx?la=ja-JP&hash=C36CFE9B7AD656C60987AAB3BE92B314052C9E19> (参照日：2021.08.31)
- 並川努・谷伊織・脇田貴文・熊谷龍一・中根愛・野口裕之 (2012). 「Big Five 尺度短縮版の開発と信頼性と妥当性の検討」 *心理学研究*, 83(2), 91-99. [https://www.jstage.jst.go.jp/article/jjpsy/83/2/83\\_91/\\_pdf/-char/ja](https://www.jstage.jst.go.jp/article/jjpsy/83/2/83_91/_pdf/-char/ja)
- 古家聡・櫻井千佳子 (2014). 「英語に関する大学生の意識調査と英語コミュニケーション能力育成についての一考察」 *武蔵野大学教養教育リサーチセンター紀要*, 4, 29-50.
- 星野敦子・牟田博光 (2003). 「大学生による授業評価にみる受講者の満足度に影響を及ぼす諸要因」

- 日本教育工学会論文誌, 27, 213-216. [https://www.jstage.jst.go.jp/article/jmet/27/suppl/27\\_KJ00003905761/\\_pdf/-char/ja](https://www.jstage.jst.go.jp/article/jmet/27/suppl/27_KJ00003905761/_pdf/-char/ja)
- 前川洋子 (2020). 「コロナ禍における学生の参加を促す双方向 Zoom 授業の実践」岡山理科大学教育実践研究, 4, 139-151.
- 文部科学省 (2020a). 「新型コロナウイルス感染症の状況を踏まえた大学等の授業の実施状況」[https://www.mext.go.jp/content/20200605-mxt\\_kouhou01-000004520\\_6.pdf](https://www.mext.go.jp/content/20200605-mxt_kouhou01-000004520_6.pdf) (参照日: 2021.08.22)
- 文部科学省 (2020b). 「新しい時代の初等中等教育の在り方特別部会 (第11回) 参考資料4」[https://www.mext.go.jp/kaigisiryu/content/20200717-mxt\\_syoto02-000008662\\_44.pdf](https://www.mext.go.jp/kaigisiryu/content/20200717-mxt_syoto02-000008662_44.pdf) (参照日: 2021.09.04)
- 文部科学省 (2021). 「新型コロナウイルス感染症の影響による学生等の学生生活に関する調査」[https://www.mext.go.jp/content/20210525-mxt\\_kouhou01-000004520\\_1.pdf](https://www.mext.go.jp/content/20210525-mxt_kouhou01-000004520_1.pdf) (参照日: 2021.08.22)
- 立教大学 (2020). 「立教大学経営学部オンライン授業に関する学生意識調査」<https://www.rikkyo.ac.jp/news/2020/09/mknpps000001bg3b.html> (参照日: 2021.08.31)
- 吉田雅巳 (1999). 「大学授業における対面グループ指導と遠隔グループ指導における交流の比較研究」日本教育工学会誌, 23, 29-32. [https://www.jstage.jst.go.jp/article/jmet/23/suppl/23\\_KJ00003905408/\\_pdf/-char/ja](https://www.jstage.jst.go.jp/article/jmet/23/suppl/23_KJ00003905408/_pdf/-char/ja)
- 和田さゆり (1996). 「性格特性用語を用いた Big Five 尺度の作成」心理学研究, 67(1), 61-67. [https://www.jstage.jst.go.jp/article/jjpsy1926/67/1/67\\_1\\_61/\\_pdf/-char/ja](https://www.jstage.jst.go.jp/article/jjpsy1926/67/1/67_1_61/_pdf/-char/ja)
- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). <https://doi.org/10.1037/0000165-000>
- Andreou, G., Vlachos, F., & Andreou, E. (2005). Affecting factors in second language learning. *Journal of Psycholinguistic Research*, 34(5), 429-438. <https://doi.org/10.1007/s10936-005-6202-0>
- Borg, M. E., Butterfield, K. M., Wood, E., Zhang, H. H., & Pinto, S. (2021). Investigating the impacts of personality on the use and perceptions of online collaborative platforms in higher education. *SN Social Sciences*, 1(40), 1-22. <https://doi.org/10.1007/s43545-020-00053-x>
- Çağatay, S. (2015). Examining EFL students' foreign language speaking anxiety: The case at a Turkish state university. *Procedia - Social and Behavioral Sciences*, 199, 648-656. <https://doi.org/10.1016/j.sbspro.2015.07.594>
- Caspi, A., Chajut, E., & Saporta, K. (2008). Participation in class and in online discussions: Gender differences. *Computers & Education*, 50(3), 718-724. <https://doi.org/10.1016/j.compedu.2006.08.003>
- Chew, S. Y., & Ng, L. L. (2016). The relevance of personality and language proficiency on the participation style of ESL Learners in face-to-face and online discussions. *The Asia-Pacific*

- Education Researcher*, 25, 605-613. <https://doi.org/10.1007/s40299-016-0288-z>
- Dag, F., & Gecer, A. (2009). Relations between online learning and learning styles. *Procedia Social and Behavioral Sciences*, 1(1), 862-871. <https://doi.org/10.1016/j.sbspro.2009.01.155>
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge. <https://doi.org/10.4324/9781410613349>
- ETS (English Testing Service) (2019). *TOEFL iBT® test and score data summary 2019*. [https://www.ets.org/s/toefl/pdf/94227\\_unlweb.pdf](https://www.ets.org/s/toefl/pdf/94227_unlweb.pdf). (参照日 : 2021.08.31)
- Fortune, M., Spielman, M., & Pangelinan, D. (2011). Students' perceptions of online or face-to-face learning and social media in hospitality, recreation and tourism. *Journal of Online Learning and Teaching*, 7(1), 1-16. [https://jolt.merlot.org/vol7no1/fortune\\_0311.pdf](https://jolt.merlot.org/vol7no1/fortune_0311.pdf)
- Goodridge, D. H., Lawanto, O., & Santoso, H., B. (2017). A learning style comparison between synchronous online and face-to-face engineering graphics instruction. *International Education Studies*, 10(2), 1-14. <https://doi.org/10.5539/ies.v10n2p1>
- Gregersen, T., & MacIntyre, P. D. (2014). *Capitalizing on language learners' individuality: From premise to practice*. Multilingual Matters. <https://doi.org/10.21832/9781783091218>
- Johnson, S. D., Aragon, S. R., Shaik, N., & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. *Jl. of Interactive Learning Research*, 11(1), 29-49.
- Kolb, D. A. (1984). *Experiential learning, experience as the source of learning and development*. Prentice Hall.
- Martin, J. K. (2020). Exploring EFL students' perceptions of the use of synchronous video conferencing to develop their speaking competence. *Seigakuin University Journal*, 33(1,2), 161-180.
- Mesri, F. (2012). The relationship between gender and Iranian EFL learners' foreign language classroom anxiety (FLCA). *International Journal of Academic Research in Business and Social Sciences*, 2(6), 147-156. <https://doi.org/10.11648/j.ijll.s.2014020601.16>
- Ng, C. H. (2020). Communicative language teaching (CLT) through synchronous online teaching in English language preservice teacher education. *International Journal of TESOL Studies*, 2(2), 62-73. <https://doi.org/10.46451/ijts.2020.09.06>
- Nurwahyuni, R. (2020). An analysis of students' perception on synchronous and asynchronous learning in speaking skill during pandemic Covid-19. *Institut Agama Islam Negeri (IAIN) Palangka Raya Indonesia*, 189-201.
- Reid, J. M. (1995). *Learning styles in the ESL/EFL classroom*. Heinle & Heinle Publishers.
- Reid, J. M. (1998). *Understanding learning styles in the second language classroom*. Prentice Hall Regents.

- Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Routledge.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Spencer, D., & Temple, T. (2021). Examining students' online course perceptions and comparing student performance outcomes in online and face-to-face classrooms. *Online Learning Journal*, 25(2), 233-261. <http://dx.doi.org/10.24059/olj.v25i2.2227>
- Weiser, O., Blau, I., & Eshet-Alkalai, Y. (2018). How do medium naturalness, teaching-learning interactions and students' personality traits affect participation in synchronous E-learning? *The Internet and Higher Education*, 37, 40-51. <https://doi.org/10.1016/j.iheduc.2018.01.001>
- West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209-231). Guilford press.

## 付録：オンラインと対面の英語授業に対する評価について

私たちは、国際日本学研究科の大学院生です。対面・オンラインでの英語スピーキング授業への学習者評価と、個人差要因（性格・学習スタイル等）と関係性について調査しております。アンケートの回答のご協力をお願いいたします。皆様から得た回答は研究の目的のみに使用することをお約束します。また、回答にあたって、セクション毎に指示を出します。指示に従って、自分に当てはまるものを選択してください。アンケートの回答は、成績とは一切関係ありません。リラックスしてお臨みください。

## (1) クラス・名前

---

## (2) 性別

男性    女性    回答しない

あなたのスピーキング能力について当てはまるものを1つ選んでください。

A → B → C の順でできることの幅が広がっています。

例) B を選んだ人は A と B ができる。C を選んだ人は ABC の全てができると解釈します。

## (3) 1人で発表する際に、

- A. 写真や絵、地図などの補助を利用しながら、一連の簡単な句や文を使って、身近なトピック（学校、地域など）について短い話をするができる。
- B. 使える語句や表現を繋いで、自分の経験や夢、希望を順序立て、話を広げながらある程度詳しく語るができる。
- C. 社会問題や時事問題に関して、具体的な例などを挙げながら、自分の視点を明確に話すことができ、話を続けることができる。

## (4) 2人以上のやりとりの際に、

- A. 補助となる絵やものを用いて、基本的な情報を伝え、また、簡単な意見交換をすることができる。
- B. 個人的に関心のある具体的なトピックについて、簡単な英語を多様に用いて、社会的な会話を続けることができる。
- C. 自分の学んだトピックや自分の興味や経験の範囲内のトピックなら、抽象的なトピックであっても議論できる。

## (5) スピーキングの授業様式に対する学習者の評価 (全19問)

以下の文に対して、「対面授業のとき」と「オンライン授業のとき」のそれぞれにおける自分の考えを、1 (全く当てはまらない) ～5 (よく当てはまる) より選んでください。

		全く当てはまらない					よく当てはまる
1	スピーキングの授業に集中して取り組みやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
2	スピーキングの授業中に意見や発言をしやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
3	スピーキングの授業に積極的に参加しやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
4*	スピーキングの授業中にミスをしたら、恥ずかしい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
5*	スピーキングの授業で英語を話している時、他者の目が気になる。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
6	スピーキングの授業に対するモチベーションがある。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
7	スピーキングの授業では、先生の声が明瞭で指示が伝わりやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
8	スピーキングの授業の流れがつかみやすい(次になにをすればよいかわかる)。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
9	スピーキングの授業では、他の学生と共に学習しているという一体感がある。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
10*	スピーキングの授業で英語を話している時に他の学生から孤立しているように感じる。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
11*	スピーキングの授業で話す時、緊張する。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
12	スピーキングの授業中、先生からの質問に答えやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
13	スピーキングの授業中、先生に質問しやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
14	スピーキングの授業中、ペアワークをしやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
15	スピーキングの授業中、3人以上でグループワークをしやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
16	スピーキングの授業中、大勢の前で1人で話す時、緊張せずに英語を話すことできる。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
17	スピーキングの授業中、発表スライドを用いずに1人で発言・発表しやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
18	スピーキングの授業中、発表スライドを用いて1人で発表しやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5
19	スピーキングの授業中、事前準備せずに1人で発言・発表しやすい。	対面授業	1	2	3	4	5
		オンライン授業	1	2	3	4	5

## (6) 学習スタイルについて (全 16 問)

以下の質問について当てはまるものを「全て」選んでください。回答は2つ以上も可能です。

1	説明書を見ながら家具を組み立てている途中でわからなくなったら？
	説明書を再度読み、理解しようとする。
	誰かに聞く、電話をして問い合わせるなど、人の助けを求める。
	部品をあれこれ触りながら、とにかく自己流でやってみる。
2	「一度しか会ったことがない人のことを思い出してください」と言われ、まず思い出すことは？
	その人の顔を思い出す。
	その人が言った言葉や話を思い出す。
	その人の行動、話し方を思い出す。
3	英単語を覚えるとき、自分が覚えやすいと思う方法は？
	何度も英単語のつづりを見て、書く。
	何度も英単語を聞き、声に出す。
	実際に会話や文章で書く。
4	メモができない状態で電話番号を覚えるときは？
	受話器の数字ボタンの位置をイメージしながら覚える。
	番号を何度も声に出し、音で覚える。
	指などを使って動作と関連づけて覚える。
5	道に迷って、誰かに教えてもらおう場合は？
	地図で書いてもらう。
	言葉で説明してもらう。
	目印となるものを教えてもらう。
6	日本語でも英語でも…
	日記をつけたり、ブログを書いたりすることが好きだ。
	音楽を聞いたり、講演を聞きに行ったりするのが好きだ。
	人と会話をすることが好きだ。
7	自分にとって一番ストレスの少ない自己紹介の方法は？
	自己紹介文を書く。
	自己紹介のスピーチをする。
	自分の特技などを披露する。
8	スーパーのお買い得商品を知りたいときは？
	チラシを見る。
	店員に聞く。
	店内を歩きながら確認する。
9	ニュースを知りたいときは？
	新聞を読む。
	音声ニュースを聞く。
	詳しい人に聞く、教えてもらう。
10	歌の内容を理解したいときは？
	歌詞を何度も読む。
	その歌を何度も聞く。
	その歌を何度も歌う。
11	英語で話すときは？
	まず、文法や単語をしっかりと下調べしてから話す。
	文法はとりあえず置いておき、とにかく話し始める。

12	英語の問題を解いていて、わからない単語があれば？	
		すぐ調べる。
		読み進めればわかるだろうと思って、調べない。
13	英語で発言する際は？	
		自分の英語が間違っているかもしれないので、何も話さずに終わることが多い。
		自分の英語が間違いかはあまり気にせず、話してみる。
14	英文法を勉強するときは？	
		文法のルールを自分なりにまとめないと、頭に入らないことが多い。
		文法のルールをじっくり覚えるより、問題文を解きながら学ぶことが多い。
15	勉強するときは？	
		1人で勉強する方がはかどる。
		グループで勉強する方がはかどる。
16	単語を覚えるときは？	
		1人で黙々と覚える方が、良い結果が出る。
		勉強仲間と問題の出し合いなどをしながら覚える方が、良い結果が出る。

## (7) 英語を話すことについて (全10問)

以下の質問について、1 (全く当てはまらない) ~ 7 (非常によく当てはまる) から選んでください。

		全く当てはまらない						非常によく当てはまる
1	私は、英語を話す力がないと思います。	1	2	3	4	5	6	7
2	私は、人と英語で話すとき緊張します。	1	2	3	4	5	6	7
3	私が話す英語は、相手に意味が伝わらないと思います。	1	2	3	4	5	6	7
4	私は、人と英語で話すことは恥ずかしいです。	1	2	3	4	5	6	7
5*	私は、英語で話してコミュニケーションできる力がある方だと思います。	1	2	3	4	5	6	7
6	私は、人と英語で話すとき、どきどきします。	1	2	3	4	5	6	7
7	私は、今の英語力では英語で話すことはできないと思います。	1	2	3	4	5	6	7
8*	私は、人と英語で話す時、リラックスしています。	1	2	3	4	5	6	7
9	私は、英語で話して、自分が考えていることを相手に伝えることができないと思います。	1	2	3	4	5	6	7
10	私は、人と英語で話す時、心配な気持ちになります。	1	2	3	4	5	6	7

## (8) 自分の性格について (全 29 問)

自分の性格について、当てはまるものを次の 1 (全く当てはまらない) ~ 7 (非常によく当てはまる) から選んでください。

		全く当てはまらない						非常によく当てはまる
1*	私は無口だ。	1	2	3	4	5	6	7
2	私は社交的だ。	1	2	3	4	5	6	7
3	私は話をするのが好きだ。	1	2	3	4	5	6	7
4	私は外交的だ。	1	2	3	4	5	6	7
5	私は陽気だ。	1	2	3	4	5	6	7
6*	私はいい加減だ。	1	2	3	4	5	6	7
7*	私はルーズだ。	1	2	3	4	5	6	7
8*	私は成り行き任せだ。	1	2	3	4	5	6	7
9*	私は怠け者だ。	1	2	3	4	5	6	7
10	私は計画性がある。	1	2	3	4	5	6	7
11*	私は軽率だ。	1	2	3	4	5	6	7
12	私は几帳面だ。	1	2	3	4	5	6	7
13	私は不安になりやすい。	1	2	3	4	5	6	7
14	私は心配性だ。	1	2	3	4	5	6	7
15	私は弱気である。	1	2	3	4	5	6	7
16	私は緊張しやすい。	1	2	3	4	5	6	7
17	私はよく憂鬱な気持ちになる。	1	2	3	4	5	6	7
18	私は多才だ。	1	2	3	4	5	6	7
19	私は物事を着々と進めるタイプだ。	1	2	3	4	5	6	7
20	私は独創的だ。	1	2	3	4	5	6	7
21	私は頭の回転が早い。	1	2	3	4	5	6	7
22	私は様々なことに興味がある。	1	2	3	4	5	6	7
23	私は好奇心が強い。	1	2	3	4	5	6	7
24*	私は短気だ。	1	2	3	4	5	6	7
25*	私は怒りっぽいと思う。	1	2	3	4	5	6	7
26	私は温和な性格だ。	1	2	3	4	5	6	7
27	私は寛大だ。	1	2	3	4	5	6	7
28*	私は自己中心的だ。	1	2	3	4	5	6	7
29	私は親切だ。	1	2	3	4	5	6	7

数字上の\*は逆転項目。



【研究論文】

# 日本語上級レベル留学生の論証型レポートに対する 自己評価・ピア評価の傾向

安高紀子  
ATAKA, Noriko

## 1. はじめに

教育評価において、従来の教師による評価に加え、学習者自身が評価の主体となる学習者主導型評価が取り入れられるようになり、学習者が自分自身に対して行う自己評価は、学習成果の進捗状況や達成度の評価を促す（パークレイ・メジャー 2020）とされている。また、形成的評価として、自己評価および学習者同士が評価を行うピア評価といった学習者主導型評価が行われることで、学習過程を評価することができる。

そこで、本研究では、日本語上級レベルの留学生を対象とするアカデミック・ライティングの授業において、学習者主導型評価である自己評価とピア評価をレポート作成のプロセスに取り入れた実践を行った。学習者が自己評価およびピア評価において、どのような評価を行っているのかを探り、その結果を基に、学習者自身の評価する力を育てるために、授業ではどのような取り組みを行っていく必要があるのか、授業の改善方法について検討する。

## 2. 研究動機

筆者は学部留学生1年生必修科目のアカデミック・ライティングの授業を担当している。本授業では、ライティングのスキル修得を目指し、課題レポートを作成するプロセスにおいて、ピア・ラーニングの実践方法のひとつであるピア・レスポンスを取り入れている。ピア・レスポンスとは、「作文の推敲のために学習者同士がお互いの書いたものを書き手と読み手の立場を交替しながら検討する活動」（池田 2007：71）である。授業において、レポートのテーマ決定、アウトラインの作成などのレポート作成の各ステップで、学習者は書き手（話し手）、あるいは読み手（聞き手）となり、質問をしたり、意見を述べるといったピア・レスポンスが行われている。

協働学習としてのピア・レスポンスの意義について、池田（2007）は認知的な観点から「批判的思考を活性化しながら進める作文学習」（p.73）としており、クラスメイトとの対話によって、新たなアイデアの獲得、思考の深まりといった効果が期待される。しかし、これまで筆者が担当していた授業では、学習者が書いたレポートに対して行ったピア・レスポンスでは、学習者はレポートの内容に関するコメントよりも、レポートの形式や表記のミス、文法的な誤りといった点を指摘する様子が見受けられた。つまり、レポートの言語的な側面に注目し、内容に関わる議論があ

まり行われておらず、内容の深まりといったピア・レスポンスの意義が実現されていないのではないかと思われた。また、本授業の目的はレポート作成スキルを身に付けることを目的としており、課題レポートで求められていることに対して、何ができており、何ができていないのかを客観的に判断する評価能力も必要だと考えられる。

### 3. 先行研究

日本語教育において、自己評価やピア評価といった学習者主導型評価を取り入れた実践が行われている。トムソン木下(2008)は、自己評価の有効な利用は利点が多いとしているが、学習者自身が自己評価能力を信頼していない場合や馴染みがない場合は成果が上がらないという点も指摘している。学習者の評価能力に関して、西岡他(2015)は、他者の具体的なテストや作品に対して評価する経験が目標である評価基準の理解を促し、学習を評価するメタ認知能力の育成につながると述べており、学習者同士が他者に対して評価し合うピア評価は評価能力を養う有効な手段であると言える。

学習者主導型評価を実践した研究として、ナズキアン(2010)、八若(2004)、市嶋(2014)がある。ナズキアン(2010)は、初級レベルの学習者のピア・ラーニングを取り入れたプロジェクトで、学習者同士が決めた評価基準に基づいて評価するピア・アセスメントを行っており、その有効性について、「学習者の内省、自己評価、そして内省に基づいた学習過程の修正を促すという点で教育的意味がある」(p.92)としている。八若(2004)は、活動中心の授業において実施した自己評価とピア評価について、学習者が書いた「学習記録」の記述を分析している。その結果について、学習者は自己の学習を振り返る活動を通じ、メタ認知知識を得ることができ、自律的学習が促進される効果があると示唆している。しかし、アンケートの結果からは、学習者は自己評価とピア評価の意義を認めているものの、自己評価に対しては客観性・妥当性に疑問を抱く者も少なくないとしている。また、市嶋(2014)は、学習者同士が決めたレポートの評価項目に基づいて、お互いのレポートを評価する相互自己評価活動を行っている。この活動に対する学習者の評価への認識を分析し、学習者は相互行為によって自己把握を促進させ、評価を意味づけていくというプロセスを明らかにしている。学習者の認識として、「学生は評価する力がない」、「教師が評価した方がいい」といったものも存在しており、相互自己評価の課題としている。

これらの研究は、学習者が自己評価・ピア評価を行うことで、評価をどのように捉えているのか、意識がどのように変容していったのかという学習者の情意的な側面に注目したものであり、実際に学習者が自己評価・ピア評価において、どのような評価を行っているのか、その評価自体については明らかにされていない。学習者主導型評価を実践し、有効な評価活動を行うためには、学習者がどのような評価を行っているのか、その傾向を明らかにする必要があるだろう。

### 4. 研究目的と課題

先行研究を踏まえ、本研究では、筆者が担当する上級日本語レベルの留学生を対象とするアカ

デミック・ライティングの授業において、期末課題として学習者が作成した論証型レポートに対して行った自己評価・ピア評価について、学習者がどのような評価を行っているのかを探ることを目的とする。本研究では、以下の2つを研究課題とする。

RQ1：学習者の自己評価とピア評価では、評価はどの程度一致しているのか。

RQ2：評価観点（説得力、読みやすさ、明晰さ、文・語句・表記、引用の適切性）によって、一致度に相違が見られるか。

## 5. 研究方法

### 5.1 対象の授業「上級日本語Ⅱ（文章表現）」と対象者

本研究の対象となる授業は、筆者が担当する学部留学生1年次必修科目「上級日本語Ⅱ（文章表現）」で、2020年度秋学期に実施された。新型コロナウイルス感染症の影響で、授業はZoomによるリアルタイムオンラインで行われた。対象者は「上級日本語Ⅱ（文章表現）」の履修者45名（1年生43名、2年生2名）で、日本語能力は日本語能力試験のN1レベル相当である。

本授業はアカデミック・ライティングのスキル修得を目的とし、期末課題である論証型レポート<sup>1</sup>（2000字～4000字）の作成を通して、論理的な文章作成の過程を経験するプロセス・ライティングの考えに基づいて設計されている。授業では、大島他（2014）を教科書として使用している。

表1に、本授業「上級日本語Ⅱ（文章表現）」全14回の内容を示す。本研究の分析対象とする自己評価とピア評価は、第13回、第14回の「レポートの点検」の授業で行われた。自己評価とピア評価の詳細については、次の5.2で説明する。

表1：「上級日本語Ⅱ（文章表現）」（2020秋学期）の授業内容

授業回	授業内容	授業回	授業内容
1	レポートとは何か	8	口頭発表をする（発表会1）
2	レポートのテーマを考える	9	口頭発表をする（発表会2）
3	文章の構想を練る	10	パラグラフを書く
4	レポートの目標を規定する	11	引用の仕方を知る
5	文章展開のパターンを知る	12	図表・数値データの使い方を知る
6	アウトラインを考える	13	レポートの点検（1回目）
7	発表の流れを考える	14	レポートの点検（2回目）

### 5.2 自己評価・ピア評価の手順

本研究の対象となる学習者による自己評価・ピア評価は、表1の通り、第13回と第14回の授業で行われた。自己評価・ピア評価の対象となる論証型レポートの執筆は、授業外の時間に、学習者が各自で取り組んだ。

1 大島他（2014）では、論証型レポートとは、ある問題に対し、複数の可能性の中から根拠（証拠・理由）と共に解決案（結論）を提示し、読み手を納得させるレポートとされている。

図1に、自己・ピア評価に関する活動の流れを示す。評価活動は、図1の通り、①～④の順に、次の手順で行われた。まず、①授業の前に、自分が執筆したレポートに対し、レポートの評価シートと照らし合わせて評価を行う（自己評価）。次に、授業では、②学習者同士でお互いのレポートを交換して読み合い、評価シートに基づき、相手のレポートに対する評価を行う（ピア評価）。③ピア評価の後、学習者同士で評価シートを交換し、何をどう評価したのかフィードバックを行う（評価フィードバック）。そして、レポートの内容に関して、修正が必要な箇所や改善案の提案を行う（相互コメント）。④授業の後、ピアからの評価フィードバック、コメントを基に、自分のレポートを見直し、修正する（自己修正）。以上の流れで実施した。授業の中で行われた②③の活動の所用時間はおよそ40分程度である。

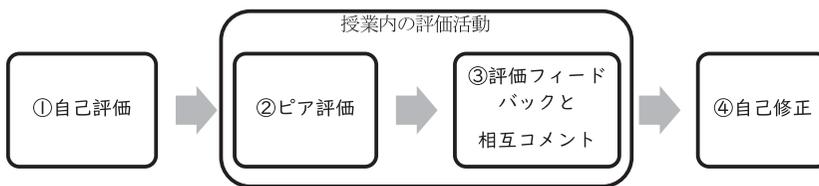


図1：自己評価・ピア評価に関する活動の流れ

### 5.3 レポートの評価シート

自己評価・ピア評価で用いられた評価シートは、大島他（2005）のレポートチェックシート（ワークシート p16）を基に作成した。大島他（2005）のレポートチェックシートはレポート提出前のチェックリストとして用いることを想定されたものだが、本研究では評価シートとして活用するために、各評価観点に学生が判定を記入する評価欄、また、各自が判定理由を記入するためのコメント欄を加え、チェックリスト形式の評価シートを作成した。レポートなどのパフォーマンス評価ではルーブリック<sup>2</sup>が用いられることが多い。しかし、本調査では、学習者自身が自己評価・ピア評価を基に、レポートの見直しを行い、修正するため、観点の尺度段階の特徴が記述されたルーブリックよりも、確認すべき点を一覧にしたチェック形式の方が利用しやすいと考え、チェックリスト形式の評価シートを用いることとした。原田他(2019)では、評価シートのルーブリックの記述欄を空欄にし、学習者の内省を自身の言葉で書ける形式にしており、これによって、規準に照らし合わせ、自身の内省を浮かび上がらせることが可能となると示唆している。そこで、本研究では、判定の記入だけでなく、コメント欄を設けることで、チェック項目を単なるチェックリストとして用いるのではなく、なぜ自分がそのように判定したのかを考えて言語化する必要を生じさせ、学生の内省を促すようにした。

2 スティーブンス & レビ（2014）によると、ルーブリックは課題、評価尺度、評価観点、評価基準の4要素から成るもので、「成功の度合いを示す数値的な尺度とそれぞれの尺度に見られる認識や行為の特徴を示した記述語からなる評価指標」（石井 2010）である。

表2に本研究の自己評価・ピア評価で使用した評価シートの記入例を示す。これは、実際に学習者が記入したものである。表2の通り、評価観点は【説得力】【読みやすさ】【明晰さ】【文・語句・表記】【引用の適切性】の5つから成る。5つの観点には、各観点の要素となる内容がチェック項目として3～5項目示されている。例えば、【説得力】の観点には「目標規定文と結論が対応しているか」「主張の展開は適切で説得力があるか」「信頼できる情報やデータを集めているか」など5つのチェック項目がある。評価欄は、チェック項目を参照しながら、該当する観点の達成度をABCDの4段階評価とし、達成度が80%以上はA、70%以上はB、60%以上はC、60%未満はDとした。

表2：レポートの評価シートの記入例

評価観点	チェック項目	評価	修正したほうがよい点、コメントなど
説得力	<ul style="list-style-type: none"> <li>・目標規定文と結論が対応しているか</li> <li>・主張の展開は適切で説得力があるか</li> <li>・反論や反駁があるか</li> <li>・信頼できる情報やデータを集めているか</li> <li>・必要な概念・語句の定義があるか</li> </ul>	B	信頼できる情報やデータを使い、全般的に説得力のある文にできている。反論・反駁が提示されている
読みやすさ	<ul style="list-style-type: none"> <li>・タイトルや見出しが適切であるか</li> <li>・1パラグラフ1話題になっているか</li> <li>・パラグラフの冒頭に中心文があるか</li> <li>・序論に動機、目標、全体の構成があるか</li> </ul>	A	全体的で、内容が明確で読みやすい パラグラフの冒頭が中心文になっておりパラグラフを把握しやすい
明晰さ	<ul style="list-style-type: none"> <li>・事実と意見がはっきり区別できるか</li> <li>・主張には妥当な理由と証拠があるか</li> <li>・主張と理由の関連性を論証できているか</li> </ul>	A	事実と意見がはっきり区別できていて読みやすい
文・語句・表記	<ul style="list-style-type: none"> <li>・文体は「だ・である体」で統一しているか</li> <li>・レポートに適切な書きことば表現を用いているか</li> <li>・誤字、脱字、変換ミスがないか</li> <li>・読点が適切に打たれているか</li> </ul>	B	「一つの要因なれると思う。」 「おわり(に)」などのミスあり
引用の適切性	<ul style="list-style-type: none"> <li>・どこからどこまでが引用か明瞭であるか</li> <li>・引用箇所の書式はルール通りであるか</li> <li>・図表やデータに出典が明記されているか</li> <li>・文献リストは整っているか</li> </ul>	A	引用の区間が明確に区切られている 引用のルールを守っている

## 5.4 データと分析方法

分析に用いるデータは、本授業で学習者が記入したレポートの自己評価・ピア評価の評価シートである。第13回、第14回の授業後、学習者から提出されたレポートの評価シートのうち、自己評価とピア評価の双方の評価シートが揃っているものを分析対象とする。1回目は37名、2回目は38名のデータを分析に用いる。データの使用にあたって、調査対象者には研究目的のために使用することを説明し、同意を得た。

データの分析方法は、自己評価とピア評価の2つの評価の判定が一致しているかどうか一致度を指標とし、重み付きカッパ係数により算出する。

## 6. 結果と考察

### 6.1 自己評価・ピア評価の一致度

学習者が行った自己評価とピア評価がどの程度一致しているかについて、重み付きカッパ係数

により一致度を算出した。カッパ係数は1.0を完全な一致とする<sup>3</sup>。

まず、自己評価とピア評価の1回目(n=37)の結果について、表3の自己評価・ピア評価(1回目)の観点別一致度に示す。5つある観点別に一致度を算出した。表3の通り、最も一致度が高かったのは【説得力】で、 $\kappa = .438$ であった。その他の4つの観点の一致度は、【引用の適切性】 $\kappa = .305$ 、【明晰さ】 $\kappa = .265$ 、【読みやすさ】 $\kappa = .262$ 、【文・語句・表記】 $\kappa = .220$ の順であった。【説得力】は中等度の一致であったが、他の4つの観点については、まずまずの一致であった。

表3：自己評価・ピア評価(1回目)の観点別一致度

	説得力	読みやすさ	明晰さ	文・語句・表記	引用の適切性
一致度 $\kappa$	.438	.262	.265	.220	.305

次に、自己評価とピア評価の2回目(n=38)の結果について、表4の自己評価・ピア評価(2回目)の観点別一致度に示す。表4の通り、一致度の高いものから順に、【説得力】 $\kappa = .475$ 、【引用の適切性】 $\kappa = .475$ 、【文・語句・表記】 $\kappa = .392$ 、【読みやすさ】 $\kappa = .389$ 、【明晰さ】 $\kappa = .183$ であった。【説得力】と【引用の適切性】の2つの観点は、中等度の一致であった。【文・語句・表記】と【読みやすさ】の2つは、まずまずの一致であった。そして、【明晰さ】は、わずかに一致で、最も一致度が低かった。

表4：自己評価・ピア評価(2回目)の観点別一致度

	説得力	読みやすさ	明晰さ	文・語句・表記	引用の適切性
一致度 $\kappa$	.475	.389	.183	.392	.475

## 6.2 評価観点別の一致度

6.1の通り、自己評価とピア評価の一致度について、評価観点によって一致度には相違があるという結果が得られた。自己評価・ピア評価のそれぞれ1回目と2回目の結果から、【説得力】と【明晰さ】の観点には、他とは異なる傾向が見られた。【説得力】の観点は、1回目も2回目も5つの観点の中で一致度が最も高く、いずれにおいても【説得力】は中程度の一致であった。そして、【明晰さ】の観点については、1回目の結果ではまずまずの一致、そして、2回目の結果ではわずかな一致であり、一致度は1回目より2回目で低い値であった。これは【明晰さ】のみに見られる傾向であった。したがって、他の観点と傾向が異なる【説得力】と【明晰さ】について、その要因を考察する。

3 一致度の解釈の基準としては、カッパ係数が0.0～0.2をわずかに一致(slight agreement)、0.21～0.40をまずまずの一致(fair agreement)、0.41～0.60を中等度の一致(moderate)、0.61～0.80をかなりの一致(substantial)、0.81～1.00をほぼ完全な一致(almost perfect)とされている(Landis & Koch, 1977)。

### 6.2.1 【説得力】の観点の一致度

【説得力】の観点の一致度は、1回目は $\kappa = .438$ 、2回目は $\kappa = .475$ で、他の観点に比べ、1回目、2回目ともに一致度が最も高かった。そして、1回目も2回目も中程度の一致度という結果であり、他の観点に比べ、自己評価とピア評価の評定が一致している度合いが高い。

そこで、この傾向の要因を探るため、【説得力】の評価において、学習者がどのような点に着目して判定を行っていたのか、学習者が記入した評価シートのコメント欄の記述内容を確認した。学習者が記入した内容と【説得力】の5つのチェック項目とを照らし合わせ、どのチェック項目に該当するのか分類した。

図2に【説得力】の項目別の評価コメント数と割合を示す。分類項目は全部で7つあり、レポートの評価シートの【説得力】の観点にあった5つのチェック項目に加え、5つのいずれの項目にも該当しないもののうち、【説得力】の観点に関するものは「6. その他」とし、【説得力】の観点に関係のないものは「0. 観点外」とした。学習者のコメントを分類した結果、図2に示す通り、「3. 反論・反駁があるか」、「4. 信頼できる情報やデータを集めているか」という2つに関するものが他の項目よりも多いことがわかる。

学習者が実際に書いたコメントの例を示す。項目「3. 反論・反駁があるか」に関しては、「反論や反駁はなかった」「自分の意見に対する反対意見はあるが、反駁が書いていない」「反論に対する反駁がないと思いますので、入れたらもっとよいと思います」といったコメントが見られた。反論を予想し、それに自分がどう対応するかという反駁を書くことは論証型のレポートでは重要であると学習者は意識していると思われる。

また、項目「4. 信頼できる情報やデータを集めているか」に関しては、「信頼できる情報やデータが足りない」「信頼できるデータを集めている」「データや情報がちょっと少ないので、もっと説得力があるデータを入れた方がいい」といったデータの有無について言及するコメントが多かった。また、「データは新聞からの情報をうまく使えていて良かったと思います」という信頼性に関するコメントも見られた。

【説得力】の観点に記されたコメントの内、「3. 反論・反駁があるか」「4. 信頼できる情報やデータを集めているか」に関するものが多く、この2つの項目を合計すると、1回目の評価では、自己評価49%、ピア評価54%、2回目の評価では、自己評価53%、ピア評価54%となり、これら2つの項目に関わるコメントだけで全体の約半数を占めている。【説得力】の項目の中で、これらのチェック項目は、レポートの文章の中で、反論や反駁があるか、データがあるかといった存在の有無を問うものであり、有無の判断は比較的容易だと思われる。これに対して、その他のチェック項目「1. 目標規定文と結論が対応しているか」は、レポートの目的が記された目標規定文の内容と最終的な結論が合っているのかという一貫性に対する判断が求められる。また、同じく「2. 主張の展開は適切で説得力があるか」については、主張に対し、理由や根拠が示されているかを確認し、かつ、それらは適切であるかどうか、説得力があると言えるかといった判断

が要求され、どの程度できているかを判断することは難しい。このように、【説得力】のチェック項目の間で、評価を行う難易度に差があると思われる。評価が難しいと考えられるチェック項目「1. 目標規定文と結論が対応しているか」と「2. 主張の展開は適切で説得力があるか」に関するコメント数は、項目「3. 反論・反駁があるか」や「4. 信頼できる情報やデータを集めているか」に比べて少ない。このことから、学習者は比較的判断が容易なチェック項目に注目し、評価をしていたのではないかと推察される。チェック項目は評価観点を構成する要素であり、評価は各チェック項目を総合的に判断して行うべきである。しかし、学習者が判定の理由として言及していたのは、チェック項目3と4に偏っており、限られた項目に基づいて評価を行っていたことが、一致度に影響している可能性が考えられる。

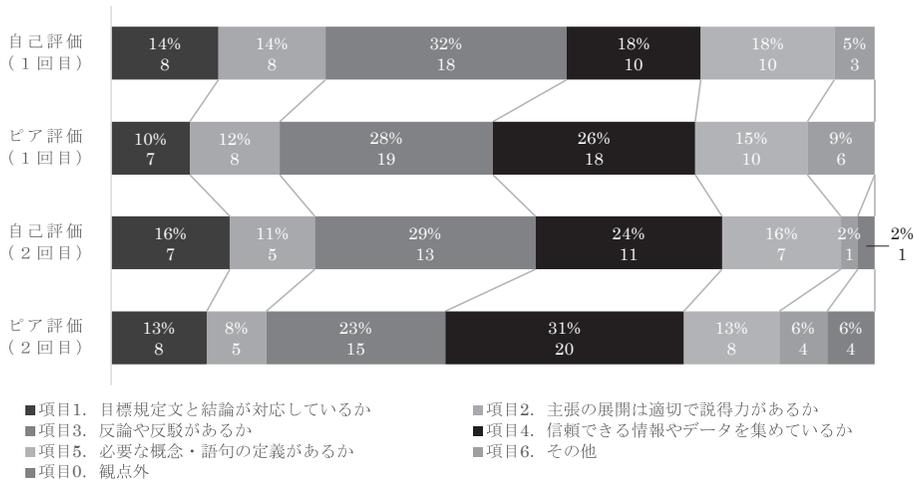


図2：【説得力】のチェック項目別の評価コメント数

### 6.2.2 【明晰さ】の観点の一致度

自己評価・ピア評価の1回目と2回目の一致度について、【明晰さ】の観点は、1回目はまずまずの一致、2回目はわずかな一致という結果で、5つの観点の中で【明晰さ】のみ、1回目より2回目で一致度は低い値であった。

一致度が低い要因として考えられることは、【明晰さ】という観点に対する学習者の理解度の問題である。明晰さは文章の論理性に関わる観点であるが、学習者にとっては明晰さとは何か、【明晰さ】という観点に対する理解が十分できていなかったのではないと思われる。それが推察される学生のコメントを示す。コメントは評価シートの【明晰さ】の観点に書かれたものである。例えば、「様々なサイトや記事を参考にした情報がある」というコメントが見られた。これは客観的な情報の有無について記されており、「信頼できる情報やデータを集めているか」というチェック項目に該当すると思われる。このチェック項目は、【説得力】の観点にあるチェック項目で、【明晰さ】に関わるものではない。学習者は、【説得力】の観点で、【明晰さ】を評価し

ていると言える。また、他の例では、「主張を支える根拠のデータの出典が示されている」というコメントが見られた。これはレポートの「根拠」について、学習者は「出典がある」という点を評価しており、その評価は【引用の適切さ】の観点に該当する。【明晰さ】の観点であれば、「根拠」について、主張に対する根拠としての妥当性を評価することになる。以上の例に見られるように、学習者は【明晰さ】が何を評価するものなのか、理解があいまいであったと思われる。このように、【明晰さ】の観点に対する理解が不十分のまま、評価が行われたと考えられ、それが一致度の低さに影響しているのではないかと思われる。

## 7 評価活動の実践における課題

本研究の結果から、今後、自己評価・ピア評価の評価活動を実践する上で、どのような改善を行っていく必要があるのか、その課題について述べる。

### (1) レポート評価シートの改良

レポート評価シートのコメント欄に学習者が記述した内容を確認したところ、学習者の評価観点に対する解釈のあいまいさ、理解の不十分さが見られ、評価に影響を及ぼしている可能性について述べた。このことから、評価シートに学習者が記述したコメントについて、さらなる分析を行うことによって、学習者の評価における問題点が浮かび上がってくるとと思われる。したがって、今後は評価シートのコメント欄の記述の分析を行い、明らかになった問題点を踏まえ、チェック項目の見直しを行う。例えば、【明晰さ】の観点については一致度が低かったことを受け、明晰さに含まれる要素は何かを再検討し、学習者が明晰さとは何かを理解できるような文言に修正する。学習者が評価シートを見て、何を評価すべきかが理解できるよう、評価シートを改良する必要があるだろう。

### (2) 自己評価とピア評価のすり合わせ

今回、実践した評価活動では、学習者同士でピア評価の結果についてフィードバックを行い、その後、各自が自己修正の際に、自己評価を振り返るという方法を取っていた。しかし、これでは自己評価とピア評価で生じたズレの要因を理解するには、不十分であったと思われる。岩田(2020)は、自己評価とピアからの評価を比較し、得点やその根拠やズレの理由を話し合う協議ワークを授業の評価活動の中に取り入れている。評価結果についてフィードバックするだけではなく、この協議ワークを取り入れたことで、自己評価力が統計的に有意に向上することを検証している。自己評価とピア評価で生じたズレに対して、なぜズレが生じたのかをピアと話し合うことが、自分自身の持つ評価基準を調整することにつながるものと思われる。したがって、今後は、この協議ワークを評価活動の一環として組み込み、自己評価力を伸ばせるようにしたい。

### (3) 評価トレーニングの実施

トムソン木下(2008)は、学習者主導型評価の提案として、学習者がピア評価の経験を積み、その経験を自己評価に移行することで、自己評価の精度を増すことが期待できるとしている。本実践を行った授業後のふりかえりにおいて、学習者のコメントには、「相手のレポートを読みながら評価したことから、自分のレポートをより客観的に見ることができた」、「自分が自分の評価をすることがとても難しいと感じました。それでピア評価をすることで、自分で知らなかった問題点を知ることができるということが大切だということを感じました」「私が相手のレポートを読みながら、引用の形式や見出しなど、私が間違えた部分について気づけることもできました」といったものが見られた。ピアへの評価を行うことによって、自分を客観的に捉えられるようになったと感じている様子が観察され、自己モニターを促進させる可能性がうかがえる。

したがって、本実践では自己評価・ピア評価は2回しか行うことができなかったが、繰り返し評価を行うことが有効であろう。評価能力を育成するためには、授業において、評価トレーニングを実施することで、良い例や悪い例など具体例を見ながら、何を評価すべきかといった評価観点の理解や到達度に関する判断が可能になる。また、サンプルを使用して実際に評価を行ってみることで、そして、学習者同士で評価結果について話し合うことで、学習者自身の中で評価基準や評価尺度の調整ができるだろう。

## 8. おわりに

本研究では、筆者が担当する日本語上級レベルの学部留学生のアカデミック・ライティングの授業において実施した自己評価とピア評価について、学習者がどのような評価を行っているのかを探るため、学習者の自己評価とピア評価の一致度を指標として分析を行った。研究課題とした以下の2つについて、次のような結果が得られた。

RQ1: 学習者の自己評価とピア評価の一致度については、自己評価・ピア評価の1回目も2回目も全体的にそれほど高くはなかった。

RQ2: 評価観点による一致度の相違については、観点によって差が見られた。【説得力】は他の観点に比べ、一致度が最も高かったが、学習者が評価の際に着目したチェック項目に偏りがあり、それが判定に影響していることがうかがえた。一方、一致度が最も低かった【明晰さ】については、明晰さという観点に対する学習者の解釈があいまいであったことが要因だと思われる。

今後の課題として、レポートの評価シートに記述されたコメントを分析し、分析結果を基に、評価シートの見直し、および改良を行う。そして、授業において、サンプルを用いた評価トレーニングの実施、評価活動に評価のすり合わせを行うステップを加えるなどの改善を行い、学習者自身が評価の主体となり、自らの学びの成果を評価できる力を育むことができる評価活動の実現を目指したい。

## 謝辞

本研究において、調査に協力して下さった本学部の2020年度「上級日本語Ⅱ（文章表現）」の履修者である留学生のみなさんに深く感謝を申し上げます。

## 参考文献

- 池田玲子（2007）「ピア・レスポンス」,池田玲子・館岡洋子『ピア・ラーニング入門 創造的な学びのデザインのために』ひつじ書房
- 市嶋典子（2014）『日本語教育における評価と「実践研究」対話的アセスメント：価値の衝突と共有のプロセス』ココ出版
- 岩田貴帆（2020）「協議ワークを取り入れたピアレビューによる学生の自己評価力向上の効果検証」『大学教育学会誌』,42(1),pp.115-124
- エリザベス・F・バークレイ、クレア・ハウエル・メジャー（2020）吉田壘（監訳）『学習評価ハンドブック アクティブラーニングを促す50の技法』東京大学出版
- 大島弥生・池田玲子・大場理恵子・加納なおみ・高橋淑郎・岩田夏穂（2005）『ピアで学ぶ大学生の日本語表現 第2版 プロセス重視のレポート作成』ひつじ書院
- ダネル・スティーブンス、アントニア・レビ（井上敏憲・俣野秀典 訳）（2014）『大学教員のためのルーブリック評価入門』玉川大学出版会
- トムソン木下千尋（2008）「海外の日本語教育の現場における評価：自己評価の活用と学習者主導型評価の提案」『日本語教育』,136号, pp.27-37
- ナズキアン富美子（2010）「ピアラーニングとアセスメント プロログを用いたピアラーニング」佐藤慎司・熊谷由理（編）『アセスメントと日本語教育 新しい評価の理論と実践』くろしお出版
- 西岡加名恵・石井英真・田中耕治（編）（2015）『新しい教育評価入門 人を育てる評価のために』有斐閣コンパクト
- 八若壽美子（2004）「活動中心の授業における評価：自己評価・ピア評価の有効性」『茨城大学留学生センター紀要』, (2), pp.13-24
- 原田三千代・浅津嘉之・田中信之・中尾桂子・福岡寿美子（2019）「アカデミック・ライティングにおける対話的評価活動の可能性」『小出記念日本語教育研究会論文集』,27, pp.37-52
- Landis, J. R., & Koch, G. G. (1977). *The measurement of observer agreement for categorical data*. *Biometrics*, 33, pp.159-174.



[Articles]

## Cognition or Context: How Do Teachers Approach Japanese University Speaking Classes?

GARSIDE, Paul

### Abstract

This interview-based study aims to provide some insight into how speaking is taught in Japanese universities, which is an under-researched and perhaps even mysterious field given the autonomy that teachers are often afforded in such classes. Two native speaker and two non-native speaker teachers of English (from the same private university in Tokyo) were interviewed for the study. They were asked what approaches they have to teaching speaking and what factors have influenced these approaches. Although differences emerged regarding attitudes to pedagogical theory, feedback, and tolerance of L1 use in the classroom, all four participants appeared committed to teaching speaking in a student-centered way, with fluency prioritized over formal accuracy. Furthermore, a combination of their own previous educational and teaching experiences, as well as contextual constraints and institutional requirements (in addition to internal factors such as age and personality) have been the biggest influences on how they approach teaching speaking classes at the university level.

**Keywords:** speaking, teacher cognition, teaching context

### Introduction

Since the mid-1990s, a good deal of research has been done into teacher cognition, which has been defined as “what language teachers think, know, believe, and do” (Borg, 2003). It is widely acknowledged that this cognition is shaped by some combination of internal (cognitive) and external (environmental) factors. Further, it influences teachers’ techniques and classroom behaviors, which has obvious repercussions for students and their learning outcomes. Teacher-based research is, therefore, an important area of investigation and forms an integral part of the attempt to understand how languages are learned in the classroom.

I have spent the vast majority of my teaching career working in Japanese universities and have developed a curiosity about the inner workings of this learning environment. Ironically, given its inherently social nature, teaching can feel like an individual enterprise, especially when teachers are given as much autonomy as is often the case in the university system. This interview-based study aims to shed some light on this teaching context and, especially, how it relates to the development of speaking skills.

## Literature Review

### The Japanese Context

Since its institutional beginnings in the nineteenth century, formal English teaching in Japan has been dominated by *yakudoku*, or the grammar-translation method. However, since the 1980s The Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) has promoted more communicative methodologies, albeit with questionable levels of success. Many high school teachers continue to teach grammar according to the way they themselves learned, reinforced by a system that has promoted grammar-based instruction for decades (Nishimuro & Borg, 2013). Even high school teachers with positive views about communicative language teaching (CLT) find themselves compromised by the reality of preparing students for university entrance exams (Nishino, 2012). So pervasive is the influence of these exams that some students are even reluctant to develop their speaking skills, perceiving them as unrelated to their chances of success (Sakui, 2004). Things could be about to change in this regard, however, as the government plans to incorporate a speaking element into university entrance exams.

Nevertheless, at present most freshman university students are likely to have a passive knowledge of at least basic vocabulary and grammar, and often far more extensive than that. Therefore, regarding university speaking classes, the main issue for many practitioners has been how to utilize the English that students already know. In other words, the goal is for students to become active users, rather than receptive learners, of language (Wiltshier & Helgesen, 2019).

With no further high stakes testing involved, universities are free to devise their own English curricula to achieve this goal. This system leads to more flexibility in the way English is taught, which translates to greater autonomy for teachers. Given that the majority of students are likely to be unfamiliar with predominantly English-speaking environments, a particularly salient issue is that of attitudes to L1 use in the classroom. A variety of approaches to this issue have been reported in the Japanese university context. For example, Ford (2009) found that native-speaker teachers varied in their tolerance of student L1 use, based on practical considerations and individual beliefs rather than institutional demands. Even if an institution

does have an English-only rule, some teachers interpret it more strictly than others depending on their own beliefs and experiences (Saito, 2014).

### **Influences on Teacher Cognition**

A variety of factors influence teacher cognition and the way it manifests itself in classroom practice. Based on recurrent themes in the literature, Borg (2003) identified four major influences: previous language-learning experience; teacher training and education; teaching practice and experience; and contextual factors. Let us examine each factor in turn.

Prior language-learning experience can have either a positive or a negative influence on teacher cognition. For example, Numrich (1996) found that over a quarter of novice teachers wanted to include a cultural component into their lessons because they had enjoyed learning about L2 culture in their own lessons. Several teachers in the same study reported a desire to avoid error correction due to their own negative experiences as learners. Moodie and Feryok (2015), in their study of Korean primary school teachers, suggested that an early commitment to language learning contributes to an increased commitment to language teaching, both in terms of professional development and English proficiency. The authors drew a parallel between the persistence needed to learn a language with the persistence needed to continually improve one's own teaching.

Although formative learning experiences can prove stubbornly resistant to change, teacher education and training can modify the influence of these experiences. For instance, Peacock (2001), in a longitudinal study, found that third-year trainee teachers were far less likely than first-year trainees to view language learning as primarily a case of learning grammar rules. They were also less likely to relate successful L2 learning to general intelligence. Furthermore, even if training and education does not lead to substantial changes in outlook it can help more experienced teachers consolidate and articulate their beliefs (Borg, 2011).

Experience gained through teaching is also likely to influence the attitudes and beliefs of practitioners. Moreover, beliefs formed through such experience are more likely to be reflected in practice, as they are more closely related to the reality of the classroom (Basturkmen, 2012). In other words, a reflexive relationship exists between practice and principles, and indeed there is evidence that teaching behaviors can change over time. For example, Richards (1998) found that experienced teachers improvised more than inexperienced teachers, suggesting that flexibility and the ability to deviate from lesson plans develops gradually.

It should be emphasized, however, that stated principles and observed practice do not always correspond. In fact, Basturkmen (2012), in a review of research into this issue, claimed that there is no more than a limited connection between the two. A stronger relationship has

been found between principles and practice among more experienced teachers, however, as teaching beliefs become embedded over time and are therefore applied more consistently (Breen et al., 2001).

One reason that changes in teacher cognition do not always result in behavioral change is that contextual factors can place limitations on practice. The demands of parents and administrators, as well as the availability of resources, can all hinder the ability of language teachers to utilize techniques and activities that reflect their beliefs. For example, Kurihara and Samimy (2007) reported that Japanese teachers who had been trained in the U.S. felt constrained by everything from exam requirements and class sizes to students' expectations and opposition from colleagues, when trying to implement communicative methods at home. The need for more class hours and smaller class sizes is a common complaint among teachers who feel unable to implement CLT within the Japanese school system (Nishino, 2008), although it has been claimed that some teachers who blame external constraints are simply afraid of abandoning their traditional teaching style (Kurihara & Samimy, 2007).

Although the greater autonomy afforded in the university context means that teachers often face fewer obstacles, unfamiliar methodologies might be resisted by some students. For example, creative activities requiring active participation might not be valued, given how much they differ from the all-too-serious business of passing entrance exams (Fujimoto, 2019). Moreover, communicating directly in an L2 is known to be a face-threatening activity, which is likely to cause anxiety among students unfamiliar with English-speaking environments. Patience and persistence are therefore required when attempting to implement pedagogical principles, even in the absence of administrative constraints.

Another explanation for the apparent mismatch between stated beliefs and observed practice is the tension between core and peripheral beliefs; the former being applied consistently across situations while the latter are applied more flexibly (Breen et al., 2001). When the two conflict, therefore, core beliefs tend to win out as they have become deeply ingrained through experience, while peripheral beliefs, even if theoretically-based, are not usually held with the same level of conviction (Phipps & Borg, 2009). For instance, teachers who have been trained to use communicative methods might revert to a more teacher-centered approach if, fundamentally, they believe that teachers should dominate classroom interaction.

### **Purposes and Research Questions**

Overall, the literature supports Borg's (2015) view that language teaching is a set of "dynamic interactions among cognition, context, and experience" (p. 275). Narrowing the focus to Japan, English teaching within the Japanese university system has been under-researched, apart from

the issue of L1 use in the classroom mentioned earlier. Given that it represents the apex of the formal education system it surely deserves greater attention.

The majority of teacher cognition studies to date have focused on structural aspects such as the teaching of grammar and literacy, but I would like to investigate how it relates to the way speaking is taught in general. Traditionally, a lot of flexibility has been afforded in university communication classes, so there is a good deal of opportunity for teachers to conduct them according to their own pedagogical conceptions.

It would be of additional benefit to observe how teacher cognition is manifested in classroom practice, and ultimately how it relates to learning outcomes (as called for by Borg, 2003), although that is beyond the scope of the current study. Data collection has, therefore, been limited to teacher interviews.

The research questions are:

1. What approaches do English teachers have to teaching speaking in a Japanese university?
2. What factors influence these approaches?

## **Methodology**

This study belongs within the constructivist paradigm, which prioritizes the perspective of the individual (Hatch, 2002). Although university teachers of English share similarities in terms of their responsibilities, they have different language learning, teacher training, and practical teaching experiences. When these diverse experiences are combined with factors such as personality and cultural background, pedagogical priorities and practices will inevitably vary. By conducting one-to-one interviews, I have attempted to relate each participant's background to their beliefs about teaching and their conceptualization of the teacher's role in a speaking class.

## **Context and Participants**

The number of participants was limited to four, all from the same private university in Tokyo. All participants had between six and twelve years' experience teaching in Japanese universities, having previously taught in other contexts. My aim when recruiting participants was to represent as many different elements of the English teaching population as possible within this institution. This form of sampling is known as maximum variation sampling, in which participants are identified according to specific attributes designated by the researcher (Roulston, 2010).

Although all participants were within the same age range, two were male (one American, one British), and two were female (one Japanese, one Chinese). As well as the two native speakers of English being from different countries, I wanted the non-native speakers to represent both

Japan and a different country. My intention was to investigate what commonalities, as well as differences, exist in their approaches to teaching. I then tried to analyze these features through the lens of their cultural background and previous learning and teaching experiences.

### **Data Collection**

One formal interview was scheduled with each participant, limited to between 30 and 40 minutes each. There were no linguistic difficulties as both of the non-native English teachers spoke English fluently. The interviews were semi-structured, therefore I had a number of open-ended questions prepared and the participants were encouraged to give full and expansive answers. These answers were then followed up with probes for further detail and explanation when necessary (Roulston, 2010). Although I tried to relate each interview question to the research questions, I also tried to avoid framing them in technical language so that they were clear and easy to answer. The interview questions are presented below, along with the relevant research question.

#### ***Background Questions***

- How long have you been teaching English? How long in Japanese universities?
- How and why did you train as an English teacher?
- What is your experience as a language learner?

#### ***Research Question One (What approaches do English teachers have to teaching speaking in a Japanese university?)***

1. Describe what you do in your speaking classes.
2. What do you think makes a good speaking class?
3. Which is more important in a speaking class, accuracy or fluency?
4. What is the role of the teacher and students in a speaking class?
5. How would you describe your relationship with your students in these classes?
6. How would you describe your teacher talk in these classes?
7. What is your view regarding L1 use in these classes?

#### ***Research Question Two (What factors influence these approaches?)***

8. How, if at all, have your experiences of language learning influenced the way you teach speaking classes?
9. How, if at all, have your experiences of teacher training and education influenced the way you teach speaking classes?
10. How, if at all, have you changed your approach to teaching speaking classes since you began teaching?
11. What do you feel are the pros and cons of teaching speaking within a unified curriculum?

### **Transcription and Analysis**

As the interviews were relatively short, I transcribed them in full. I did not adopt conversation analysis levels of detail as I was not concerned with the co-construction of talk during the interviews. However, I did include paralinguistic features such as laughter to portray the mood of the participants when I felt it was significant. This kind of non-verbal information brings an interview to life on the page and allows the reader to understand how the participants felt, which is an important aspect of the analysis.

After transcription, I conducted a thematic analysis of the data, which entailed sorting the data into categories based on topics raised during the interviews. This process was conducted inductively (i.e., without reference to a priori predictions and hypotheses) to allow the flexibility to consider topics initiated by the participants as well as the researcher (Roulston, 2010). Nevertheless, the interview questions were designed in such a way as to elicit comments relevant to the research questions and interviewee comments were coded accordingly. Topics therefore included practical and theoretical aspects of the participants' teaching approaches, as well as the factors that have influenced them, such as their own language learning and teaching experience. Codes or labels were not prepared in advance, in order to allow themes to emerge from the data.

### **Positionality**

Ethical considerations were minimized in this study as there was no status difference between myself and the participants. At the time, we were all teachers in the same department of the same university and roughly within the same age bracket. All four participants were known to me in a professional capacity, which made gaining access straightforward. Pseudonyms have been used to ensure confidentiality.

Nevertheless, there are potential drawbacks to this kind of research context. First of all, there is a danger that familiarity with the participants could make it difficult to maintain objectivity, although I had never seen any of them teach and had not discussed their approach to teaching communication classes prior to the study. I was also conscious of the need to be well-prepared and to maintain focus during the interviews, just as would be the case for participants with whom I had no prior contact or relationship.

A further issue is how the participants oriented themselves to me as a researcher. For example, if they had sensed that I favor a particular approach to teaching speaking this could have influenced how they responded to my questions. In addition, although there was no status differential, as professional colleagues they might still have been inclined to overemphasize the pedagogical or theoretical framework behind their approach.

## Results and Findings

### Approaches to Teaching Speaking

Several themes emerged from the interview data regarding overall approaches to teaching speaking. One thing that came across strongly was that all four participants favor a student-centered classroom, in which the teacher keeps a low profile during communication activities. According to Xia,

I think if it's a good speaking class, usually I feel lonely (laughs). I feel lonely means all my students are talking, so they are really enjoying their conversation and they ignore me. I mean I'm just standing there and watching them speaking and, because I'm the teacher, I cannot join them because my job is just watching and organize and give some advice if necessary.

It is significant that Xia even feels she “cannot join” her students, as if she is compelled to take a back seat. Similarly, Rika stated that “I consider myself as a facilitator, because I don't want to be like ‘a teacher’ because students are the main part. So I just help them when they need help”. The fact that she doesn't want to be like ‘a teacher’ suggests that she is rejecting the traditional role of the educator that she was familiar with in her own schooldays. Tom, referring to the teacher's role as “the guide on the side”, also stated that “you definitely don't want the teacher to be the central focus of a communication or speaking class”. Richard was less specific about this issue, but mentioned that he allows students to choose some of the topics to be discussed.

Another area of agreement was that fluency should be prioritized over accuracy in speaking classes. According to Richard, “In those kind of discussion classes... the communication is much more important, I think, than accuracy.” Tom added, “ideally you have them focused on meaningful interaction and to minimize their concerns about formal accuracy.” Rika and Xia agreed, with the latter relating a speaking class to the entertainment industry: “When people are talking it's like kind of showbusiness, and if the student can speak really fluently, it shows their confidence.”

The previous comments hint at one area of difference among the participants; that is, a greater tendency to evoke linguistic theory among the native-speaker teachers. Tom referred to theory most explicitly: “My approach is based on the hierarchy of willingness to communicate pyramid... If you want them to speak, you have to establish willingness to communicate so I approach the class from the bottom up, establishing rapport with each other”. Richard also highlighted theory, albeit positioning himself in opposition to a purely “communicative approach”, which he said was popular when he began his career. Rika and Xia, in contrast, did not directly mention theory and were less clear overall about their approaches. For example, Xia stated, “I

think for the past year here, because of the discussion program, I do speak less than before, but I'm not quite sure if it's my own approach or whether it's because of the discussion program."

There were also differences in terms of approaches to giving feedback. Richard seemed to prioritize feedback most clearly, stating, "I think it's a big part of the lesson. We'll focus on things they used well and we'll focus on my suggestions. But there will be some grammatical points as well, and vocabulary." Tom mentioned that in a previous course he would have liked to provide more feedback regarding intercultural issues, but felt obliged to base his feedback on that course's fixed method of assessment: "We were really handcuffed in terms of the amount and quality of the feedback we were giving because of the assessment that we were required to use." This issue seemed less of a priority to Xia and Rika, with the latter commenting, "I sometimes give verbal feedback right away when it's necessary, but not much."

A final theme that emerged was that of attitudes to L1 use in the classroom. Tom was the most positive, commenting that "I am probably more lenient than most teachers about L1 use... especially with lower proficiency classes, I want them to meta-reflect about language targets and what's going on in the class to make sure everybody understands." Rika, too, was quite tolerant of L1 use: "I think it's okay. Students think it's okay because they want to make sure [of] some important information in Japanese, and I think it's alright." Richard also indicated acceptance in limited situations: "It depends what the point is... if it's just for checking understanding quickly amongst themselves, I don't really have a problem with it at all." Xia conveyed the strictest attitude of the four, although again it was mainly due to contextual factors: "At the first university I was teaching it's not [that] I was very strict, it's the school was very strict. No Japanese in class, even during the break!" The discrepancies in attitude to L1 use are not surprising, however, given the variation among university teachers in Japan identified by Ford (2009) and Saito (2014).

### **Factors Influencing these Approaches**

Having established the participants' approaches to teaching speaking, let us now examine the factors that have influenced these approaches, using Borg's (2003) model of teacher cognition. Language learning experiences were referenced most directly by Xia, who mentioned having students read aloud in class, which was a technique she had used herself when learning English in China. Rika also brought up her own schooling, although more negatively, saying that she was given very few opportunities to actually speak and that "It was just like a school for translation. Teachers taught me how to translate English into Japanese every single day (laughs)!" Richard and Tom were less specific about their own experiences of school, although Tom stated, "I think it's vitally important for a language teacher to be a language learner. I think if you're

no longer a language learner you're going to lose touch with what it's like from a student's perspective." He then talked about the benefits of taking Japanese discussion lessons while employed as an English teacher. This comment resonates with Moodie and Feryok's (2015) suggestion that a commitment to language learning is linked to a commitment to language teaching and professional development.

Teacher training and education also seemed to have had a major impact on the non-native speaker participants. In particular, Rika was very positive about her experiences in the United States when studying for her Master's degree, and cited the pair and group work activities that she was exposed to there: "I think definitely these ways are something I learned in the U.S., not in Japan. In Japanese school, when I was a student, I almost never experienced that." Xia also mentioned some positive advice she was given when training to be a high school teacher in Japan: "I still remember the teachers at that high school told me 'You talk too much!' (laughs)... so that's why I always remind myself to talk less in class." Tom and Richard, on the other hand, were already experienced teachers before undergoing formal training. However, even experienced teachers can benefit from such training, as it can help to consolidate what they have learned in the classroom (Borg, 2011). Tom appeared to support this point when talking about his Master's program: "I think if you go out and try and do it and make your mistakes, and then get feedback and the theoretical grounding about why this works the way it works... it made sense to me."

All four participants were very clear about the impact of teaching experience. Xia cited the influence of her current teaching context, as well as her training, in her efforts to reduce her teacher talking-time. Tom appeared to learn a similar lesson through experience: "Initially, it's just so hard to get out of the way. It's so hard to stop talking. You're afraid of silence..." Richard mentioned his move away from pure CLT and towards a more skills and function-based approach to teaching, while Rika stated that she has become less strict as a teacher over time: "I changed a lot! When I started, I didn't know anything so I was very strict... I kind of regret (laughs). But now I'm not - as long as students enjoy learning it's okay." As experienced teachers, their principles and beliefs are likely to evolve more slowly from now on, as they become more consistent and entrenched over time (Breen et al., 2001).

Nevertheless, all beliefs and principles, whether entrenched through experience or not, are potentially subject to institutional demands and contextual constraints. Tom, for instance, expressed his frustrations about not being able to give the kind of feedback he wanted: "I felt restricted constantly. I felt an obligation to teach the class the way that they wanted it taught." Xia mentioned the strict English only policy at a previous workplace, but she also spoke positively about reducing her teacher-talk due to the requirements of the discussion course she

was currently teaching. In addition, all participants referred to the restrictions, but also the practical benefits, of teaching according to a unified curriculum. There are many references in the literature to the drawbacks and frustrations caused by such constraints (e.g., Kurihara & Samimy, 2007; Nishino, 2008) but the comments from this study's participants offer a reminder that institutional requirements are not always perceived negatively.

## Discussion

Having analyzed the data in terms of emerging themes, what patterns or commonalities can be observed among teaching approaches? Most strikingly, all four participants appeared committed to teaching speaking in a student-centered way, with fluency prioritized over formal accuracy. This finding is good news for those who despair at the enduring legacy of the grammar-translation or *yakudoku* method in Japan as teachers of all backgrounds, at least in this institution, seem to recognize the importance of meaning-focused output (Nation, 2007). It is interesting that Rika has broken the cycle of teacher-fronted lessons which she was exposed to at school in Japan. It might, however, be instructive that it was her experiences in the United States that enabled her to do so.

On the other hand, differences emerged regarding the extent to which these approaches were based on theoretical considerations. While both of the native-speaker teachers referred to theory, whether positively (regarding willingness to communicate) or negatively (regarding pure CLT), neither of the non-native-speaker teachers made any reference to pedagogical literature. This finding also mirrors the differing attitudes to feedback, with both of the native-speaker teachers seemingly more meticulous in their desire to have students reflect on activities. There were also differences regarding attitudes to L1 use, although most participants expressed tolerance towards it for the purposes of confirming or checking understanding.

The major influence on the participants' approaches to teaching speaking seems to have been practical experience, with all four reporting major changes in approach since they began teaching. Richard has become more focused on teaching skills, through presenting and analyzing functional language, rather than simply expecting students to learn by talking. Rika has become less strict, although she partly attributed this change simply to getting older and feeling less need to maintain a distance between herself and her students. Both Tom and Xia have learned to reduce their talk-time in class, even though Xia admitted she has a naturally talkative character. However, the influence of age and personality factors makes it difficult to draw conclusions about the effect of teaching experience alone.

All participants mentioned the effect of language learning and teacher training, although these experiences seem to have been more formative for the non-native-speaker teachers. As

well as Rika's study in the United States, Xia mentioned techniques from her language classes in China and her training in Japan that she was still trying to employ. In contrast, the fact that both Tom and Richard had been teaching for a number of years before becoming qualified seems to have reduced the impact of that training, further highlighting the significance of practical teaching experience. Contextual factors are also relevant, as all participants mentioned the need to adhere to the demands of particular institutions or programs in which they have taught.

## Conclusion

It is clear from this study that a combination of education, experience, and contextual factors, as well as internal factors such as age and personality, are responsible for the development of teacher cognition. Although these elements are dynamic and inter-connected, the result in this institution is that, regardless of background, teachers appear committed to maximizing student talking-time and fluency development. It is to be hoped that this trend exists among other universities too, given the pervasive focus on formal grammar and vocabulary study prior to university in Japan.

I hope that this study contributes to what is known about how speaking is taught in general, and especially at universities in Japan, both of which have been under-researched areas to date. Although I was unable to do so in this study, it would also be beneficial to compare stated principles and approaches to teaching with actual teaching practice (via classroom observations) and to relate both to learning outcomes.

## References

- Basturkmen, H. (2012). Review of research into the correspondence between language teachers' stated beliefs and practices. *System* 40, 282-295.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36 (2), 81-109.
- Borg, S. (2011). The impact of in-service teacher education on language teachers' beliefs. *System*, 39 (3), 370-380.
- Borg, S. (2015). *Teacher cognition and language education: Research and practice*. (2nd ed.). London, UK: Continuum.
- Breen, M. P., Hird, B., Milton, M., Oliver, H., & Thwaite, A. (2001). Making sense of language teaching: Teachers' principles and classroom practices. *Applied Linguistics* 22 (4), 470-501.
- Ford, K. (2009). Principles and practices of L1/L2 use in the Japanese university EFL classroom. *JALT Journal* 31 (1), 63-80.
- Fujimoto, D. T. (2019). The Japanese student and the university English teacher. In P. Wadden

- & C. C Hale (Eds.), *Teaching English at Japanese universities: A new handbook* (pp. 43-53). Abingdon, UK: Routledge.
- Hatch, J. A. (2002). *Doing qualitative research in education settings*. Albany, NY: State University of New York Press.
- Kurihara, Y., & Samimy, K. K. (2007). The impact of a U.S. teacher training program on teaching beliefs and practices: A case study of secondary school level Japanese teachers of English. *JALT Journal* 29 (1), 99-122.
- Moodie, I. & Feryok, A. (2015). Beyond cognition to commitment: English language teaching in South Korean primary schools. *The Modern Language Journal*, 99 (3), 450-469.
- Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1 (1), 2-12. doi:10.2167/illt039.0
- Nishimuro, M., & Borg, S. (2013). Teacher cognition and grammar teaching in a Japanese high school. *JALT Journal*, 35 (1), 29-50.
- Nishino, T. (2008). Japanese secondary school teachers' beliefs and practices regarding communicative language teaching: An exploratory survey. *JALT Journal*, 30 (1), 27-51.
- Nishino, T. (2012). Modeling teacher beliefs and practices in context: A multimethods approach. *The Modern Language Journal*, 96, 380-399.
- Numrich, C. (1996). On becoming a language teacher: Insights from diary studies. *TESOL Quarterly*, 30 (1), 131 – 53.
- Peacock, M. (2001). Pre-service ESL teachers' beliefs about second language learning: A longitudinal study. *System*, 29, 177 – 95.
- Phipps, S., & Borg, S. (2009). Exploring tensions between teachers' grammar teaching beliefs and practices. *System*, 37, 380 – 390.
- Richards, J. C. (1998). Teacher beliefs and decision making. In J. C. Richards (ed.), *Beyond training* (pp. 65 – 85). Cambridge, UK: Cambridge University Press.
- Roulston, K. (2010). *Reflective interviewing: A guide to theory and practice*. London, Sage.
- Saito, Y. (2014). Japanese university teachers' beliefs and practices regarding an English-only policy. *The Journal of Rikkyo University Language Center*, 31, 29-42.
- Sakui, K. (2004). Wearing two pairs of shoes: Language teaching in Japan. *ELT Journal* 58 (2), 155-163.
- Wiltshier, J., & Helgesen, M. (2019). Tearing down the wall of silence: Constructing the English conversation class at a Japanese university. In P. Wadden & C. C Hale (Eds.), *Teaching English at Japanese universities: A new handbook* (pp. 43-53). Abingdon, K: Routledge.



【研究ノート】

# デジタルアーカイブ・ポータル の構築・運営・展開の日米欧比較

－利活用における「ジャパンサーチ」の特徴と展望－

田中 絵麻  
TANAKA, Ema  
大江 宏子  
OE, Hiroko

## I はじめに 「ジャパンサーチ」の利活用に向けて

本稿のテーマは、2020年8月に運用が開始された日本のデジタルアーカイブ・ポータル・サイト（以下アーカイブ・ポータルと略す）である「ジャパンサーチ（Japan Search）」を取り上げ、その構築・運営・展開の観点から欧州連合（EU）の「ユーロピアナ（Europeana）」と米国の「米国デジタル公共図書館（Digital Public Library of America：DPLA）」と比較し、特に利活用における特徴と課題を考察することである。ユーロピアナの本運用の開始は2008年、DPLAは2013年であり、日本より12年と7年先行して開始されている（DPLA, 2013; Europeana, 2010）。

なお、デジタルアーカイブ・ポータルについての定義はまだ定まっていない模様であり、Rubin (2014)によると、「より最新の試みがウェブポータルだ。そこには何も定義はなく、ただポータルが多様な情報資源を集約する、ということがあるだけだ」としている（p.249）。文化情報学事典を参照すると、ポータルデータベースを「個別のデータベースを丸ごと含み込んで、データベース化し、典拠に当たる個別のデータに直接アクセスできるようにする形のデータベース」であるとしている（赤間, 2019, p. 399）。ポータル化が進められている要因としては、より信頼性の高い選別された情報、データへのアクセスが改善するという利点が存在していると考えられる。Rubin (2014)は、インターネット上の情報は組織化・選別されておらず、情報過多の状態にあるが、ポータル導入により、上質なオンラインの情報源へのアクセスを改善すると捉えている（p.251）。

本稿が比較分析を行う三つのポータルは、技術的には複数の図書館、美術館、博物館等のデータベースのメタデータを集約し、ポータル・サイトで提供するという点では共通している。また、いずれもメタデータの形式や著作権の面で国際標準に準拠している。しかし、技術面以外では、構築の経緯や関連政策等を含めた組織面、収蔵データの分類面、利用者に提供するサービス等の利用面では異なる点もある。McKemmish et al. (2019)は、各図書館・美術館において収蔵品が増加するにつれ、紙や実物での管理やレコードキーピングが困難となり、1970年代以降のデジ

タル化・電算化の潮流のなか、アーカイブの構築の組織や分類も現代化が進み、伝統的な方法や組織からのパラダイム・シフトが起きていると指摘する (pp.121-126.)。その意味で、アーカイブ・ポータルは、このパラダイム・シフトの渦中にあるといえ、組織面、方法面での分析が求められていると考える。

ジャパンサーチと先行する欧米のデジタル・ポータルを比較する視点をもつ先行研究としては、大向 (2020) が挙げられるが、海外の動向については概略の紹介となっている。大向 (2020) によると、ジャパンサーチの構築は、日本版ヨーロッパアーナとして位置づけられており、海外におけるデジタル・ポータルの潮流を念頭においているという (p.331)。また、「国内に存在するあらゆる種類の文化資源を対象とするジャパンサーチは、日本のデジタルアーカイブの系譜におけるひとつの到達点」であるとして、その重要性を指摘するとともに、2014年以降に検討が本格化し、運用開始に至った経緯をまとめている (p.329)。ただし、ジャパンサーチが、欧米のデジタル・ポータルと比較して、どのような特徴を持つかについては、管見の範囲では分析が途上にある。

奇しくも、ジャパンサーチの運用が開始された2020年は、世界でコロナ渦が拡大した年であり、テレワークやオンライン教育を始めとして、日本では進展の遅れが指摘されていたデジタル利活用が否応なく進展した。美術館や博物館も一時的な閉鎖や時短運用、入場人数の規制が行われた (文化庁, 2021d)。一方で、収藏品等のアーカイブのデジタル化を進めてきた美術館や博物館では、それらの資源を活用し、デジタル・ミュージアムによりコンテンツを配信する動きも見られるようになっている (田良, 2020)。田良 (2020) が指摘するように、従来の実物展示が中心の美術館・博物館は暗に来訪者を前提としていたが、デジタル化によりオンラインでのアクセス (ユーザー) の裾野の拡大、引いては美術館・博物館が目指す文化遺産の継承や活用にも資する。

ただし、個々の美術館や博物館が個別にデジタル化を推進したとしても、どこにどのようなデジタル情報 (作品) があるかを探すことはユーザーからみて手間がかかる。その意味で、複数のデジタルアーカイブのデータベースを連結するデジタル・ポータルの運用が開始されたことは、今後のデジタル・ミュージアムの利用拡大にも資するものと考えられる。実際、先行する欧米のアーカイブ・ポータルでは、利用者がアクセスしやすく、利用しやすい工夫を行っている。また、こうしたデジタル・ミュージアムやデジタル・ポータル利活用の推進を担う人材育成も重要になっている。

世界的にアーカイブ・ポータルの構築や運用が進展する理由としては、技術面の進歩も背景にある。詳細は本文で議論するが、現代におけるデジタル・ポータルの特徴としては、以下が挙げられる。

- a. ポータルサイト (Europeana、DPLA など) へのメタデータの集約
- b. 洗練された検索機能 (種類、テーマ、年、分野、フォーマット等) の提供
- c. 柔軟なデジタル・ライセンス (クリエイティブ・コモンズ・ライセンス等を採用)
- d. API (Application Programming Interface) の提供などの開発ツールの提供

- e. パーソナライゼーション（個人リストの作成、マイページ等）機能
- f. デジタル展覧会（デジタル・キュレーション、デジタル教育教材の提供）の提供
- g. 複数のツール（ブログ、SNS）によるユーザーへのリーチ拡大
- h. プロフェッショナルなネットワークとユーザーの活用支援

出所：Rubin (2014, p. 250) を参照し筆者作成。

こうした技術的特徴により、アーカイブ・ポータルを通じて、利用者にとってデジタル収蔵品の価値が向上すると考えられる。ジャパンサーチもユーロピアナ、DPLA も技術面では、共通する点が多い。ただし、各技術がどの程度活用されているかについては、差も見られ、a. から h. のうち、下にいくほど欧米のポータルが先行している。

一方で、日本のジャパンサーチには、欧米に見られない特徴や可能性もあると考える。日本のジャパンサーチの利活用では、「地域情報発信」での活用が目指されている点は、欧米のポータルとは異なる特徴である。また、ジャパンサーチの構築が急速に進展した背景として、IT 技術や AI 技術の活用を推進する一環として予算が割り当てられた点も注目される。加えて、欧米のアーカイブ・ポータルの収蔵品は、それぞれの文化的蓄積や文脈が反映されていることから、日本においても日本の特徴をもつアーカイブ・ポータルとして発展していくことが望ましいと考える。これらの点についても、欧米のアーカイブ・ポータルの点から議論する。

以上を踏まえて、本論では、先行するユーロピアナと DPLA、ジャパンサーチを対象として、①構築の経緯とミッション（目的）、②活動内容・参加機関・収蔵数、③各ポータルの特徴の三つの観点から比較することで特徴を抽出する。また、抽出した特徴を、技術面、組織面、分類面、利用面の四つの観点から整理したうえで、ジャパンサーチの利活用の展望を考察する。なお、各ポータルの調査では、2021 年 10 月現在の実サイトを参照している。

## II EUの「ユーロピアナ」の構築・運営・展開

### 1 構築の経緯とミッション

#### (1) 構築の経緯

ユーロピアナの構築の経緯等の基礎的な情報は、サイト上で公表されている。ユーロピアナによると、同ポータルの構築の経緯は以下のとおりである (Europeana, n.a.-a)。

- ・2005 年：フランスのシラク大統領を含む 5 人の国家元首が署名した書簡により、欧州連合の関係者に欧州デジタルライブラリーの開発を支援するよう要請。
- ・2008 年 11 月 20 日：ユーロピアナのプロトタイプ公開。博物館、視聴覚資料室、ギャラリーなどがライブラリーに加わり、ヨーロッパの文化遺産への共通のアクセスポイントとなる。開始時の収蔵デジタルオブジェクト数は 450 万。

- ・2011年1月：欧州委員会は「ニュー・ルネッサンス」レポートを発表。ユーロピアナを「ヨーロッパのオンライン文化遺産の中心的な参照レポート」として支持を表明。
- ・2012年9月：ユーロピアナのメタデータをクリエイティブ・コモンズ・ライセンス（Creative Commons CC0 1.0 Universal Public Domain Dedication）に基づいて公開。メタデータ自体の自由な利用が可能になる。
- ・2015年5月：欧州委員会はユーロピアナをデジタルサービスインフラ（DSI）の一つに位置づけ。DSIは市民、企業、公共機関のためにネットワーク化された国境を越えたサービスを提供。

なお、「ニュー・ルネッサンス」レポート内では、Europeanaが219回、文化遺産（cultural heritage）の用語が450回言及されており、ユーロピアナにおいて、文化遺産が重要視されていることがわかる（Reflection Group on Bringing Europe's Cultural Heritage Online, 2011）。

## (2) ミッションとビジョン

ユーロピアナのミッションとビジョン以下のように掲げられている（Europeana, n.a.-c）。

### ・ ミッション

ユーロピアナは、デジタル変革（Digital Transformation）のさなかにある文化遺産部門を支援する。デジタル変革と連携をもたらす専門性、ツール、政策を構築し、イノベーションをうながす。人々が、教育、研究、創造、娯楽のために、文化遺産をより簡単に利用できるようにする。私たちの活動は、開かれた、知識豊富で創造的な社会の実現に貢献する。

### ・ ビジョン

ユーロピアナは、文化遺産部門がデジタルにより活性化され、また、欧州が文化により活性化され、回復力がもたらされ、経済が成長し、雇用が増加し、ウェルビーイングと欧州人としてのアイデンティティを改善することを想い描いている。

## 2 活動内容・参加機関・収蔵数

### (1) 活動内容

ユーロピアナの活動状況についても同サイトを参照すると、以下のようにまとめられている。活動内容は、ポータル構築にとどまらず、関係者の連携促進やオーディエンスへのリーチ促進に加えて、技術開発や利用の裾野を拡大する活動まで射程に含み、包括的な活動を行っているといえる。

- ・ 協力体制の構築（We collaborate）：多様なプロジェクトやパートナーシップ、リアルとデジタルのイベント・プログラムにより、欧州のあらゆる分野の文化遺産の専門家と協力。
- ・ 啓発活動（We advocate）：文化遺産の開放性、透明性、再利用を支援するよりよいデジタル

活動について啓発。

- ・ オーディエンスへのリーチ (We reach audience) : 文化遺産機関が、既存・新規の視聴者とオンラインでつながることを可能にする。すべての文化遺産機関が標準化されたフォーマットで質の高いデジタル資産を作成することを支援し、21世紀のミッションにあった形でコレクションを共有、探索、照会、使用可能にする。
- ・ キャパシティビルディング (We build capacity) : 文化遺産分野の機関や個人に、デジタル技術や実践を開発する機会を提供する。文化遺産組織のデジタル変革を支援するツールやサービス、イベントを開発するために協力。
- ・ 技術開発 (We build technology) : デジタル文化遺産を展示、共有、利用するための技術的なソリューションを開発、運営。だれでも利用可能な一連の API、Europeana Pro などのほか、組織内向けのシステムやプロセスを設計。

なお、ユーロピアナの運営は、ユーロピアナ・ファウンデーション (Europeana Foundation) が担い、主に EU からの補助金で運営されており、年間予算は 600 万ユーロ程度である。ただし、近年の予算は減少傾向にあり、2017 年の年間予算は約 637 万ユーロ、2020 年は約 592 万ユーロである (Europeana Foundation, 2018, 2021)。

## (2) 参加機関とアクセス可能な収蔵品数

ユーロピアナは、欧州の政策の枠組み内で構築・運営されており、EU 加盟国の参画により実現した。文化遺産機関がそれぞれの国の文化政策の一環として行っている活動を直接補完し、支援している。これにより、ユーロピアナには、欧州全土から約 4,000 の機関が参画しており、2020 年初頭の時点で、書籍、音楽、芸術作品などの 5,800 万点のデジタルオブジェクトへのアクセスを提供している。

## 3 ユーロピアナの特徴

### (1) ユーザー向けポータルとしての特徴

ユーロピアナは、リニューアルされているが、従来のサイトもクラシック・サイトとしてアクセスが提供されている。リニューアルされたサイトと従来のサイトと異なる部分もあるため、両方のサイトを確認しつつ、その特徴を抽出する。

まず、クラシック・サイトでは、図 1 に示すように、ナビゲーションで、「コレクション」、「探索」、「展示」、「ブログ」の分類が示されており、その下に横断的な検索が可能なサーチ・ボックスが配置されている。一方、新サイトでも「コレクション」は共通しているが、「ストーリー」のナビが追加されており、検索は一階層下に移動している。また、新サイトでは、最近の展示や取り組みをトップページで紹介しており、より新しい情報提供を意識していることがわかる。「コレク

ション」からアクセス可能な情報は、収蔵作品の分類（1914-1918、考古学、アート、ファッション、産業遺産、文献、地図と地理学、移民、音楽、自然史、新聞、写真、スポーツ）であり、新旧サイトで多少順番は異なるが内容は同じである。新サイトで追加されている「ストーリー」は旧サイトの「展示」と「ブログ」を統合したコーナーとなっている。

収蔵コンテンツのメディア別の収録数は、画像が最も多く約 3,680 万件、次いでテキストが約 2364 万件、動画が約 111 万件、音声約 85 万件、3D が 2 万 4,000 件である（2021 年 10 月現在、ユーロピアナの検索ページの表示より）。

そのほか、新サイトで追加されている要素としては、図 2 の左上にある横三本線をクリックすると表示される「教員向け」と、ログイン機能が挙げられる。いずれも旧サイトのトップページになく、新機能といえる。教員向けのコンテンツはまだ数は多くないが、遠隔教育やデジタル学習、教育方法のトレンドといった分類で提供されており、デジタル教育を行ううえで教員が参考になるとと思われるコンテンツが揃えられている。また、ログイン機能は、メールアドレスでユーザー登録すると利用可能になる機能で、アクセスしたコレクションやデジタル収蔵品を「My Collection」として記録しておくことが可能である。

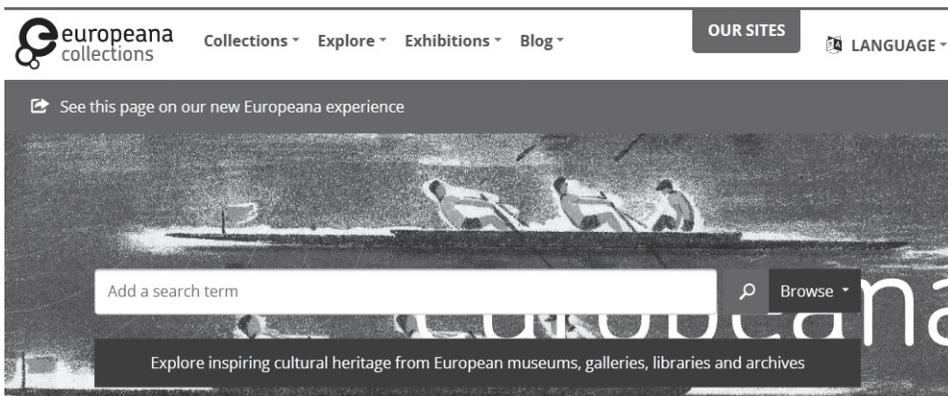


図1 ユーロピアナのクラシック・サイトのトップページ  
出所：Europeana classic site (<https://classic.europeana.eu/portal/en>)

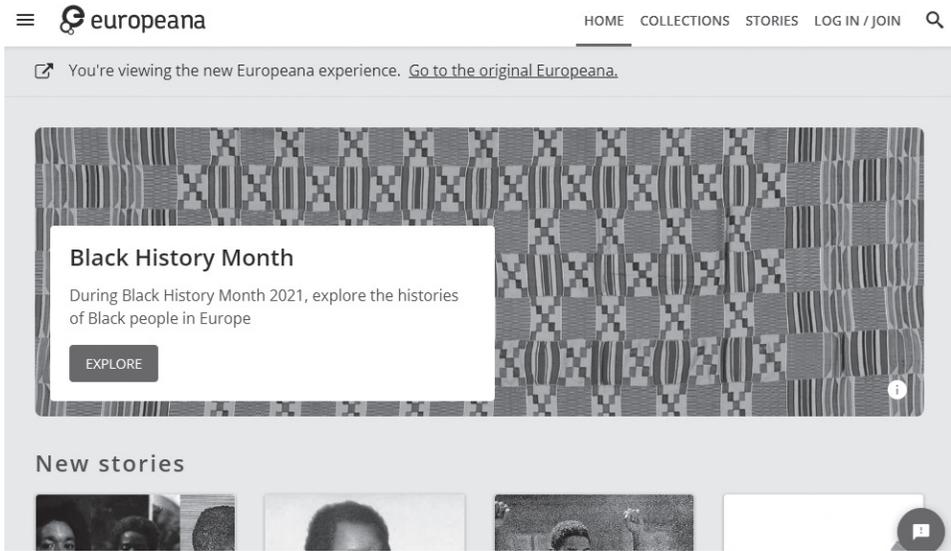


図2 ユーロピアナの新サイトのトップページ  
出所：Europeanana new site (<https://www.europeana.eu/en>)

## (2) 参加機関向けネットワーク・ハブとしての特徴：Europeanana Pro

ユーロピアナのミッションに掲げられているように、ユーロピアナは単にデジタル化された収蔵品へのアクセスをユーザーに提供するに留まらず、参加機関向けのネットワーク・ハブとしての機能も有している。参加機関向けには、ユーロピアナ・プロのウェブサイトが別に用意されており、サービスやツール、イベントの情報の他、欧州の戦略文書や参画方法などの情報が掲載されている。

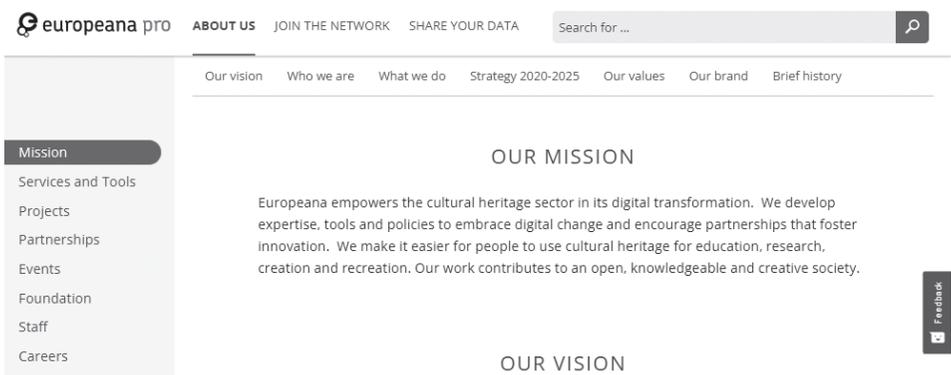


図3 ユーロピアナ・プロのトップページ  
出所：Europeanana Pro site (<https://pro.europeana.eu/>)

参加機関は、このプロサイトにアクセスすることで、必要な情報にアクセスできるだけでなく、各種サービスやオンライン・ツールも利用可能になっている。サービスに含まれるアイテムとしては、「デジタル・キャパシティ・ビルディング」、「著作権とデジタル文化遺産」、「教育」、「インパクト」、「ユーロピアナ」、「ウェビナー」、「研究」、「Covid-19における活動」があり、参加機関の能力向上や連携を促進するサービスが充実しているといえる。

なお、ユーロピアナの2020年から2025年の期間を対象とした戦略では、基盤強化、データの品質の向上、キャパシティ構築の三つの目標が掲げられており、さらにユーロピアナを中核とした文化遺産部門のデジタル変革を推進する姿勢が読み取れる (Europeana, 2020b)。

### (3) 開発者向けツールの特徴

加えて、ユーロピアナでは開発者向けのツールも充実している。2021年10月現在では、以下の15種類のツールが提供されている (Europeana Pro, n.a.)。アプリケーション開発に用いることができるAPIのほか、ユーロピアナのコンテンツを魅力的に伝えるために活用可能な各種ツールが揃えられている。

- ・ API：ヨーロッパ中の機関から集められた豊富なコレクションからアプリケーションを構築するために Europeana API を利用可能
- ・ CrowdHeritage Platform：デジタル化された文化遺産のメタデータを充実させるためのオンライン・クラウドソーシング・プラットフォーム
- ・ Culture Chatbot：文化遺産を扱う機関のための仮想アシスタント（チャットボット）を構築するためのプラットフォーム
- ・ Curated datasets：新聞、書籍、写真、美術品、工芸品、オーディオクリップなど、優れたコンテンツをまとめたデータセット
- ・ Dashboard：3Dモデルやメタデータを共有し、Europeanaでのレビューや公開に活用可能なオンライン・ツール
- ・ Enhanced Unified Playout Service：AV文化遺産コンテンツを再利用するための一連の機能を提供するユニバーサルオーディオビジュアルプレーヤー
- ・ Europeana Essentials Slide Deck：Europeana イニシアチブの活動に関するプレゼンテーションを支援するための、さまざまなトピックを網羅したスライド資料集
- ・ Europeana Archaeology Vocabulary Services：メタデータ用語とオープンデータ資源を相互にリンクするためのオンライン・マッピング・ツール
- ・ Events toolkit：デジタル文化遺産に関するオンラインイベントを成功させるためのガイダンス、ヒント、リソース
- ・ Historiana：文化遺産コレクションを教育に活用するためのオンライン学習プラットフォーム
- ・ LinBi Enrichment Platform：Europeana Data Model ガイドラインに沿って、リンク、拡張、新

規オブジェクト作成のためのツールを提供し、データの拡張をサポートするプラットフォーム

- ・ MotionNotes：アートやダンスのパフォーマンスにおけるデジタル技術の利用を促進する、マルチモーダルなビデオアノテーター<sup>1</sup>
- ・ MovesCollect and MovesScrapbook：Europeana上のコンテンツを利用するための新しい方法を提供する、相互に接続されたツール
- ・ Seven tips for digital storytelling：オンライン上の文化遺産を用いて魅力的で感動的なストーリーを語るためのリソース
- ・ Storymaker：専門家でなくても、3Dコンテンツを使ってストーリーを制作可能なツール
- ・ Transcribathon：デジタル化された資料をより豊かにするためのプラットフォーム

### Ⅲ 米国の「DPLA」の構築・運営・展開

#### 1 構築の経緯とミッション

##### (1) 構築の経緯

米国のDPLAの構築が本格化したのは、2010年10月にマサチューセッツ州ケンブリッジで開催された会議が契機となっている(DPLA, n.a-e)。同会議には、図書館、財団、学会、テクノロジープロジェクトのリーダー40名が参加し、それまで個別に構築されサイロ化していた文化遺産のデジタル・コレクションを集約・活用するための、オープンな分散型ネットワークを構築することで合意した。その後、2010年12月には、ハーバード大学パークマン・クラインセンターが、アルフレッド・P・スローン財団の支援を受け、図書館、テクノロジー、法律、教育の各分野の専門家を招集して構築プロジェクトを開始、同センターを中心として、草の根コミュニティが形成され、数百人の公共・研究図書館員、イノベーター、デジタル・ヒューマニスト、その他のボランティアが6つの活動領域を分担し、DPLAの範囲、設計、構築を行った。DPLAは、開発開始から約2年半後の2013年4月に公開され、フリーで、オープンで、アクセス可能なナショナル・デジタル・リソースとなった。

なお、2021年10月現在では、DPLAを支援する財団は、アルフレッド・P・スローン財団のほか、ナイト財団、アンドリュー・W・メロン財団、ピボタル・ベンチャーズも名を連ねている(DPLA, n.a-d)。

##### (2) ミッションと長期的成功

DPLAの2019年から2022年の期間を対象とした戦略によると、DPLAのミッションと長期的成功は次のとおりである(DPLA, 2019)。

---

1 動画への注釈や描画を追加可能な機能。

#### ・ ミッション

DPLA は、共有する歴史、文化、知識へのアクセスを最大化することで、人々が学び、成長し、多様でより機能的な社会に貢献することを後押しする。

#### ・ 長期的成功

DPLA は、すべてのプロジェクトが持続可能であるように活動する。責任を持ってリスクを取ることを信念としている。小規模で機敏な非営利団体である DPLA は、現場を代表して新しい技術を試し、新しいアイデアを模索することができるユニークな立場にある。DPLA は、パートナーやより広い領域のため、注意深く計算し、リスク、コスト、利益のバランスを考慮する。私たちの価値と使命を追求する能力は、私たちの組織と協力者が私たちの活動を維持する能力に直接つながっている。

## 2 活動内容・参加機関・収蔵数

### (1) 活動内容

DPLA の説明によると、DPLA は米国で最も信頼されている知識の共有源である図書館と文化機関の価値を高めることを目的として、パートナーと協力して、図書館が情報をより利用しやすくするための革新的なツールやアイデアを推進している。また、DPLA は、全米のパートナーと協力して以下の活動を行っている (DPLA, n.a.-a)。

- ・ 国内の図書館、アーカイブ、博物館、その他の文化施設から提供される何百万もの資料を、誰もがワンストップで発見できるようにする。
- ・ 図書館が電子書籍やその他の電子コンテンツを購入、整理し、利用者に提供できるように、図書館が管理する市場とプラットフォームを提供する。
- ・ 図書館のリーダーや実務者を集め、コミュニティに貢献し、情報を提供し、力を与える技術を探求し、発展させる。

なお、DPLA は、独立した非営利組織であり、財団基金のほか、有料会員収入、寄付によって運営されている。2018年度(2018年7月1日 - 2019年6月末まで)の総収入は249万429米ドル、総支出は224万9,071米ドルであった(DPLA, 2021)。運営は主に支援財団からの拠出金に支えられているが、以下のような収入も計上されている。

- ・ ハブ・ネットワーク有料会員収入：26万7,353米ドル
- ・ eブックからの配分：3万5,070米ドル
- ・ ソフトウェア開発：1万9,500米ドル

### (2) 参加機関とアクセス可能な収蔵品数

DPLAには、直接DPLAに接続してメタデータを提供するのではなく、中間アグリゲータのサービス・ハブ経由で参加する機関が多い。直接DPLAに接続するのは、大規模なコレクションを持つコンテンツ・ハブである(神崎, 2020, p. 343)。2021年10月現在のサービス・ハブは34機関、コンテンツ・ハブは15機関である(DPLA, n.a-g)。

なお、本稿の調査の範囲では、DPLA参加機関数は不明であるが、時実(2015)によると2015年時点で約1,600機関が参加していたという(p.314)。また、当時の収蔵コンテンツ数は約996万件となっていた。なお、DPLAトップページによると、2021年10月現在、DPLAの収蔵コンテンツ数は約4,492万件に達しており、拡大を続けている(Ⅲ 3. (1)の図4参照)。

### 3 DPLAの特徴

#### (1) ユーザー向けポータルとしての特徴

図4は、DPLAのトップページである。ユーロピアナのクラシック・サイトのように、サーチボックスが中央に配置されている。これにより、米国中のテキスト、画像、動画、音が検索可能であることが一目でわかる。また、トピック別の検索も提供されており、南北戦争、航空、野球、公民権運動、フード、1840年代以降の想像力、写真、科学における女性の8つのテーマが示されている。さらに、リンクされている「展示」のページに遷移すると、米国に特徴的な32のテーマ別(月への競争、パーソナルカメラの発展、米国のラジオの黄金時代等)でコンテンツへアクセスできるようになっている。

DPLAの収蔵コンテンツのメディア別の内訳は、テキストが最も多く約2,338万件、イメージが約1,180万件、動画が約26万件、音声が約1万5,000件、物理的オブジェクトが約8万である(2021年10月現在、DPLAの検索ページより)。

また、DPLAの特徴として、電子書籍に力を入れていることも挙げられる。DPLAでは、オープン・ブックシェルフの名称で、約1万冊の電子書籍へのアクセスを提供している。この1万冊は、図書館員や専門家が選定したもので、幅広く読まれ支持されているタイトルが選ばれている(DPLA, n.a-f)。

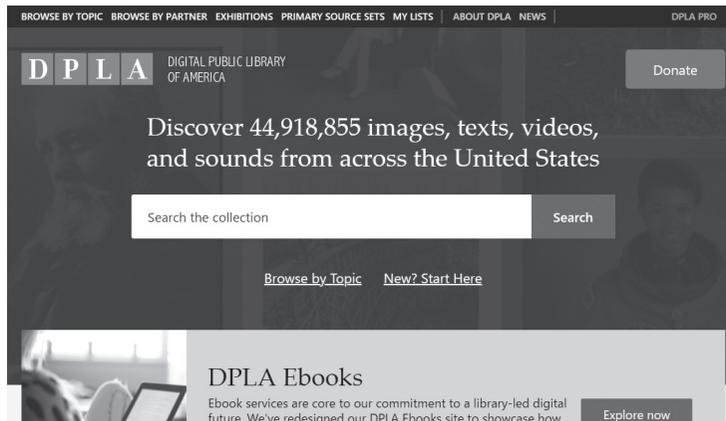


図4 DPLAのトップページ  
出所：DPLA (https://dp.la/)

## (2) 教育向けコンテンツの特徴

教育コンテンツの開発にも積極的で、ナビゲーション・バーの左から四つ目の「一次ソースコレクション (Primary source collections)」のコーナーでは、教育用にトピック別にまとめられたセットが多数提供されている。各セットには、クラス内で利用可能な指導要領 (Teaching Guide) も添付されている。一次ソースコレクションの分類は、米国の歴史、アメリカ文学、科学技術、世界史、芸術、アフリカン・アメリカン、アジアン・アメリカン、労働史、ラテン・アメリカン、法律と政府、移民、ネイティブ・アメリカン、社会運動、女性の14テーマとなっている。各テーマには、サブテーマが設定されており、関連する画像、動画、テキストといった一次資料のほか、二次資料へのリンク、指導要領がまとめられている。

また、トップページの下の方には、「DPLA の使い方」のコーナーがあり、DPLA の用途として、教育、ファミリー・リサーチ、生涯学習、学術研究の分類で、使い方のガイドが提供されている。

## (3) 開発者向けツールの特徴

DPLA においても、開発者向けの以下の四種類のツールや情報が提供されている。文書は充実しているが、開発者や技術者向けの内容が多いと言える。

- ・ API コーデックス：DPLA の API は、DPLA カタログのすべてのアイテムへのプログラムによる検索とアクセスを提供。API を使用し、アプリケーションの強化、DPLA をカスタマーのポータルに接続、データ取得による独自構築が可能
- ・ 一括ダウンロード (bulk downloads)：一括ダウンロードにより、分析や調査を行うほか、DPLA のアグリゲーション機能を用いて独自のキュレーション体験を構築可能
- ・ 技術文書：技術スタック、プロセス、および展開についてのドキュメントを wiki で公開
- ・ サンプルコードとライブラリー：DPLA API による開発用サンプルコードとライブラリー

## IV 日本の「ジャパンサーチ」の構築・運営・展開

### 1 構築の経緯とミッション

#### (1) 構築の経緯

日本においては、2000年代前半より、国立国会図書館がデジタルアーカイブ・ポータルの構築に取り組んでおり (中山, 2004)、2005年にはプロトタイプ的一般公開が行われた (久古 et al., 2006)。その後、2014年には文化庁に「文化関係資料のアーカイブに関する有識者会議」が設置され、組織や分野を超えたアーカイブの利活用の推進が進められた。ただし、当時のポータル・サイトの仮称は、「文化ナショナル・アーカイブ」となっている (大向, 2020)。大向 (2020) の分析で注目されるのが、2014年の検討では、政府は国立国会図書館がつなぎ役となり、全国の博物館・美術館、大学・研究機関・民間施設等と連携する方針を示したものの、なぜ図書館以外の美術館等の連携において、国立国会図書館がつなぎ役になるのかという疑問が提示されつつも明

示的には整理されなかったとの指摘である (p. 331)。

大向 (2020) も整理しているように、ジャパンサーチの構築が本格化した契機としては、2017年6月に閣議決定された「未来投資戦略 2017」が挙げられる。未来投資戦略の第2項目の具体的施策の一つである、「Ⅲ 地域経済好循環システムの構築」のなかの「3. 観光・スポーツ・文化芸術・iii) 文化芸術資源を活用した経済活性化」のなかに、「デジタルアーカイブの構築」が置かれている。なお、iii) 文化芸術資源を活用した経済活性化には以下の三つの項目があり、デジタルアーカイブ構築の位置づけを読み取ることができる (閣議決定, 2017, p.163)

### iii) 文化芸術資源を活用した経済活性化

#### ① 文化芸術資源の活用の更なる促進に向けた体制・制度の整備

#### ② 文化芸術資源を核とした地域活性化・ブランド力向上

#### ③ コンテンツを軸とした文化芸術産業の強化

- ・ 地域コンテンツの新たな市場開拓のため、急拡大するアジアのコンテンツ市場開拓に向けた各国との官民対話を拡充するとともに、国内外におけるビジネスマッチングイベントの開催や、業界団体等とともに VR/AR 等の先進的なコンテンツ技術を活用するためのガイドラインを整備する。
- ・ コンテンツ産業や観光の振興、地方創生等につながる映画やマンガ・アニメ・ゲーム等のメディア芸術分野の国内外への発信機能の強化等を図る。
- ・ 我が国の知的資源・文化芸術資源を一元化し新規ビジネス・サービスを創出するため、各分野でのデジタルアーカイブ化や、国立国会図書館を中心とした分野横断の統合ポータル構築を推進する。

なお、ジャパンサーチ構築経緯にかかる文書や資料は、ジャパンサーチのウェブページでも公開されているが、ジャパンサーチ (仮称) が用いられることとなった2017年以降の文書が掲載されており、上述の未来投資戦略には言及がない (ジャパンサーチ, 2021a)。同ページに掲載されている「知的財産推進計画 2017」以降の政策文書では、国立国会図書館をハブとした各機関の連携によりジャパンサーチを構築することが既定路線となっており、同時に、知的財産戦略本部にデジタルアーカイブジャパン推進委員会と実務者検討委員会が設置され、これらの委員会の検討に基づいてジャパンサーチの構築が進められた (首相官邸, n.a.)。

## (2) 概要と目的

ジャパンサーチのウェブページによると、ジャパンサーチは、書籍等分野、文化財分野、メディア芸術分野のデジタルアーカイブを連携させ、日本の多様なコンテンツのメタデータをまとめて検索できる「国の分野横断型統合ポータル」であると説明されている。また、集約したメタデータを利活用しやすい形式で提供し、コンテンツの利活用を促進する基盤 (プラットフォーム) と

しての役割もあるとしている(ジャパンサーチ, 2020a)。

また、ジャパンサーチの目的としては、デジタルアーカイブを文化の保存・継承・発信のみならず、観光、地方創生、教育研究、ビジネスへの利用等の新たな価値創出、イノベーション推進にも貢献するものとしている。加えて、デジタルアーカイブの連携と横断型検索による、デジタル情報資源の効率的活用が目指されている(Ibid.)。

## 2 活動内容・参加機関・収蔵数

### (1) 活動内容

本稿執筆時点において、ジャパンサーチは正式に公開されてから、1年強の期間しか経過しておらず、連携機関の拡充や広報活動が活動の中心となっている。ジャパンサーチのウェブサイトによると以下の活動が行われている(ジャパンサーチ, 2021b)。

- ・ 正式版の公開：2019年2月27日に試験を公開し、2020年8月25日に正式版を公開。正式版公開時にはマイノットの機能拡張とギャラリーの共同編集機能を追加。
- ・ 連携状況：2020年度に9機関、44データベースと新たに連携。
- ・ 利活用データへの変換：2021年3月末現在、連携機関がAPI提供を許可している100データベース由来のメタデータ全件(約2,000万件)を利活用データに変換済。
- ・ ギャラリーの拡充：2020年7月にギャラリー100件を追加公開。2021年3月末現在、公開ギャラリー数は219件。
- ・ 利活用促進に向けた取り組み：2020年12月から連携機関以外の個人や組織もプロジェクト機能、ワークスペース機能が利用可能に。
- ・ GIF IT UP 2020：GIFアニメーションの作成コンテスト「GIF IT UP 2020」に協力。
- ・ 広報活動：イベントの開催、講演、記事執筆、SNSでの発信、表彰。

本稿の調査の範囲では、ジャパンサーチの運営費は国立国会図書館予算のデジタルアーカイブ事業で計上されている。2018年の同予算額は1億6,092万円、2019年は約1億9,145万円であった(国立国会図書館, 2019)。また、デジタルアーカイブの利活用の促進の項目で計上されている2020年度の予算額は5,544万円、2021年は6,996万円となっている(国立国会図書館, 2021)。

### (2) 参加機関とアクセス可能な収蔵数

ジャパンサーチのウェブサイトによると、2021年10月現在、連携している機関は70機関である(ジャパンサーチ, n.a.-c)。表1は、70機関を組織の形態別に分類したものである。

表1 ジャパンサーチの連携機関（組織形態別）

組織形態	連携組織名
独立行政法人	独立行政法人国立公文書館、独立行政法人国立美術館、奈良文化財研究所、独立行政法人国立文化財機構、東京文化財研究所
公益財団法人・一般社団法人	公益財団法人放送番組センター、公益財団法人 渋沢栄一記念財団、一般社団法人 全国美術館会議、公益財団法人ポーラ美術振興財団 ポーラ美術館
国立	九州国立博物館、国立科学博物館、国立国会図書館、国立教育政策研究所教育図書館、国立民族学博物館、奈良国立博物館、国立歴史民俗博物館、国立国語研究所、東京国立博物館、京都国立博物館
大学共同利用機関	国文学研究資料館、大学共同利用機関法人 人間文化研究機構、総合地球環境学研究所、国際日本文化研究センター
県立図書館・県立等	岩手県立図書館、群馬県立図書館、福井県文書館・福井県立図書館、茨城県立図書館、徳島県立図書館、千葉県立中央図書館、札幌市中央図書館、宮城県図書館、大阪市立図書館、鳥取県立図書館、青森県立図書館、山梨県立図書館、県立長野図書館、佐賀県立図書館、鳥取県、三重県、和歌山県立文書館、奈良県立図書情報館
大学・大学付属組織・大学図書館	東京農工大学科学博物館、慶應義塾美術品管理運用委員会、琉球大学附属図書館、東京大学、京都大学研究資源アーカイブ、新潟大学 地域映像アーカイブ研究センター、早稲田大学 坪内博士記念演劇博物館、慶應義塾大学 文学部 民族学考古学専攻、慶應義塾ミュージアム・commons、慶應義塾大学、関西大学アジア・オープン・リサーチセンター（KU-ORCAS）、九州産業大学図書館、立命館大学アート・リサーチセンター、慶應義塾福沢研究センター、慶應義塾大学メディアセンター
官庁	総務省行政管理局、国土交通省国土地理院、文化庁
美術館	東京富士美術館、大阪市立東洋陶磁美術館、愛知県美術館
その他	縄文遺跡群世界遺産保存活用協議会、青森県環境生活部、NHK アーカイブス、特定非営利活動法人映像産業振興機構、日本写真保存センター、TRC-ADEAC 株式会社、沖縄県南城市教育委員会、南方熊楠顕彰館

出所：ジャパンサーチより筆者作成。

また、連携しているデータベース数は2021年10月現在142であり、メタデータ件数は約2343万件である（ジャパンサーチ，n.a.-b）。

### 3 ジャパンサーチの特徴

#### (1) ユーザー向けポータルとしての特徴

図5は、ジャパンサーチのトップページである。Japan Searchの文字の上に、スライドショー形式で主要な収蔵コンテンツのイメージと説明が表示されている。サーチ・ボックスは左上に配置されており、最初にアクセスしたときの情報量は少ないが、日本のイメージと検索サイトであることが直感的にわかる構成といえる。



図5 ジャパンサーチのトップページ  
出所：ジャパンサーチ (<https://jpsearch.go.jp/>)

スクロールすると、図6に示すようにトップページ株には、ジャパンサーチの概要や利用方法、ギャラリーや連携機関、収録データベースの情報が並べられている。さらに下部にスクロールすると新着情報のコーナーがあり、その下には、「ギャラリー」、「データベースを見つける」、「テーマ別検索機能を使う」、「施設・機関を見つける」のナビゲーションが置かれている。



図6 ジャパンサーチのトップページ下部  
出所：ジャパンサーチ (<https://jpsearch.go.jp/>)

テーマ別検索機能では、18のテーマが設定されているが、最初に「日本のエンタメコンテンツを調べる」が置かれた次に「南画検索」が、その下に「刀剣について調べる」があるなど、発

展途上の様子となっている。また、テーマ別検索となっているが、「教育・商用利用検索」、「全国美術館会議横断検索」、「国立国会図書館サーチ」等の機関・用途別の分類も含まれている。

ジャパンサーチでは、検索ページにおいてメディア別の収録件数の情報は提供されていないが、総収録数は約 2,343 万となっている。うち、書籍等が約 1,110 万と公文書が約 404 万含まれるため、テキストが占める割合が大きいと言える(表 2)。

なお、ジャパンサーチの統計情報によると、以下の分野別にメタデータを収録している。上述の「エンタメコンテンツを調べる」には、アニメ、映画、ゲームのメタデータへのアクセスが提供されているが、収録されているのは JACC アーティスト、JACC 脚本検、JACC クリエイターの情報であり、俳優や脚本家などのクリエイターの人物情報となっている。

表 2 ジャパンサーチの分野別メタデータ件数と収録データベース例・数(重複分類あり)

分野	メタデータ件数	収録データベース例	データベース数
アニメ	3,320	JACC アーティスト検索、JACC クリエイター検索、日本アニメーション映画クラシックス	3
映画	81,650	JACC アーティスト検索、国立映画アーカイブ所蔵映画フィルム検索システム等	7
ゲーム	3,256	JACC アーティスト検索、JACC クリエイター検索	2
公文書	4,045,880	国立公文書館デジタルアーカイブ等	3
自然史・理工学	5,410,643	魚類写真資料データベース、西表文献、サイエンスミュージアムネット、地球研アーカイブズ等	12
書籍等	11,105,506	国立国会図書館デジタル・コレクション、全国書誌等	43
人文学	3,388,819	映像資料目録、音響資料曲目、音響資料目録等	64
地域	548,984	三重の歴史・文化散策マップ(地域文化資産)、 にいがた 地域映像アーカイブ データベース等	34
地図	3,288	古地図コレクション等	2
データセット	27,674	データカタログサイト、映像資料	2
美術	525,040	愛知県美術館コレクション、ポラ美術館コレクション、国立美術館所蔵作品総合目録等	15
舞台芸術	80,844	演劇博物館名品セレクション、3D データベース等	8
文化財	412,977	三重県埋蔵文化財センター所蔵品、大阪市立東洋陶磁美術館収蔵品画像オープンデータ等	20
放送番組	81,168	放送ライブラリー公開番組DB、JACC 脚本検索	6

出所：ジャパンサーチより集計・作成(ジャパンサーチ, n.a.-b)

### (3) 連携機関向け・開発者向け情報

ジャパンサーチにおいても、連携機関向け、開発者向け情報を提供している。しかし、連携機関向けには、連携方針や連携手順、データの登録方法の情報が中心となっている(ジャパンサーチ, 2020c)。また、開発者向け情報としては、利活用スキーマ概説、SPARQL エンドポイント解説、簡易 Web API 概説の三つが、開発者向け API としては、SPARQL エンドポイント(<https://jpsearch.go.jp/rdf/sparql/>)と EasySparql(<https://jpsearch.go.jp/rdf/sparql/easy/>)の二つが提供されており、利活用に向けた開発環境の整備が進められている(ジャパンサーチ, n.a.-a)。

### (3) 開催イベントにみる地域振興との関わり

ジャパンサーチは、本格運用開始から日が浅いこともあり、連携機関数の拡充や、連携機関向けサービスの提供、開発者向けツールの開発は、その途上にあると考えられる。一方で、立ち上げ・拡充時期における活動ではジャパンサーチに特徴的な活動も見られる。ジャパンサーチが参画・開催しているイベントで、ジャパンサーチのサイトで紹介されているのは、以下の4イベントである。

- ・ 2020年9月：正式版公開記念オンラインイベント「ジャパンサーチの挑戦～ポストコロナ社会とデジタルアーカイブ～」(内閣府知的財産戦略事務局)
- ・ 2020年11月：「ジャパンサーチ正式版公開～書籍等分野の連携及び利活用拡大に向けて～」(第22回図書館総合展で開催)
- ・ 2020年10月～12月：「2020アーバンデータチャレンジ京都：ジャパンサーチ・タウン」
- ・ 2021年3月：「ジャパンサーチを使ってみた！～教育・研究・地域情報発信の現場から～」

上の二つはジャパンサーチの運用開始にかかる概説的な内容となっており、下の二つはジャパンサーチの利活用促進のためのイベントである。

## V 欧米との比較からみるジャパンサーチの特徴

### 1 ポータルとしての技術面：国際標準への準拠

「はじめに」で整理したように、2000年代以降に構築が進化したアーカイブ・ポータルは、複数のオンラインデータベースをメタデータで連結し、利用者がワンストップで検索等を可能にし、利便性を高めている。その際、技術面でみると、公共的な性質を持つアーカイブ・ポータルでは、オープン・フォーマットの採用が望ましいとの指摘もある(デジタル文化財創出機構, 2016, pp. 59-60)。

技術面でみると、ユーロピアナ、DPLE、ジャパンサーチとも共通点がみられる。例えば、ユーロピアナのレコードAPIのページでは、サポートしているアウトプット・フォーマットとして、JSON、JSON-LD、RDF/XMLの三種類を挙げている(Europeana, n.a.-b)。同様にDPLAでもAPIで利用する構造化データのフォーマットとしてJSON-LDを採用している(DPLA, n.a.-b)。ジャパンサーチの利活用スキーマでは、JSON-LD形式(.jsonまたは.jsonld)、RDF/XML形式(.rdf)、Turtle形式(.ttl)をサポートしている。このうち、JSON-LDは、2014年にW3C勧告として国際標準化が進められたデータフォーマットで、記述方法の規定が緩やかで外部からも利用しやすく、分野を超えた応用の可能性があるとの指摘されている(神崎正英, 2017, p. 622)。

なお、本稿では詳細は議論しないが、デジタル著作権の取り扱いについては、ユーロピアナ、DPLA、ジャパンサーチとも、収蔵コンテンツの著作権の状態に応じて、可能な場合にはクリエイ

イティブコモンズ (CC) のデジタル著作権を利用しており、収録デジタルコンテンツの二次利用も可能にしている (DPLA, n.a.-c; Europeana, 2020a; ジャパンサーチ, 2020b)。

以上のように、技術面をいくつかのポイントで比較すると、日本のジャパンサーチは、世界的に利用されている国際標準やフォーマット、デジタル著作権を採用しており、ポータルとしての技術面では先端的取組みを行っていると考えられる。

## 2 構築経緯と組織面：トップダウン型、草の根型

構築経緯と組織面の観点でみると、ユーロピアナ、DPLA、ジャパンサーチには違いがみられる。ユーロピアナはEUの枠組みで構築されてきたトップダウン型と言える。一方、DPLAは、サイトでも説明されているように、草の根型で発展してきた。なお、DPLAに米国議会図書館が参画したのは2016年となっている (DPLA, December 2, 2016)。米国議会図書館は世界最大規模の図書館に成長したが、藤野 (1998) は、その設立経緯のなかからその性質は利用者のための公共図書館の色彩が強いと指摘している。藤野 (1998) は、図書館の位置づけを、①立法府の参考図書館、②議員＝利用者のための「公共図書館」、③歴史・文化の保存所としての国家の図書館に分類し、米国は②に該当するとしている (pp.9-10.)。この分類に従えば、Europeanaは、③の延長線にあると言える。

一方、ジャパンサーチを見ると、政府文書や既存研究では「日本版ユーロピアナ」の文言が散見されるが、構築の初期段階では、草の根的活動も行われたと考える。デジタル文化財創出機構 (2016) は、ジャパンサーチの構築・公開以前の様々な動きをまとめているが、2010年代に官民ともに個別のデジタル化やデータベースの整備を進めているなか、ポータル構築の機運やデジタル文化財有効活用のための取り組みが広がっていったことがわかる。その後、閣議決定された「未来投資戦略2017」を契機にトップダウンでジャパンサーチの構築が進展した。その意味では、日本のジャパンサーチは、草の根とトップダウンのハイブリッド型と言える。

予算面では、ユーロピアナはEUからの拠出金、DPLAは財団からの拠出金を主な収入源としている。本稿の調査の範囲では、ジャパンサーチは、国立国会図書館の予算の枠組みのなかで構築・運営費が出されている。予算規模としては、ユーロピアナは日本円にして年間6億～7億円規模、DPLAは年間2億～3億程度といえる。国立国会図書館の予算のうちデジタルアーカイブの構築には1億5,000万から2億円程度、利活用の促進には5,000万円から7,000万円が配分されている。ただし、2021年度の国立国会図書館の情報システム費の予算要求額は、31億3,542万円であり、デジタル・アーカイブ・システムの運用・更新に必要な経費が約10億円計上されていることから、直接的に言及されている以外の設備投資も行われていると考える。

## 3 収録コンテンツの分類面：項目・分類、フォーマット

本稿で比較した三つのアーカイブ・ポータルは、収録コンテンツの分類にも特徴が見られる。ユーロピアナでは、文化遺産の継承を重視しているが、II 3. (1) 項で確認したように、考古学

やアートといった分類がある一方で、ファッション、写真、スポーツ等の近代・現代的な項目も大項目となっている。また、DPLAは、Ⅲ 3. (1) や (2) 項で紹介したように、南北戦争や公民権運動といった米国の文化的特徴が読み取れる項目立てとなっており、野球や移民、フードといったユーロピアナでは見られない項目がある。一方、ジャパンサーチでは、日本の文化・コンテンツに焦点が当てられており、エンタメコンテンツ、刀剣、南画といったテーマが切り出されている。ただし、エンタメコンテンツについては、アニメとゲームのデータベースと接続されているものの、コンテンツや作品のデータではなく、クリエイターやアーティストの情報が収録されている。

トップページからアクセス可能なコンテンツにも違いが見られる。文化遺産継承を全面に押し出すユーロピアナのリニューアル後のポータルでは、収蔵件数は多いもののトップページからアクセス可能なページでは、コレクションやストーリーといったキュレーションが行き届いたコンテンツが用意されており、利用者が関心をもったテーマやトピック別にアクセスしやすいように配慮されている。なお、リニューアル前のクラシック・サイトでは、サーチ・ボックスが中央に配置され、より検索重視の構成となっている。このクラシック・サイト型の構成は、DPLA とジャパンサーチも採用している。

収蔵コンテンツのメディア別でみると、ユーロピアナではイメージが占める割合が最も大きく、デジタル文化遺産の継承の観点からもイメージ重視であるといえる。DPLA はテキストの割合が大きく、教育コンテンツやeブックに力をいれている特徴がある。ジャパンサーチはユーロピアナ的なテーマ別分類も志向しつつ、現時点ではテキストや目録情報が多い。

なお、日本のポップカルチャーにかかるデータベースとしては、文化庁が2021年10月にベータ版の運用を開始した「メディア芸術データベース」がある(文化庁, 2021c)。同データベースには、文化庁が2010年以降に行ってきた調査等を元に、マンガ・アニメーション・ゲーム・メディアアートの作品情報や所蔵情報を収録している(図7)。表1で示したように、文化庁もジャパンサーチの連携機関として名を連ねており、メディア芸術データベースが、ジャパンサーチと連携していくかが注目される。加えて、同データベースには、大学等が収集してきたデータが活用されており、マンガ分野の協力所蔵館として明治大学 米沢嘉博記念図書館もデータを提供している(文化庁, 2021a)。メディア芸術データベースもOpen APIを公開しており、検索結果は、JSON形式で返される仕様となっている(文化庁, 2021b)。収蔵データ数は、総計694,727件(マンガ:495,506件、アニメーション:136,768、ゲーム:48,118、メディアアート:14,335)と充実している(2021年10月現在、トップページの検索ボックスを空のまま検索ボタンを押すと件数表示)。



図7 メディア芸術データベースのトップページ  
出所：メディア芸術データベース (<https://mediaarts-db.bunka.go.jp/>)

#### 4 コンテンツの利用面：ユーザーコミュニティの構築

アーカイブ・ポータルでは、既存のデータベースを集約しワンストップで検索できるという利便性向上のみならず、利活用による新たな価値創出が重要になっている。ユーロピアナとDPLAでは、それぞれProサイトも運用しており、ユーザーの利活用や価値創出、連携を支援する情報、サービス、ツールが提供されている。特に、ユーロピアナでは、デジタル文化遺産を活用するキャパシティブUILDINGや、ツールやサービスの開発に力を入れていることがうかがわれる。

一方、後発のジャパンサーチでは、連携機能数を増やすことや利用者向けサービスの拡充は今後の課題として位置づけられている。デジタルアーカイブジャパン推進委員会・実務者検討委員会(2020)では、ジャパンサーチの公開に合わせて、これまでの取り組みと今後の課題を報告している。同総括報告書によると、今後の課題として、デジタルコンテンツの拡充、利活用のための基盤整備、連携促進、つなぎ役の支援、人材育成支援が言及されている。

なお、2017年には、デジタルアーカイブ学会 (<http://digitalarchivejapan.org/>) が発足しており、学術・研究面のさらなる発展に向けて取組まれている。同学会の大会や学会誌では、ジャパンサーチの利活用の実践的な取り組み事例も報告されている(大井・渡邊, 2021)。

## VI おわりに ジャパンサーチにおける利活用の展望

本論では、ジャパンサーチの特徴について、先行する欧米のアーカイブ・ポータルであるユーロピアナとDPLAとの比較から考察した。その際、二次資料や各ポータルのトップページやナビゲーション、構築の経緯やミッション、活動内容やポータルの構成といった情報を手がかりに、技術面、組織面、分類面、利用面の四つの観点からそれぞれの特徴は次のようにまとめられる。

- ・ 技術面では、ジャパンサーチは、ユーロピアナや DPLA と共通する国際標準を採用し、API も公開しており、海外のアーカイブ・ポータルに近い利便性を提供するものと言える。参加機関向けやユーザー向けのツールの開発と提供は、10年以上先行するユーロピアナに一日の長がある。
- ・ 組織面では、トップダウン型で構築されたユーロピアナ、草の根型でネットワークが拡大した DPLA と比較して、ジャパンサーチは検討段階では草の根で必要性の認識が拡大し、構築段階ではトップダウンの色彩が濃いハイブリッド型と考える。
- ・ 分類面では、ジャパンサーチはテーマ別のコンテンツも用意しつつ、トップページの中央にサーチ・ボックスを配するユーロピアナのクラシック・サイト型の構成となっている。また、メディアフォーマットはテキストが多い。
- ・ 利用面では、ユーロピアナや DPLA は、連携機関や開発向けの情報・サービス・ツールを提供する Pro サイトを提供しているが、ジャパンサーチはこうした別サイトはなく、連携機関やユーザーの拡大を図ることが課題といえる。

今後の方向性としては、①ユーロピアナのようにキュレーションされたコンテンツを増やし文化遺産にかかる啓発を推進する、② DPLA のように教育コンテンツを拡充する、③ジャパンサーチの独自の方向性を進む、が想定される。この点について、デジタルアーカイブジャパン推進委員会・実務者検討委員会 (2020) は、「我が国におけるデジタルアーカイブの構築・共有と活用の推進は、文化の保存・継承・発展だけでなく、コンテンツの二次的な利用や国内外への情報発信の基盤となる取組である」(p.3)としており、利活用に重きがおかれている。その際、教育、学術・研究、観光、地域活性化、防災、ヘルスケア、ビジネス等の様々な分野での利活用により、社会的、文化的、経済的発展につなげていくことを重視している。2021年には、2021年から2025年の期間を対象としたジャパンサーチ戦略方針の策定が進められており、国立国会図書館電子情報部からは、ミッション、目標、KPI、コレクション・ポリシーを盛り込む必要性が指摘されている(国立国会図書館電子情報部, 2021)。こうした総括や戦略の検討からは、日本は上述の③の方向にあると考える。

遡って2017年に示された「我が国におけるデジタルアーカイブ推進の方向性」を確認すると、デジタルアーカイブを文化的遺産の継承と、その対外的な発信の基盤として位置づけられている(デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会, 2017, p. 3)。また、ユーザーが単にアクセスするだけでなくアプリ開発や地域情報発信に活用することで、新たな価値を発言できるようになるとしている (Ibid. p.4)。こうした点から、ジャパンサーチという日本のアーカイブ・ポータルは、欧米のミッションやビジョンにはないユニークな視点を含んでいると言える。その一つが、観光や地域情報発信、防災等の個別分野での利活用を想定している点である。実際、IV 3 (3) で取り上げたように、ジャパンサーチの地域情報発信での活用事例が報告されている。

ジャパンサーチの利活用を促進するためには、ユーザーコミュニティの構築の重要性が増して

いると考える。ユーロピアナやDPLAのミッションや目標は、日本とは異なる部分があるものの、アーカイブ・ポータルの活用を推進するためには、ユーザー数とスキルやリテラシーの両面を上げていく必要がある。その際、Proサイトのような情報やサービス、ツールを集約し、共有するプラットフォームが果たす役割は大きいと考える。ユーロピアナでは、キュレーションされたテーマ別コンテンツを、DPLAでは教育向けのコンテンツの拡充が進められているが、ジャパンサーチでは、「テーマ別検索」のページではキーワード検索がうながされるに留まり、ユーザーがより知りたい、調べたいと感じる手がかりや導き手となるコンテンツが乏しい。アーカイブ・ポータルの利活用を推進するためには、一般ユーザーとアーカイブの架け橋となるキュレーションされたコンテンツの役割が高まると考えられる。Proサイトで提供されるツールやサービス、ガイダンスに加えて、アーカイブ・ポータル活用の事例や方法、キュレーションされたコンテンツの増加を図っていくことで、利活用の拡大につながっていくと考えられる。ジャパンサーチは本稿執筆時点で発足して1年あまりであり、今後の発展によって日本の知的財産の活用と拡大に寄与することが期待される。逆説的ではあるが、ジャパンサーチのユニークな発展に向け、ユーロピアナやDPLAで提供されているテーマ別コンテンツや教育コンテンツやその制作ノウハウから学ぶべき点は残されていると考える。

以上のように、本稿では、ジャパンサーチの特徴と展望について、先行するポータルとの比較から考察を行った。その際、現代的なアーカイブ・ポータルの要件をまとめただけで、比較分析の視点として、技術面、組織面、分類面、利用面の四つの分析軸を設定した点に本稿のオリジナリティがあると考えられる。これまでの先行研究では、先行するユーロピアナやDPLAについて個別の紹介や部分的な比較や参照が行われているが、本稿のように、構築の経緯や活動内容、ポータル・サイトの特徴について項目を揃えて比較を行ったものは管見の範囲では見られなかった。また、一次情報として政策文書や各ポータル、また、その公開文書を参照しつつ、今後のジャパンサーチに求められる要件について考察を行った点も一定の意義があると考えられる。

ただし、本稿には、テーマ別コンテンツや教育コンテンツやその制作ノウハウについてはさらなる研究の余地が残されている。加えて、ユーロピアナやDPLAにおける連携方法や組織ごとの役割分担やそのための資源（人員や資金）の配分についても詳細には踏み込んでいない。いうまでもなく、ユーロピアナのリニューアルに見られるように、各国・地域のポータルは発展を続けており、継続的な研究が必要である。また、本論では、比較分析対象のアーカイブ・ポータルとして、ユーロピアナとDPLAのみを取り上げているが、アーカイブ・ポータルの構築はアジア圏の中国や韓国でも取組まれている（デジタル文化財創出機構, 2016, pp. 56-57）。そのため、ユーロピアナやDPLAにかぎらず、ジャパンサーチに参考になる取り組みは他にもあると想定される。こうした点も今後の研究課題として残されている。

## 謝辞

本稿は、国際共同研究を行っている共著者の大江が、英国のボーンマス大学において開催した Eventbrite seminar 「Digital archiving by museums and libraries: Japan, the EU, and the UK」(2021年6月18日)での討議を発展させ、とりまとめたものである。セミナーにおいて討論者としてコメントをくださった Mr. Benjamin White (Centre for Intellectual Property Policy and Management, BU, Chair of Legal Working Group of Ligue des Bibliothèques Européennes de Recherche) 及び Mr. Neil Fitzgerald (Head of Digital Research, British Library) に、特に記して謝辞とする。

## 参考文献・資料

- DPLA. (2013). *Digital Public Library of America (DPLA) Launches Today*. DPLA. <https://dp.la/news/digital-public-library-of-america-dpla-launches-today/>
- DPLA. (2019). *Digital Public Library of America Strategic Roadmap 2019-2022 - Collaborating for Equitable Access to Knowledge for All*. <https://dpla.wpengine.com/wp-content/uploads/2019/06/DPLA-Strategy-2019-2022.pdf>
- DPLA. (2021). *2019 Form 990 for Digital Public Library of America*. [https://www.causeiq.com/organizations/view\\_990/461160948/d65334c76cac3978f28d33d680a3f1fb](https://www.causeiq.com/organizations/view_990/461160948/d65334c76cac3978f28d33d680a3f1fb)
- DPLA. (December 2, 2016). *Mapping the Future of DPLA and Library of Congress Collaboration* <https://dp.la/news/mapping-the-future-of-dpla-and-library-of-congress-collaboration>
- DPLA. (n.a.-a). *About Us*. Retrieved October, 17 from <https://dp.la/about>
- DPLA. (n.a.-b). *API Codex*. Retrieved October, 17 from <https://pro.dp.la/developers/api-codex>
- DPLA. (n.a.-c). *Copyright Status Filter*. Retrieved October, 17 from <https://dp.la/about/rights-categories>
- DPLA. (n.a.-d). *Funders*. Retrieved October from <https://dp.la/about/funders>
- DPLA. (n.a.-e). *History*. DPLA. Retrieved October, 17 from <https://pro.dp.la/about-dpla-pro/history>
- DPLA. (n.a.-f). *Open Bookshelf*. Retrieved October, 17 from <https://ebooks.dp.la/open-bookshelf/>
- DPLA. (n.a.-g). *Our Hubs*. <https://pro.dp.la/hubs/our-hubs>
- Europeana Foundation. (2018). *Europeana Foundation Annual Report and Accounts 2017*. <https://pro.europeana.eu/post/annual-report-and-accounts-2017-answering-the-call-to-culture>
- Europeana Foundation. (2021). *ANNUAL REPORT & ACCOUNTS 2020*.
- Europeana. (2010). *Background*. European Union. [https://web.archive.org/web/20101225175701/http://europeana.eu/portal/aboutus\\_background.html](https://web.archive.org/web/20101225175701/http://europeana.eu/portal/aboutus_background.html)
- Europeana. (2020a). *Creative Commons for cultural heritage*. <https://pro.europeana.eu/event/creative-commons-for-cultural-heritage>

- Europeana. (2020b). *Europeana Strategy 2020-2025 - Empowering Digital Change*. [https://pro.europeana.eu/files/Europeana\\_Professional/Publications/EU2020StrategyDigital\\_May2020.pdf](https://pro.europeana.eu/files/Europeana_Professional/Publications/EU2020StrategyDigital_May2020.pdf)
- Europeana. (n.a.-a). *BRIEF HISTORY*. Retrieved October, 17 from <https://pro.europeana.eu/about-us/mission#brief-history>
- Europeana. (n.a.-b). *EUROPEANA RECORD API*. Retrieved October, 17 from <https://pro.europeana.eu/page/record>
- Europeana. (n.a.-c). *MISSION*. Retrieved October, 17 from <https://pro.europeana.eu/about-us/mission>
- Europeana Pro. (n.a.). *SERVICES AND TOOLS*. Retrieved October, 17 from <https://pro.europeana.eu/about-us/services-and-tools>
- McKemmish, S., Piggott, M., Reed, B., & Upward, F. (2019). アーカイブズ論：記録のちからと現代社会（翻訳：安藤正人，石原一則，坂口貴弘，塚田治郎，保坂裕興，森本祥子）. 明石書店. <https://ci.nii.ac.jp/ncid/BB2952112X>
- Reflection Group on Bringing Europe's Cultural Heritage Online. (2011). *The new renaissance - Report of the Comité des Sages on bringing Europe's cultural heritage online*. <https://op.europa.eu/en/publication-detail/-/publication/79a38a23-e7d9-4452-b9b0-1f84502e68c5>
- Rubin, R. (2014). *図書館情報学概論*（根本彰，Trans.）. 東京大学出版会.
- 赤間亮. (2019). A5-1 文化・芸術と文化情報学. In 村上征勝 等 (Ed.), *文化情報学事典*. 勉誠出版.
- 大井将生, 渡邊英徳. (2021). [42] ジャパンサーチのワークスペース機能を活用した協働キュレーション授業：「問い」と資料を接続するデジタルアーカイブの活用法. *デジタルアーカイブ学会誌*, 5 (s1), s67-s70. [https://doi.org/10.24506/jsda.5.s1\\_s67](https://doi.org/10.24506/jsda.5.s1_s67)
- 大向一輝. (2020). ジャパンサーチの経緯と文脈. *デジタルアーカイブ学会誌*, 4(4), 329-332. [https://doi.org/10.24506/jsda.4.4\\_329](https://doi.org/10.24506/jsda.4.4_329)
- 閣議決定. (2017). *未来投資戦略 2017—Society 5.0 の実現に向けた改革*. [https://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/miraitousi2017\\_t.pdf](https://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/miraitousi2017_t.pdf)
- 神崎正英. (2017). リンクの機能を柔軟に生かすデータのウェブ. *情報の科学と技術*, 67(12), 622-627. [https://doi.org/10.18919/jkg.67.12\\_622](https://doi.org/10.18919/jkg.67.12_622)
- 神崎正英. (2020). ジャパンサーチ利活用スキーマの設計と応用. *デジタルアーカイブ学会誌*, 4(4), 342-347. [https://doi.org/10.24506/jsda.4.4\\_342](https://doi.org/10.24506/jsda.4.4_342)
- 久古聡美, 吉田暁, 中山正樹. (2006). 国立国会図書館デジタルアーカイブ・ポータルの現状と今後—Web2.0 世代のデジタルアーカイブ・ポータルの提供を目指して—. *情報管理*, 49(6), 313-323.
- 国立国会図書館. (2019). *平成31年度国立国会図書館予算（案）の概要*. [https://www.ndl.go.jp/jp/aboutus/outline/h31\\_outline\\_budget.pdf](https://www.ndl.go.jp/jp/aboutus/outline/h31_outline_budget.pdf)

- 国立国会図書館. (2021). 令和3年度国立国会図書館予算(案)の概要. [https://www.ndl.go.jp/jp/aboutus/outline/r03\\_outline\\_budget.pdf](https://www.ndl.go.jp/jp/aboutus/outline/r03_outline_budget.pdf)
- 国立国会図書館電子情報部. (2021). ジャパンサーチの戦略方針(2021-2025)について. [https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive\\_suisiniinkai/jpsearchwg/dai3/siryoul.pdf](https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/jpsearchwg/dai3/siryoul.pdf)
- ジャパンサーチ. (2020a). ジャパンサーチの概要. デジタルアーカイブジャパン推進委員会・実務者検討委員会. <https://jpsearch.go.jp/about>
- ジャパンサーチ. (2020b). デジタルコンテンツの二次利用条件表示について. <https://jpsearch.go.jp/policy/available-rights-statements>
- ジャパンサーチ. (2020c). 連携をご希望の機関の方へ. デジタルアーカイブジャパン推進委員会・実務者検討委員会. <https://jpsearch.go.jp/cooperation>
- ジャパンサーチ. (2021a). アーカイブのページ. デジタルアーカイブジャパン推進委員会・実務者検討委員会. <https://jpsearch.go.jp/archive>
- ジャパンサーチ. (2021b). ジャパンサーチの活動報告(2020年度). デジタルアーカイブジャパン推進委員会・実務者検討委員会. <https://jpsearch.go.jp/arfy2020#cehznblajeag5>
- ジャパンサーチ. (n.a.-a). 開発者向け情報. Retrieved October, 17 from <https://jpsearch.go.jp/static/developer/ja.html>
- ジャパンサーチ. (n.a.-b). 現在のデータ. デジタルアーカイブジャパン推進委員会・実務者検討委員会. Retrieved October, 17 from <https://jpsearch.go.jp/stats>
- ジャパンサーチ. (n.a.-c). 連携機関. デジタルアーカイブジャパン推進委員会・実務者検討委員会. Retrieved October, 17 from <https://jpsearch.go.jp/organization?from=0>
- 首相官邸. (n.a.). デジタルアーカイブジャパン推進委員会及び実務者検討委員会. [https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive\\_suisiniinkai/index.html](https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/index.html)
- 田良島哲. (2020). 行かない/行けない人のためのデジタル・ミュージアムと、それを支えるデジタルアーカイブ. アートスケープ(2020年07月01日号).
- デジタルアーカイブジャパン推進委員会・実務者検討委員会. (2020). 3か年総括報告書-我が国が目指すデジタルアーカイブ社会の実現に向けて [https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive\\_suisiniinkai/pdf/r0208\\_3kanen\\_houkoku\\_honbun.pdf](https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_suisiniinkai/pdf/r0208_3kanen_houkoku_honbun.pdf)
- デジタルアーカイブの連携に関する関係省庁等連絡会・実務者協議会. (2017). 我が国におけるデジタルアーカイブ推進の方向性. [https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive\\_kyougikai/houkokusho.pdf](https://www.kantei.go.jp/jp/singi/titeki2/digitalarchive_kyougikai/houkokusho.pdf)
- デジタル文化財創出機構. (2016). デジタル文化革命!: 日本を再生する"文化力". 東京書籍. <https://ci.nii.ac.jp/ncid/BB21087631>
- 時実象一. (2015). DPLAfest 2015. 情報管理, 58(4), 313-318. <https://doi.org/10.1241/johokanri.58.313>
- 中山正樹. (2004). 国のデジタルアーカイブ・ポータル構築: 国立国会図書館「電子図書館中期計画2004」の実施に向けて(&lt;特集&gt; デジタル情報資源のアーカイビング). 情報の科学

と技術, 54(9), 453-460. [https://doi.org/10.18919/jkg.54.9\\_453](https://doi.org/10.18919/jkg.54.9_453)

藤野幸雄. (1998). アメリカ議会図書館：世界最大の情報センター (Vol. 1428). 中央公論社.  
<https://ci.nii.ac.jp/ncid/BA36819033>

文化庁. (2021a). マンガの協力所蔵館など. <https://mediaarts-db.bunka.go.jp/library#anc04>

文化庁. (2021b). メディア芸術データベース (ベータ版) WebAPI 仕様. [https://mediaarts-db.bunka.go.jp/resources/pdf/mediaartsdb\\_webapi\\_documents.pdf](https://mediaarts-db.bunka.go.jp/resources/pdf/mediaartsdb_webapi_documents.pdf)

文化庁. (2021c). メディア芸術データベース (ベータ版) について. <https://mediaarts-db.bunka.go.jp/about>

文化庁. (2021d). 国立美術館・博物館・劇場の開館状況等について. [https://www.bunka.go.jp/koho\\_hodo\\_oshirase/sonota\\_oshirase/covid19\\_info/kyukan.html](https://www.bunka.go.jp/koho_hodo_oshirase/sonota_oshirase/covid19_info/kyukan.html)

\* 各 URL の最終アクセス日は 2021 年 10 月である。



【研究ノート】

## 上級日本語科目改革のための調査に関する報告

安 高 紀 子

ATAKA, Noriko

黄 叢 叢

HUANG, Congcong

黄 秀 智

HWANG, Sooji

吴 梅

WU, Mei

### I はじめに

現在、本学部では1年次必修の初年次教育科目に関して、新カリキュラムでの開講に向け、検討が進められている。これに伴い、初年次教育科目を履修する学部1年生の留学生に対して行われている上級日本語科目においても、初年次教育の新しいシラバスと照らし合わせ、見直しを行うとともに、これまでの上級日本語科目の授業実践を振り返り、改善を行うこととなった。大学での学部留学生に対する日本語教育は、アカデミック・ジャパニーズと呼ばれ、日本留学試験<sup>1</sup>導入にともなって用いられるようになった表現（門倉 2003:123）である。日本留学試験では、「日本の大学での勉強に対応できる日本語力（アカデミック・ジャパニーズ）」と定義されているものの、概念が不明瞭である（門倉, 2000; 三宅, 2003; 森, 2005）との指摘もあり、門倉（2003）はアカデミック・ジャパニーズという教育領域を「シラバス」として明示することは困難だとするも、日本語教育としての議論の必要性を示唆している。

そこで、本学部の留学生に対するアカデミック・ジャパニーズとしての日本語教育において、何が必要であるのかを改めて検討するために、これまでの上級日本語科目の履修者に対して、アンケート調査、およびヒアリング調査を実施した。本稿では、そのアンケート調査とヒアリング調査で得られた履修者からの意見、また、調査の結果、明らかになった点について報告する。

---

1 日本留学試験（EJU）は、2002年より独立行政法人日本学生支援機構によって実施されている試験で、主催である独立行政法人日本語学生試験機構のHPによると、「外国人留学生として、日本の大学（学部）等に入学を希望する者について、日本の大学等で必要とする日本語力及び基礎学力の評価を行うことを目的に実施する試験」である。https://www.jasso.go.jp/ryugaku/eju/about/about\_a/about\_eju.html（2021年10月17日アクセス）

## Ⅱ 上級日本語科目の概要

明治大学国際日本学部は、在籍する学生のおよそ2割を留学生が占めており、本学部に在籍する留学生に対しては、それぞれの日本語能力に合わせ、「入門」から「上級」まで全7レベルでの日本語教育が行われている。本稿の対象である上級日本語科目は、7レベルの中で最も高いレベルに位置づけられ、日本語能力試験のN1レベル相当である。履修者の大半は本学部留学生1年生で、留学生が大学で勉強するための日本語や一般的な教養としての日本語を学ぶための科目として、「読む」「聴く」「話す」「書く」の技能別に教育が行われている。授業科目としては、「読解」「聴解」「口頭表現」「文章表現」の4科目あり、春学期に「上級日本語Ⅰ」、秋学期に「上級日本語Ⅱ」として開講され、各科目1単位、春学期・秋学期合わせて合計8単位が付与される。本学部の正規学部留学生（English Trackの学生を除く）に対しては、外国語必修科目となっている。

## Ⅲ 本調査の目的

上級日本語科目の改革を進めていくための検討材料として、これまで行われてきた上級日本語科目の授業について、履修者である留学生がどのように考えているのか、また、本学部での日本語教育に対して、どのような要望があるのかを把握しておく必要がある。そこで、本調査ではこれまでに上級日本語を履修した留学生からの意見や要望を収集する目的で、アンケート調査、およびヒアリング調査を実施する。

## Ⅳ 調査方法

本調査は2020年度以前に上級日本語科目を履修した留学生（調査時に2～4年次に在籍する留学生）を対象とし、アンケート調査、およびヒアリング調査を実施した。まず、アンケート調査を実施した。次に、アンケート調査の回答者のうち、「ヒアリングへの協力が可能」との回答があった留学生を対象に、ヒアリング調査を実施した。Ⅴ章ではアンケート調査について、次に、Ⅵ章ではヒアリング調査について報告する。

## Ⅴ アンケート調査

### 1. アンケート調査の概要

アンケート調査<sup>2</sup>は本学部に在籍する留学生で、2020年度までに上級日本語科目を履修した留学生を対象に行った。表1にアンケート調査の概要を示す。

---

2 アンケート調査で使ったアンケートフォームについては、本稿の最後に参考資料として示す。

表 1：アンケート調査の概要

対象者	国際日本学部に在籍する留学生（2年生～4年生）
実施期間	2021年6月7日～18日（12日間）
実施方法	Microsoft Forms を利用したウェブアンケート アンケートの URL をメールで全対象者に送付
質問項目数	15 項目

アンケートでは、上級日本語科目の改革を行う上で、履修した留学生の意見を踏まえる必要がある点について、全 15 項目の質問を作成した。質問内容は、授業の週当たりのコマ数と科目構成（3 項目）、現行の授業内容に関する評価（4 項目）、上級日本語科目の履修で身に付けることができた知識やスキル（1 項目）、上級日本語科目に対する要望（1 項目）、授業の履修に対する不安（2 項目）、レポートや発表などの支援の必要性（2 項目）、外国語必修科目としての英語科目との比較（2 項目）である。質問への回答は、質問項目によって、多枝選択式と自由記述式を併用した。また、アンケートへの回答は無記名とした。

## 2. アンケート調査の結果と考察

アンケートへの回答数は 21 件であった。回答者の学年別内訳は、2 年生 8 名、3 年生 9 名、4 年生 4 名である。出身地域別では、韓国の 12 名が最も多く、中国 5 名、香港 2 名、台湾 1 名、その他 1 名である。

以下では、質問項目ごとに結果、また学生のコメントを記す。

### （1）週当たりの授業回数

週当たりの授業回数について、「現在、週 4 回の授業が行われていますが、この回数について、週に何回が適当だと思いますか」という質問を行った。回答形式は 0 回（不要）～6 回以上から該当するものを選ぶ多枝選択式である。

回答で最も多かったのは、4 回で 48% (10) であった。以下、2 回が 29% (6)、3 回が 14% (3) であった。少数ではあるが、現行以上の 5 回を希望する者が 5% (1)、一方で、0 回が 5% (1)、つまり必修科目として不要だとする回答もあった。この結果から、現行の週 4 コマの授業回数に対して、回答した留学生の半数は適当だと捉えていることがわかった。

### （2）科目構成

現行の科目構成について、履修者はどのように考えているのかを知るため、「現在は、読解・聴解・口頭表現・文章表現の 4 科目の構成ですが、この科目構成について、どう思いますか」という質問を行った。回答は多枝選択式である。

回答は、「ちょうどよい」は 67% (14)、「科目が多すぎる」は 33% (7) という結果であった。この結果から、およそ 7 割は現行の科目構成で良いと考えていることがわかった。

さらに、「科目が多すぎる」と回答した者に対して、「どのような科目構成がよいと思いますか。科目名を具体的に教えてください」という質問を行った。その結果、「聴解と口頭表現」と「読解と文章表現」の2科目構成を希望する回答が6件、「読解と聴解」「文章表現」「口頭表現」の3科目構成を希望する回答が1件であった。

### (3) 現行の上級日本語の授業に対する評価

上級日本語の授業を履修し、留学生はどのように考えていたのかを知るため、「授業で、良かった点、悪かった点はどんなことですか。具体的に教えてください」という質問を行った。回答は自由記述式とした。

以下では、それぞれの科目別に、回答のコメント（原文のまま）を記す。

#### ①「読解」

- ・ 新聞記事や小説など色んな種類の資料が読めて良かったです。
- ・ 春学期の最後にやったビブリオバトルが印象的でした。
- ・ 資料を通して、日本の歴史や有名な人物などのことが学べること。
- ・ 新聞、小説、論文など様々な文章を読みながら新たな単語を学習することができました。
- ・ 文の内容が学部と特に関係がなかった。
- ・ 読解スキルをあまり身に付かなかった気がする。
- ・ 課題が少し多くて、最初読んで単語を理解するのが難しくて、きつかったです。

評価されている点として、新聞や小説などの多様なタイプの文章を教材として扱っており、それら読むことで新しい語彙の学習ができたとするコメントが多く見られた。教材とした文章のトピックについては、日本の歴史などに関する新しい知識が習得できたという肯定的な意見が見られる一方で、本学部の各領域との関連性という視点からの教材選択の必要性に関する意見もあった。また、文章をどう読むかという読解スキルの育成に対する要望もあった。その他には、文章の難易度、課題の負担といった点を問題点として指摘するコメントも見られた。

#### ②「聴解」

- ・ 多様な媒体からの映像を見て、問題を解くことで、自分の聴解力を確認するいい機会になったと思いました。
- ・ 色々な番組や、ドキュメンタリー、ドラマなどを見て、普通の日本人が話す言葉や言葉の速さを勉強できる。
- ・ 就活関連、ニュース、歌舞伎、ドラマなど色んな分野の単語が学べました。
- ・ 単語や聞き取りのポイントなど適当な速さで説明して下さっていいと思います。
- ・ 「読解」と同様、内容が特に学科の内容と関係がなかった。
- ・ 授業の目的と受講者が求めていることが一致していないこと。

授業では、ドキュメンタリー、ドラマ、ニュースなど多様なジャンルの視聴覚教材に触れることができ、ナレーションや会話など様々な形態の発話に触れることができたこと、また内容に関する新しい語彙が学べたことなどが評価されている。一方で、素材とした視聴覚教材の内容に関して、「読解」と同様に、本学部の各領域との関連性という点で指摘があった。

### ③「口頭表現」

- ・ 敬語の授業が一番役に立ったと思う。
- ・ 言葉遣いや日本語独特の言い回しなどを学べてよかったと思います。
- ・ 授業が、実際に活用できる内容で構成されていた点。具体的に、他の授業で行われるであろう、発表に必要な準備と練習の方法を学ぶことができたこと。
- ・ 発表などの時、先生から一人ずつ個人フィードバックのかみがもらえて自分の日本語発信力のどの部分が良くてできたのか、どんなところが弱いかわかって良かったです。
- ・ グループ活動が多く、クラスの友達と話す機会が多くて、とてもよかったと感じました。
- ・ 人数が多かったので時間が足りなかった。

まず、敬語や日本語の言い回しなどの言語表現に関する学習、また、プレゼンテーションやディスカッションなど、実際の授業活動において求められるスキルの運用練習などが評価されている。また、授業内のグループ活動に対して肯定的な評価が見られた。その他、授業運営上の問題点として、1クラス当たりの人数が多いため、個々の発表に十分な時間が取れないという点が問題点として挙げられていた。

### ④「文章表現」

- ・ 序論、本論、結論、引用、参考文献、資料の正しい作成方法を詳しく学びました。
- ・ 一つのレポートを時間をかけて完成することで、しっかり練習ができた点。
- ・ 体系的にレポートの書き方について学ぶことができてよかったと思いました。特に、引用の仕方を学び、他の科目でのレポートにも有用に活用しているので、1年次の時に学んでおいてよかったと思いました。
- ・ この授業で学んだレポートの書き方を他の授業でよく活用することができました。
- ・ ペア活動で、お互いに文章をチェックすることでさらに日本語を学べること。
- ・ 課題が少し負担でした。

レポートの構成や引用の仕方などのレポート作成方法に必要な知識、また、レポートのテーマ設定から完成まで、作成のプロセスを体験できたことを評価するコメントが見られた。そして、レポート作成のスキルを他の授業で活用できたというコメントも多かった。一方で、レポート作成課題に対する負担感や、困難さを訴えるコメントもあった。

## (4) 上級日本語科目の履修で身に付けることができた知識やスキル

上級日本語科目を履修することで、履修者はどのような知識やスキルを身に付けることができたと考えているのかを把握するために、「上級日本語を履修して、自分ができるようになったと思うことはどんなことですか。具体的に教えてください」という質問を行った。回答は自由記述式で、以下のようなコメントが見られた。

- ・大学のレポートを書けるようになりました。
- ・レポートを書くスキルが身につけられました。より高いレベルの日本語が発信できるようになりました。
- ・日本語の本を読むことがより好きになったこと。正しいフォーマットでレポートを書けること。ドラマや番組を見るとき、聞き取れない単語数が減ったこと。
- ・新しくできるようになったというより、私がおもっていた日本語を磨くことができる機会になったと思う。
- ・新しく知ったものはなく、学んだものを復習している感じがした。
- ・留学生たちとコミュニケーションができるようになりました。
- ・顔も知らない、互いのことを知らない相手に先に話しかける勇気を得たこと。

上記の通り、「レポートが書けるようになった」「プレゼンテーションの方法」「より高いレベルでの日本語での発信」といったコメントが多数見られた。レポートの作成や発表といったアウトプットを実際に行うことで、自分自身の成果が実感でき、スキルを習得できたと感じていることがわかる。その他には、全体的な日本語能力の向上を挙げるものもいた。また、「留学生とのコミュニケーション力」「知らない相手に話しかける勇気」といったコメントについては、授業の中で、ペアやグループで意見交換や共有を行うなど、クラスメイトとの協働学習を通して培われたものと思われる。

## (5) 上級日本語科目に対する要望

本科目は大学1年次に履修する科目であり、留学生がどのようなことを学びたいと考えているのか、彼らのニーズを知るために、「上級日本語の授業で、こんなことを学びたかったということがありますか。それはどんなことですか」という質問を行った。回答には、以下のような要望が見られた。

- ・私生活や学生生活のサポートになる内容
- ・日常生活でよく使うフレーズ
- ・若者言葉
- ・敬語

- ・ 日本語の方言
- ・ 就職関連の日本語
- ・ ビジネスメールの書き方
- ・ 書く時のルール、様々な文書のテンプレート
- ・ 日本人と交流するチャンスが増えるとよい

上記の通り、要望としては、大学での授業活動自体に関わるものではなく、クラスメイトや友人と話すための若者言葉、日本人学生との会話で触れる方言、また、アルバイト先でお客様に対する敬語表現などで、留学生にとって、授業外の日常場面で用いられる日本語に対するニーズがあることがわかった。また、就職のための日本語やビジネス日本語など、将来に向けた準備としても日本語学習の必要性を感じているようである。その他に、授業での日本人学生との交流を望む意見もあった。本学部 English Track の留学生対象の日本語の授業では、日本人学生をゲストに招き、日本人学生に対するインタビューやアンケート調査の実施、留学生のプレゼンテーションを聞いて、コメントをもらうなど、日本人学生と留学生が交流する場を設けている。それに対し、現行の上級日本語の授業では、授業内で留学生と日本人学生が共に学ぶような活動は行っていない。留学生にとっては、日本人学生との交流の機会が少なく、イメージしていた留学生活とは異なる環境だと感じている可能性がある。

#### (6) 授業の履修に対する不安

本科目を履修する留学生の日本語学習に関する背景として、日本国内の日本語学校等を経て、本学に入学する者のほかに、母国の高等学校卒業した後、直接本学に入学するという者も多い。そのため、日本語のみでの学習経験が乏しいと思われ、授業の履修に対する不安を抱えている可能性が考えられる。そこで、「上級日本語の授業で、授業活動（プレゼンテーション、ディスカッション、ペア／グループワークなど）を日本語で行うことに不安がありましたか」という質問を行った。

回答は多枝選択式で、「不安があった」は10% (2)、「不安はなかった」は48% (10)、「どちらともいえない」は43% (9) という結果であった。この結果から、「不安はなかった」という回答がおよそ半数を占めており、授業活動に問題なく取り組んでいると思われる。しかし、「どちらともいえない」という回答もほぼ同数あり、その理由については、今回のアンケートでは知ることができておらず、今後、調査する必要があるだろう。

一方で、「不安があった」と回答した者に対しては、さらに「どんな点で不安を感じていましたか」という質問を行った。自由記述式の記述では、「人々の前で話すのが苦手」「間違いが多く、直しきれなかった」との回答が見られた。

### (7) レポートや発表などの支援の必要性

留学生にとって、レポート作成や発表を母語ではない外国語で行うことは容易ではない。そのため、どのように対応しているのかを把握するため、「レポートの提出やプレゼンテーションの前に、あなたの「日本語」について、日本人の友人など誰かにネイティブチェックをしてもらっていますか」という質問を行った。回答は多枝選択式で、「まったくしてもらっていない」が62% (13)で最も多かった。次に「ときどきしてもらっている」が33% (7)、「いつもしてもらっている」は5% (1)という結果であった。

次に、支援の必要性に関して、「レポート作成のサポートが受けられるシステムがあれば、利用したいと思いますか」という質問を行った。回答は、「ぜひ利用したい」が62% (13)で、「ときどき利用したい」が38% (8)という結果であった。

これらの結果から、留学生はレポート作成などにおいて、サポートを受けたいという希望はあるものの、本学部では学部生に対するサポートのシステムはないため、実際には、友人やその他の人からの助けも得ることができておらず、支援が必要な状況にあることがわかった。

### (8) 外国語必修科目としての英語科目との比較

上級日本語科目と同様に外国語の必修科目となっている英語科目との比較において、2つの点について質問を行った。

まず、英語科目と上級日本語科目の負担の度合いについて、「英語の必修と比べて、上級日本語の授業を履修する負担はどうでしたか」という質問を行った。回答は多枝選択式で、「英語の方が大変だった」が43% (9)、「英語も日本語も同じくらい大変だった」が29% (6)、「日本語の方が大変だった」が24% (5)、「英語も日本語もどちらもあまり大変ではなかった」が5% (1)であった。英語の方が大変だという回答が4割を超えており、上級日本語科目と比べ、英語科目の方に負担を感じていることがわかる。しかし、「英語も日本語も同じくらい大変だった」、「日本語の方が大変だった」という回答も一定数あり、反対に「日本語も英語もどちらもあまり大変でなかった」という意見が少数であることから、必修科目としての英語および日本語の履修に対して、学生は負担を感じていることが示された。

次に、クラス編成に関して、英語科目は履修者の英語レベルによって、レベル別のクラス編成が行われているが、日本語科目は日本語レベルによるクラス分けは行われていない。そこで、クラス分けの必要性という点について、「英語はレベル別に授業が行われていますが、日本語に関してもレベル別に授業を行ったほうがよいと思いますか」という質問を行った。回答は多枝選択式で、「レベル別でなくて問題ない」が52% (11)、「レベル別がよい」が48% (10)で、意見が半分に分かれた。英語科目は留学生・日本人学生のいずれにとっても1年次の外国語必修科目であるため、履修者数が非常に多い。一方、日本語科目は留学生のみが対象で、履修者数は少ないため、英語と比較するには、状況に大きな相違がある。今回のアンケートでは、多枝選択式の回答で、回答理由は不明であるため、今後は理由についても調査する必要があるだろう。

## Ⅵ ヒアリング調査

### 1. ヒアリング調査の概要

アンケート調査の回答者の中で、ヒアリングへの協力が可能と回答のあった者に対して、ヒアリングを実施した。表2にヒアリング調査の概要を示す。表2の通り、ヒアリング調査では、参加者が自分の考えを十分に述べることができるよう、留学生の各言語（中国語・韓国語）で実施した。また、参加者同士の率直な意見交換を促す目的で、ヒアリング実施時に教員は同席せず、SA（スチューデントアシスタント）がファシリテーター役を務めた。なお、ヒアリング調査を行うにあたっては、実施前にSAとのヒアリング模擬練習、ファシリテーションの方法に関するワークなど、事前に綿密な打ち合わせを行い、事前準備を行った上でSAはヒアリングを担当した。

ヒアリング調査はコロナ禍の状況を鑑み、Zoomを利用したリアルタイムオンライン形式で行った。ヒアリング調査中は参加者に同意を得て、ビデオ録画をした。ヒアリングの内容については、ファシリテーターのSAが要点をまとめた記録を作成した。さらに、韓国語・中国語それぞれの母語話者である筆者らが日本語翻訳を行った。

表2：ヒアリング調査の概要

	韓国語母語話者	中国語母語話者
実施日	2021年7月16日	2021年7月14日
実施方法	オンライン（Zoom）	オンライン（Zoom）
参加者	5名（4年生1名、3年生4名）	6名（3年生5名、2年生1名）
実施時間	約40分	約45分

### 2. ヒアリング調査の結果

ヒアリングでは、アンケート調査の結果を踏まえ、履修者からさらに意見を抽出する必要があると思われる点をトピックとして立て、意見交換を行った。ヒアリングでのトピックは、「週当たりの授業回数と科目構成」、「授業に対する評価」、「大学1年次に身に付けておくべきスキル」、「上級日本語に対する要望」の4つである。各トピックに関して、言語別にヒアリングで挙げられたコメントの要点をまとめたものを表3に示す。ヒアリングで参加者から挙げられた具体的な意見については、以下でトピックごとに報告する。

表3：各トピックに関する言語別のコメント

トピック1：週当たりの授業回数と科目構成	
【韓国語】 ・履修の負担 ・科目構成の見直し	【中国語】 ・履修の負担 ・履修時期や規定の変更
トピック2：上級日本語の授業に対する評価	
【韓国語】 ・レポート作成に関する学び ・内容の適切さ	【中国語】 ・レポート作成に関する学び ・難易度の低さ
トピック3：大学1年生に、身につけておくべきこと	
【韓国語】 ・日常生活に必要な基本的な日本語能力 ・教員とのコミュニケーションのための日本語能力	【中国語】 ・レポート作成、プレゼンテーションのスキル ・教員とのコミュニケーションのための日本語能力
トピック4：上級日本語の授業に対する要望	
【韓国語】 ・実生活で役立つ日本語（敬語、ビジネス日本語） ・初年次教育科目との重なる調整 ・課題の軽減	【中国語】 ・難易度の向上 ・初年次教育科目との重なる調整 ・レベル差の解消 ・成績の評価方法の変更 ・履修時期や規定の変更

### (1) 週当たりの授業回数と科目構成

週当たりの授業回数、および科目構成について、①履修の負担、②科目構成の見直し、③履修時期やルールの見直しといった3点に関するコメントが見られた。

まず、①履修に対する負担について、韓国語母語話者、中国語母語話者の両者から「必修科目は課題が多い」、「日本語の必修科目が多く、他に履修したい科目を取ることができない」といった負担を感じている理由が挙げられている。そして、「大学で学ぶための日本語の習得には週4コマでは足りないが、他にも必修科目が多数あるため、4コマが適切だ」とし、履修の必要性は認めているものの、他の科目との兼ね合いから、現状維持が適切とするコメントが見られた。

次に、②科目構成の見直しについては、「本学部に入學するには、一定レベル以上の日本語力が要求されており、その基準をクリアした留学生に対し、4技能すべての科目を履修させる必要があるのか」という疑問が示されていた。また、プレゼンテーションの方法やレポートの書き方は、大学での学習活動で必須のスキルであるため、これらのアカデミックスキルを中心とした科目構成にすることを希望する意見が見られた。

最後の③履修時期や規定の変更については、上級日本語科目に対する要望の節で詳しく述べる。

### (2) 上級日本語の授業に対する評価

上級日本語の授業に対する評価としては、①レポート作成に関する学び、また、授業内容の問題点として、②内容の不適切さ、③難易度の低さの3点が挙げられている。

①レポート作成に関する学びについては、大学の様々な授業でレポートが課されることが多く、「何よりも早く学ぶべき重要なこと」として、学習活動に必要なスキルと認識されている。また、大学入学以前にレポート作成について学んだ経験がないため、新たな学びであり、自分自身の伸びが感じられたという理由も示されていた。

次に、授業内容の問題点として挙げられた②内容の不適切さについては、現行の授業内容に対し、「日本語以外の科目の学習に役立つと感じられない」、「それぞれの技能のスキルをどうすれば獲得できるのかを学びたい」といった意見があった。そして、③難易度の低さについては、「授業内容が簡単で、日本語学校の授業と変わらない」、「独学でできる内容だ。もっとチャレンジしたい」という意見が見られた。

以上の通り、授業の内容と難易度については、本学におけるアカデミック・ジャパニーズとして必要な学びとは何か、そして、履修する留学生の日本語レベル考え、何をどのように扱うべきかを見直す必要があるだろう。

### (3) 大学1年次に身に付けておくべきこと

大学1年次に身につけておくべきこととして、①レポート作成・プレゼンテーションの能力、②教員とのコミュニケーションのための日本語能力、③日本で暮らすために必要な基本的な日本語能力という意見が挙げられている。

①レポート作成・プレゼンテーションの能力は、大学での学習活動において、必要不可欠なスキルと捉えられている。また、②教員とのコミュニケーション能力については、口頭、およびメールでのやりとりにおいて、失礼のない適切なコミュニケーションできているかという不安を感じているようである。そして、③日本で暮らすために必要な日本語能力については、日常生活において、留学生は日々の生活、またアルバイトなどで、多様な人々とコミュニケーションをする機会があり、大学という場での学びに限らず、幅広く対応できる日本語力の必要性を感じているとともに、それらの日本語力を身に付けるための学習機会を望む声もある。

### (4) 上級日本語に対する要望

上級日本語に対する要望には、授業に対するものと、科目の枠組みに関するものと二種類あった。授業に対する要望としては、①実生活で役立つ日本語の学習、②難易度の向上、③初年次教育科目との重複の調整、④課題の軽減の4点である。また、科目の枠組みに関する要望として、⑤クラス内のレベル差の解消、⑥成績評価の方法の変更、⑦履修時期や規定の変更などの3点が挙げられている。

授業に対する要望として、まず、①実生活で役立つ日本語の学習があり、例えば、アルバイトでの必要性から敬語表現の学習、そして、将来の就職を視野に入れたビジネス日本語の学習が必要だとする意見があった。そして、②難易度の向上は、「上級日本語」の授業の問題点として指摘されている通り、本学部の留学生の日本語レベルに相応しい内容、難易度の授業が求められて

いる。③初年次教育科目との重なる調整については、初年次教育科目には「日本語表現（口頭）」「日本語表現（文章）」、上級日本語科目には「口頭表現」「文章表現」があり、いずれも大学での学習に必要な日本語力の育成を目的としている。そのため、上級日本語科目で「口頭表現」「文章表現」を履修した留学生は、初年次教育科目の履修は不要ではないかという意見もあった。現行では、初年次教育科目の「日本語表現（口頭）」「日本語表現（文章）」と、上級日本語科目の「口頭表現」「文章表現」の間で、学習項目や活動などの調整は行われていないことは事実であり、指摘の通り、今後、調整が必要であろう。最後に、④課題の軽減については、1年次は全体的に必修科目数が多いため、日本語科目の課題の負担を軽減してほしいとの訴えがあった。

次に、科目の枠組みに関しては、⑤クラス内のレベル差の解消として、レベル別のクラス編成を希望する意見があった。アンケート調査の結果では、レベル分けの必要性について意見は半々であったが、ヒアリング参加者からは、英語必修科目のようにレベル分けを行い、クラス内の履修者間のレベル差を是正すべきだという意見が見られた。そして、履修者間のレベル差に関連する問題として、⑥成績の評価方法の変更を望む意見がある。理由として、入学時点で留学生の日本語能力に差があり、どんなに努力をしても、元々の差が障壁となり、努力が成績に反映されないため、現行の相対評価ではなく、絶対評価で成績評価をするべきだとの意見が出されている。

最後に、⑦履修時期や規定の変更として、まず、履修時期を英語科目と同様、春学期はインプット（読解・聴解）、秋学期はアウトプット（口頭表現・文章表現）とし、科目別に集中して履修できるように変更してほしいとの要望があった。また、履修の規定について、一点目は英語科目のように、一定レベル以上の日本語能力を有するものは履修科目数を減らすべきだという意見、そして、二点目は日本人学生に比べ、留学生は必修科目数が多いため、必修科目数を軽減すべきだという意見であった。

以上がヒアリング参加者からの要望である。これらは上級日本語履修者の総意を表すものではないが、ヒアリング参加者が自らの経験に基づいて考えた意見や提案がなされており、上級日本語科目の改革を行うための貴重な示唆を得ることができた。

## Ⅶ 総合考察

上級日本語科目を履修した留学生に対するアンケート調査の結果、およびヒアリング調査の結果を踏まえ、今後の改革に向けて特に検討すべき点として、「週当たりの授業回数と科目構成」、「上級日本語の授業に対する要望」の2点について考察する。

### 1. 週当たりの授業回数と科目構成

週当たりの授業回数について、アンケート調査の結果では、週4回を希望する回答が半数を占めたが、残り半数は、週2回、週3回という授業回数の削減を希望する回答であった。ヒアリング調査では、週4回という回数は、大学で学ぶための日本語力を身に付けるためには必要だとの意見がある一方で、必修科目の課題の多さ、履修希望科目と必修科目の時間割上の重なりによる

制約という点から、週4回の授業は負担が大きいとする声があった。

また、科目構成について、アンケートの結果では、現行の「読解」「聴解」「口頭表現」「文章表現」でよいとする回答が約7割を占める一方で、科目数が多いとする回答は3割程度あり、科目数の削減を希望する履修者からは、「聴解と口頭表現」「読解と文章表現」といった技能の統合を望む声があった。同じく、ヒアリング調査でも、本学部に入学できるレベルの日本語力があれば、4技能の全てを履修する必要はないのではないかとの意見があり、技能を統合した科目を希望する声もあった。

これらの結果から、週当たりの授業回数については、アンケートでもヒアリングでも、現行の4回を希望する意見と、授業回数を減らすことを希望する意見の両者がほぼ同数で存在しており、今後の方向性としては、これまでの春学期・秋学期を合わせて8コマを維持、もしくは負担を減らしていくことを考えたい。そして、科目構成においても、授業回数の削減との関連により、再構成が必要であろう。実際の大学の授業活動においては「読む」「聞く」「書く」「話す」を組み合わせて運用する力が必要であり、アンケートおよびヒアリングでも意見として挙がっていた通り、技能を統合した型での授業実践を行うことを検討する。

その他、ヒアリングにおいて、自身の履修希望科目と必修科目が時間割上重なり、希望科目を履修できないという問題に対し、履修期間の延長を求める声もあった。履修期間を延長すれば、選択の自由度が高くなるというメリットはあるだろう。しかし、上級日本語科目が大学で学ぶための基礎となる日本語力の育成を目的とした科目である以上、4年間の大学生活の中で、大学1年次の段階で履修し、日本語能力を身に付けておくことが望ましいと考える。履修者である留学生に対して、上級日本語科目の目的や意義を十分に説明し、理解を得ていきたい。

## 2. 上級日本語の授業に対する要望

留学生が上級日本語の授業に求めていることとして、今回の調査項目の「大学1年次に身に付けておくべきこと」、「上級日本語に対する要望」への回答から、特に検討すべき点として、以下の3点が挙げられる。

### ①「実生活で役立つ日本語」に対する要望

アンケート調査、およびヒアリングのいずれにおいても、留学生が日本社会で暮らすために、日常生活に必要な語彙や表現、また、他者とのコミュニケーションのための若者言葉や敬語表現などに対するニーズが挙げられていた。これまでの授業では、「大学での勉学に対応できる日本語力」という大学におけるアカデミックな場面で必要な日本語という観点から、授業での学習項目は選定されていた。そのため、要望として挙げられた大学外の日常生活における言語活動については、あまり扱われてこなかったと思われる。今回の調査で留学生が安定した留学生活を送る基盤となる日常生活において、サポートとなるような日本語学習を必要としていることが明らかとなった。

## ②「教員とのコミュニケーションのための日本語」

今回のヒアリング調査で、「口頭でもメールでも、先生とコミュニケーションする場合、どのようにしたらいいかを勉強したい」との声があった。

留学生にとって大学に必要な日本語力について、堀井(2003)はこれまでの大学学部留学生への日本語教育で使用されてきた教材の分析や調査の結果に基づき、「『アカデミック・ジャパニーズ』が日本の大学での勉学に必要な日本語力であるならば、そこには、まず、①学習を行うための、講義理解、レポートを書く、口頭発表をするなどの日本語能力(スキル)と、②学習を行うためのネットワーク作りのためのコミュニケーション力が含まれている」(p116)と述べている。つまり、今回の留学生から要望のあった「教員とのコミュニケーションのための日本語能力」は、堀井(2003)の言う②学習を行うためのネットワーク作りのためのコミュニケーション力に相当するものであり、「アカデミック・ジャパニーズ」とは別に位置づけられ、「キャンパス・ジャパニーズ」と分類されるものである。しかし、実際の学部での日本語教育では、アカデミック・ジャパニーズのスキルを獲得するための授業が行われていることが多い(堀井2003)という。本学部の上級日本語科目でも、この「キャンパス・ジャパニーズ」の部分はあまり扱われてこなかったものであり、今後、検討が必要であろう。

## ③「初年次教育科目との重なる調整」

上述の通り、本学部の初年次教育科目の「日本語表現(口頭)」「日本語表現(文章)」と、留学生の上級日本語科目の「口頭表現」「文章表現」の授業内容には、レポート作成やプレゼンテーションといった部分で重なりが見られ、ヒアリング調査で問題点として指摘されている。三宅(2003)は、日本人大学生に対する大学で学ぶための日本語力は、留学生のアカデミック・ジャパニーズと関連しているとし、「非母語話者に母語話者の言語能力と同等のものを期待することはできないものの、大学におけるアカデミック・ジャパニーズの最終目標は、母語話者、非母語話者ともに同じであるとはいえないだろうか」(p108)と述べている。このように、留学生も日本人学生も目標とするものが同じであれば、重なりが生じるのは当然であるが、留学生と日本人学生とで大きく異なる点は、留学生はそれを「外国語」で運用しなければならないという点である。したがって、留学生に対する日本語教育においては、「外国語」であることをサポートすべく、より言語表現面に重点を置き、運用のための練習を行うことが重要ではないかと考えられる。

そして、このような初年次教育と学部留学生に対する日本語教育の重なりの問題は、多数の大学において初年次教育が実施されている現状では、本学部のみの問題ではない。久川(2013)は留学生向け日本語科目担当者の立場から、日本語科目の改善へのプロセスとして、初年次教育との協働の中で取り組むことが不可欠だとしている。初年次教育との重なりを単に避けるということではなく、大学で学ぶために必要な能力の育成という日本人学生も留学生も共通する目的に向け、共にどのように実現することができるかを考えていくことも必要であろう。

## Ⅷ おわりに

本稿は、学部留学生を対象に行われている上級日本語科目について、履修した留学生にアンケート調査・ヒアリング調査を実施し、履修者の留学生が現在行われている本学部の日本語教育に対し、どのように考え、何が必要だと考えているのかを明らかにすることが目的であった。アンケート調査の結果、ならびにヒアリング調査の結果から、現行の授業の回数や科目構成に対する考え、そして、授業に関する問題点が示された。そして、履修者からのニーズとして挙がっていた「実生活で役立つ日本語」や「教員とのコミュニケーションのための日本語」については、これまでの上級日本語科目の授業では十分に扱われていなかったことが明らかとなった。

今回の調査によって、上級日本語科目の改革において、検討すべき重要な課題を得ることができた。これらの結果を基に、本学部の留学生に対するアカデミック・ジャパニーズとして、どのような教育を行っていくのかを検討していきたい。

## 謝辞

本調査を行うにあたって、アンケートにご回答くださった上級日本語科目履修者の皆様、ヒアリングにご参加くださった留学生の皆様、ヒアリングにおいてファシリテーター役を務めてくださったSAのユンファヨンさん、ソルヨンファンさん、侍子安さん、張慶瑞さん、そして、本調査の実施にご支援くださった中野教務事務室内田則之様、菊池真理栄様、本調査の実施に際し、数多くのご助言をくださった本学部の小森和子教授、柳澤絵美准教授、すべての方に心よりお礼を申し上げます。

## 参考文献

- 門倉正美 (2001) 「『日本留学試験』の狙いと問題点—『日本留学試験』の「最終報告書」を読む」『横浜国立大学留学生センター紀要』8, pp.93-112
- 門倉正美 (2003) 「アカデミック・ジャパニーズとは何か」『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語予備教育と大学日本語教育の連携のもとに—』平成14年度～16年度 科学研究費補助金基盤研究費研究成果報告書, pp.123-132
- 久川伸子 (2013) 「日本の大学における学部留学生向け日本語プログラムの現状と課題：改善のための協働」『東京経済大学人文自然科学論集』134, pp.41-50
- 堀井恵子 (2003) 「留学生が大学入学時に必要な日本語力は何か—『アカデミック・ジャパニーズ』と『日本留学試験』の『日本語試験』を整理する」『日本留学試験とアカデミック・ジャパニーズ 日本留学試験が日本語教育に及ぼす影響に関する調査・研究—国内外の大学入学前日本語予備教育と大学日本語教育の連携のもとに—』平成14年度～16年度 科学研究費補助金基盤研究費研究成果報告書, pp.113-122
- 三宅和子 (2003) 「留学生・日本人大学生のアカデミック・ジャパニーズとは」『日本留学試験と



6 どのような科目構成がよいと思いますか。科目名を具体的に教えてください。(例：読解と文章表現)

7 「読解」の授業で、良かった点、悪かった点はどんなことですか。具体的に教えてください。

8 「聴解」の授業で、良かった点、悪かった点はどんなことですか。具体的に教えてください。

9 「口頭表現」の授業で、良かった点、悪かった点はどんなことですか。具体的に教えてください。

10 「文章表現」の授業で、良かった点、悪かった点はどんなことですか。具体的に教えてください。

11 「上級日本語」を履修して、自分ができるようになったと思うことはどんなことですか。具体的に教えてください。

12 「上級日本語」の授業で、「こんなことを学びたかった」ということがありますか。それはどんなことですか。

13 「上級日本語」の授業で、授業活動（プレゼンテーション、ディスカッション、ペア/グループワークなど）を日本語で行うことに不安がありましたか。

不安があった      不安はなかった      どちらともいえない

14 どんな点で不安を感じていましたか。

15 レポートの提出やプレゼンテーションの前に、あなたの「日本語」について、日本人の友人など誰かにネイティブチェックをしてもらっていますか。

いつもしてもらっている      ときどきしてもらっている  
まったくしてもらっていない

16 レポート作成のサポートが受けられるシステムがあれば、利用したいと思いますか。

ぜひ利用したい      ときどき利用したい      利用しなくてよい

17 「英語」の必修と比べて、「上級日本語」の授業を履修する負担はどうでしたか。

英語の方が大変だった      日本語の方が大変だった  
英語も日本語も同じくらい大変だった      英語も日本語もあまり大変ではなかった

18 「英語」はレベル別に授業が行われていますが、「日本語」に関してもレベル別に授業を行ったほうがよいと思いますか。

レベル別がよい

現在のままでよい（レベル別ではない）

どちらともいえない

19 「上級日本語」の授業について、感想やコメントや要望などがあったら、自由に書いてください。

### Ⅲ. 「上級日本語」授業に関するヒアリングへのご協力をお願い

本学部の日本語科目について、今後の改善のため、みなさんから率直な意見を伺いたいと思っています。ヒアリングは、オンライン（小グループ）で行います。時間は1時間程度、実施時期は6月中旬以降の予定です。（教員は立ち会いません）

20 「上級日本語」ヒアリングに参加できますか。

参加を希望する

参加してもよい

参加しない

21 お名前とメールアドレスを教えてください。（ご記入のアドレスに、ヒアリングの日程などのご連絡をいたします）

【資料紹介】

## クリエイタービジネス論における音楽視聴調査 (2018～2021)

A Research of contact to music in the Creator Business Class  
(2018-2021)

原田悦志  
HARADA, Nobuyuki

### I はじめに

明治大学国際日本学部で開講されているクリエイタービジネス論では、この講座を受講した明治大学生を対象にした音楽視聴調査を、2018年から継続して実施している。

この調査は、筆者がチーフ・プロデューサーを務めていたNHKワールドTVの音楽番組『J-MELO』で行われていた『J-MELO リサーチ』<sup>(1)</sup>を原型として始まった。全世界の日本音楽ファンの実像を解明すべく、2010年から2017年にかけて東京藝術大学研究科（2010～2015）及び慶應義塾大学アート・センター（2016～2017）の協力を得て、番組ウェブサイトに掲載されたアンケートに視聴者が記入し、返信する形式で実施されたものである。

授業におけるリサーチの当初の目的は、調査研究の基礎を教えることであった。初回である2018年においては、「学年、年齢、性別」等の定点観測的な項目に、調査を担当する学生の興味や関心を加味した、基礎的な形で試行された。それ以降、年を追うごとに学生たちが改善を繰り返し、質問項目も増えていった。2020年からは大手広告代理店である博報堂の調査研究員の助力も得て、より多様で深化した音楽視聴調査へと発展した。

実施年ごとの回答者数は、61名（2018年）、53名（2019年）、60名（2020年）、82名（この講座を受講していない明治大学生22名を含む、2021年）であった。本稿においては、計256名の明治大学生が回答した4年間の調査から、筆者が特に重要だと考えた3点「主たるメディアの交代がもたらした、音楽視聴のあり方の変化」「セレンディピティ（偶然の幸せな発見）が鍵となる選択の基準」「コロナ禍で加速化したリモートアクセスの増加」について論じたい。さらに、2021年の調査に基づき、「サウンドの一部としての日本語歌詞に関する考察」を行ったが、これは、筆者の主要な研究テーマである「日本音楽のさらなる世界伝播」の嚆矢となり得るものだと考えている。

なお、各年度の調査報告は、明治大学国際日本学部のウェブサイトに掲載している。<sup>(2) (3) (4) (5)</sup>

## Ⅱ クリエータービジネス論における学生たちの音楽視聴の特徴

### 1 主たるメディアの交代がもたらした、音楽視聴のあり方の変化

まず、第一に挙げられる事象は、学生が支持する音楽家の多様化である。

2020年の調査において、名が挙がったアーティストは89組に上った（複数回答）。しかし、その中での最高得票は、わずか4票にすぎなかった（星野源、ONE OK ROCK、SEKAI NO OWARI、RADWIMPSの4組が、同票で第1位）。

2021年の調査では、「最も好きなアーティストは誰か」という単一回答の設問において、最高得票は3票となった（防弾少年団（以下、BTS）、SEKAI NO OWARIの2組が、同票で第1位）。また、「好きなアーティストを全員答えてください（複数回答）」という設問では、実に314組も名が挙がったが、最高得票のYOASOBIですら、8票にとどまった。

つまり、教室の中で、圧倒的に支持を受けるアーティストは存在しなかったのだ。この特徴は、音楽視聴のあり方の変化と密接に結びついている。放送(broadcast)から、通信(communications)へという、メディアの主役の交代だ。

かつては放送、とりわけテレビ番組は、新しい歌手や楽曲を知る主たる手段だった。世帯視聴率のオンライン調査が始まった1977年9月26日以降のデータ<sup>(6)</sup>によると、レギュラー番組においては、歴代第1位の『ザ・ベストテン』（TBS系列）が41.4%（1981年9月17日放送分）、第2位の『夜のヒットスタジオデラックス・スペシャル』（フジテレビ系列）は36.0%（1988年2月10日放送分）を記録している。

スペシャル番組では、第1位の『第35回NHK紅白歌合戦』（1984年12月31日）は78.1%を記録し、日本のおよそ8割の世帯が視聴した。小泉今日子、チェッカーズらが初めて出場し、都はるみの最後のステージとなった紅白である<sup>(7)</sup>。

それに対し、2021年においては、レギュラーの音楽番組が番組平均世帯視聴率で10%を超えることは稀である。例えば、2021年9月20日から26日にかけてのビデオリサーチにおける音楽番組の平均世帯視聴率のトップは『うたコン』（NHK）の7.1%だが、番組平均個人視聴率（4才以上の個人全体の視聴率）では3.8%と、さらに低い数値になっている<sup>(8)</sup>。また、2020年12月31日に放送された『第71回紅白歌合戦』は、第一部が34.2%、第二部が40.3%<sup>(9)</sup>と、健闘しているとはいえ、第35回の時点からはほぼ半減している。

では、現代の学生たちは、どのように音楽の情報を得て、視聴しているのだろうか。2021年の調査から紐解いてみよう。

「あなたが普段、アーティストや楽曲、イベントなどについて知る情報源としているものは何ですか。あてはまるものを全てお選びください（複数回答）」という質問の答えは、第1位がYouTubeで、全体の75%を占めた。以下、Twitter（68%）、アーティストの公式SNS（67%）と続き、ソーシャルメディアが主たる情報源となっていることが分かる。

それでは、彼らはテレビを観なくなったのだろうか。「あなたはテレビを観ますか（単一回答）」

という問いに対して、図1の通り、「毎日観る」と答えた学生は、「1時間以上」「1時間未満」を合わせると、52%に達した。

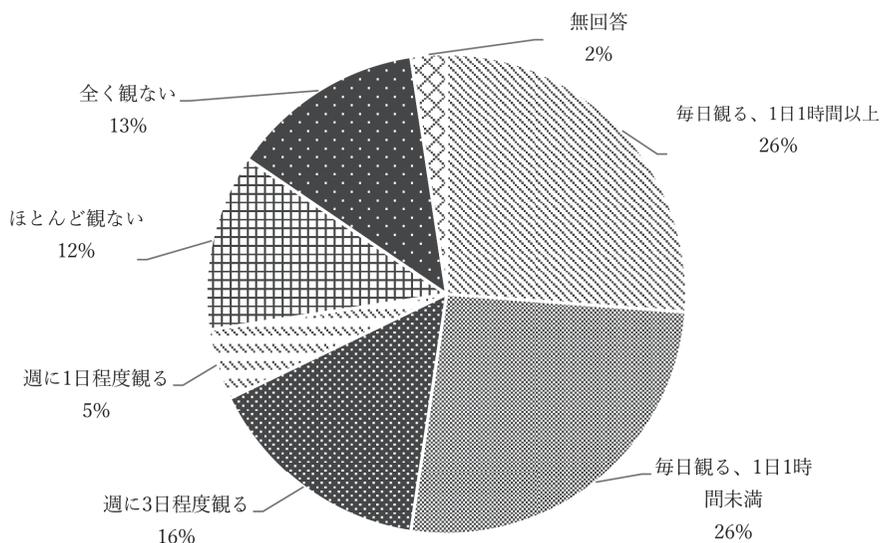


図1 テレビの視聴頻度 (地上波とBS)  
出所 2021年の調査資料より

また、「あなたが音楽を視聴する時に、ここ一年間で利用したことのあるメディアは何ですか (複数回答)」と尋ねたところ、第1位はストリーミングサービス (24%)、第2位は動画・ライブ配信アプリ (23%) だが、テレビ番組も第3位 (15%) につけている。

現時点においては、まだテレビを観る学生が少なからず存在していることが分かる。だが、音楽情報の取得、及び音楽の視聴源としては、既にテレビは主役の座をSNSに譲っていることがうかがえる。「あなたが音楽を視聴する時に、ここ一年で利用したことのある機器は何ですか (複数回答)」という設問で、最も頻度が高い機器であったのは、90%を占めたスマートフォン (以下、スマホ) であった。さらに、「今後減ると思われるものは何ですか (複数回答)」という問いに対しては、「テレビ番組の視聴」(32%) が、「CDやDVDレンタルショップの利用」(34%) に次いで、僅差の第2位となった。スマホを主たる音楽の聴取源としている学生たちは、将来的には、さらにテレビ番組を視聴することが減少すると予想している。

このような行動形態を裏付けるように、「どんな時に音楽を聴きますか (複数回答)」という問いで第1位となったのは「通勤・通学の行き帰りの時」であり、実に96%を記録した。「テレビ・ラジオを視聴している時」と答えたのは、わずか3%にとどまった。

この結果から浮かび上がってくるのは、1980年代のようにテレビ番組を通じて一つのヒットチャートを共有しているのではなく、個人が独自に情報を得て、スマホにイヤホンを接続して楽

曲を楽しむという、1人1ジャンルともいえる音楽視聴のあり方である。

転機をもたらしたのは、学生が音楽視聴に際し、最も頻繁に利用しているメディアである、ストーリーミングサービス（定額聴き放題）の発達だ。日本での開始は、Apple Musicは2015年、Spotifyは2016年のことであった。以後、ヒットの基準は、「売れた枚数」から「聴かれた回数」へと大きく変わった<sup>(10)</sup>。サービスが始まった時期は、調査に応えた学生の多くが、中学生から高校生ぐらいの時にあたる。彼らは、多感な年頃に、国境や時代を越えた音楽をスマホで無尽蔵に聴き、好きな音楽家や楽曲を発見していった。つまり、「J-Pop」や「ロック」といった既存の枠組みではなく、各自が見つけて構成した多種多様な楽曲のプレイリストこそが、自らにとっての音楽ジャンルとなったのだ。

## 2 セレンディピティ（偶然の幸せな発見）が鍵となる選択の基準

次に、学生たちが何を基準に好きなものを選択しているかについて、分析したい。キーワードはセレンディピティ、すなわち「偶然の幸せな発見」である。

2018年の調査において「『魅力的な店舗や商品』を選ぶ基準は何ですか（複数回答）」という問いかけを行った。その結果、「身近な人のおすすめ」（87%、第2位）、「SNSでの口コミ」（66%、第3位）を抑え、第1位となったのは「直観・雰囲気/偶然」（93%）だった。それらに対し、放送は16%、キュレーターは8%しか選択基準としている者はいなかった。つまり、テレビで識者が勧めるものより、自分の直観や偶然の出会いを重視する学生が、はるかに多かったのである。翌2019年の調査でも同じ質問を行ったところ、「直観・雰囲気/偶然」を最も重要視するという学生は、92%に達した。

2021年の調査では、学生たちのセレンディピティをより深く掘り下げる質問を設けた。「キュンとくる感情」、つまり、恋心のようにときめく思いを、どのように抱いたかを尋ねた。図2は芸能人や著名人にそのような感情を持った経緯を表したものだ。70%を占める新発見とは、ドラマや映画、CMなどを偶然見て、新たにその人を見つけたことを意味する。個人選択とは、好きなグループやチームに所属している者に対してキュンとくる感情が発展した、いわゆる「推しメン」のことだ。おススメは、他者からの推薦のことである。

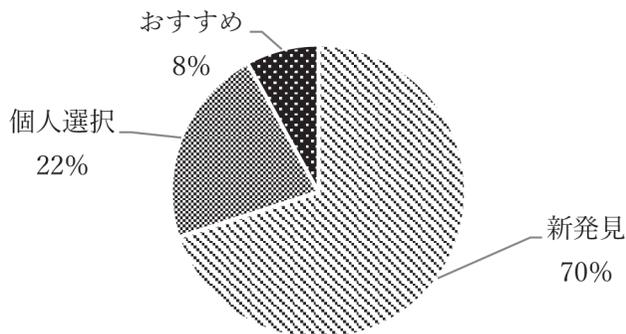


図2 芸能人や著名人にキュンとした経緯  
出所 2021年の調査資料より

その結果、男性、女性共にほとんど割合は変わらず、「作品などとの偶然の出会いによってキュンとくること」が、圧倒的多数だった。つまり、「自分で見つけ出した」という感覚が重要なのだ。

それでは、音楽に関するセレンディピティを分析してみよう。「曲にキュンときた経路は何ですか（複数回答）」という質問に対して、第1位となったのはYouTubeだった。第2位はテレビや映画だが、僅差の第3位が音楽系アプリ（ストリーミングサービス）となっている。

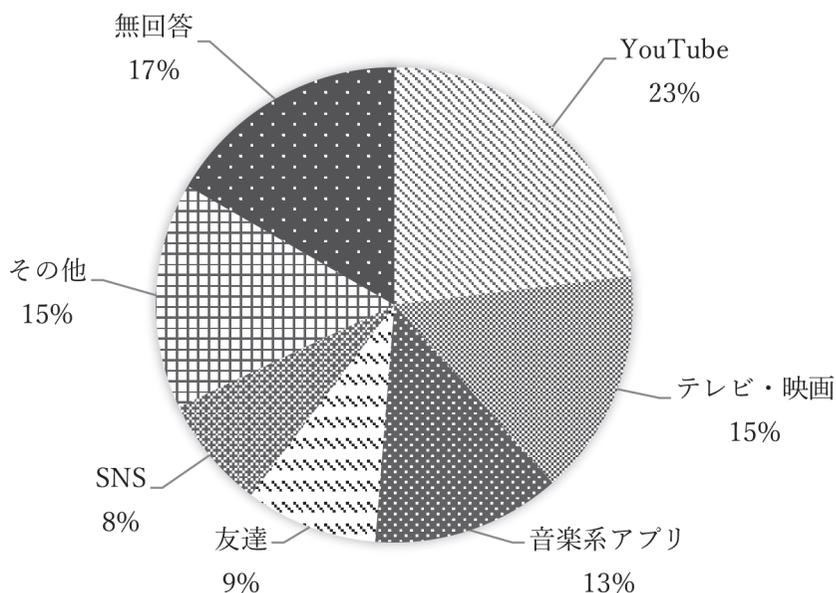


図3 曲にキュンときた経路  
出所 2021年の調査資料より

YouTube やストリーミングサービスには AI によるデータ解析に基づくおすすめ機能が標準的に装備されている。動画や楽曲を再生すると、視聴傾向に沿った関連コンテンツが次々と紹介されるというビジネスモデルである。II -1 で、新しいヒットの基準は「売れた枚数」から「聴かれた回数」へと遷移したと指摘した。それに加えて、従来のテレビ視聴率に代わり、頻繁に支持の尺度として用いられているのが、動画再生サイトでの視聴回数だ。

一例として、YouTube で 2019 年 11 月に開始された「一発録り」でパフォーマンスすることをルールとしている『THE FIRST TAKE』を取り上げたい。このコンテンツは、2021 年 9 月末時点で総再生回数が 13 億回を超えており、中でも第 1 位である LiSA の『紅蓮華』は、1 億 1 千万回以上再生されている<sup>(11)</sup>。『紅蓮華』は、テレビアニメ『鬼滅の刃』のオープニング曲である。LiSA のことを知らなくても、アニメ作品を観てこの曲を知り、「いいな」と感じた人も多いと推察する。そこで、この曲をもっと聴きたいと思って YouTube で探すと、『THE FIRST TAKE』での歌唱が上位となって見つかる。そして、その画面には、LiSA の他の曲や、『THE FIRST TAKE』に参加している数多くのアーティストの作品が「おすすめ」として紐づけられている。このようにして、『紅蓮華』をきっかけとして、他の曲にもたどり着く人が増え、視聴回数は加

算されていく。

つまり、「偶然見つけて、好きになったもの」は、実は「偶然を装い、あてがわれたもの」である可能性があるのだ。自分で実際に見たり、聞いたりすることより、スマホを入り口にして、インターネットを介して調査や発見を行うことに慣れている者は、バーチャルな空間にあるセレンディピティは演出されているということを知っておく必要がある。

2021年にセレンディピティに関する項目の調査を行った学生は、当初、コロナ禍で外出の自粛が要請される中、お気に入りのモノやコトを偶然に発見する経緯は、全面的にメディアを介しているという仮説を立てていた。しかし、確かにメディアを通して見つけるという答えが多数ではあったが、「店内のBGMで、偶然、お気に入りの音楽を見つけた」という学生も4人いるなど、外出先でお気に入りのモノやコトを発見するケースがなくなるわけではないということが分かった。この結果を受けて、「実際に見て、触れて感じたセレンディピティは、メディアを介して得たそれよりも、実体験として強い印象を与えるのではないか」と考察した。

だが、もしかしたら、その店で、そのタイミングで、その曲が流れたことが意図的だったという確率もゼロではない。セレンディピティは、元来、作り出されるものではなく、主体的かつ自然に生まれてくるものだ。しかしながら、偶然に出会ったかのように装うという仕掛けは、音楽においても、ヒットを作り出す鍵となっているのだ。

### 3 コロナ禍で加速化したリモートアクセスの増加

続いて、2020年に発生したコロナ禍が、学生たちの音楽との関わり方に与えた変化について分析する。

2020年の調査において、「COVID-19（新型コロナウイルス感染症）の流行により、あなたの音楽生活に変化があれば教えてください（複数回答）」と質問した。すると、プラスの変化という答えが57%と過半数となり、マイナスの変化の18%を大きく上回った。ポジティブな答えの内訳では「音楽を聴く時間が増えた」（29%）、「聴くジャンルが増えた」（19%）、「ライブ映像視聴」（16%）が上位3件であった。

これに対し、ネガティブな答えの上位3件は「ライブが中止になった」（43%）、「音楽を聴く時間が減った」（28%）、「カラオケに行けなくなった」（14%）だった。

コロナ禍において多くの者に共通している体験は「外出の自粛、在宅時間の増大」「対面イベント等の減少、縮小」などであろう。ポジティブな変化は、自分でコントロールできる時間が増えたため、それを音楽視聴に充てたというものが上位を占めていたのが分かる。「初心者が楽器を始めた」「経験者の楽器の練習時間が増えた」という楽器の習得についても、両者を合算すると17%に達した。何かを始める、あるいは何かに打ち込むために、空いた時間を利用する者も現れたのだ。

ではなぜ、ネガティブな変化の第2位が、「音楽を聴く時間が減った」という答えだったのだろうか。担当した学生は、「音楽は、通勤通学時にスマホで聴くという者が大多数を占めている。

その時間がなくなったことにより、音楽聴取の時間も減少したのでは」と考察した。

そして、カラオケに行けなくなったことは、新しい音楽を聴く機会を奪ったことも意味した。2018年の調査では、「どこで音楽を聴くか(複数回答)」という問いに対し、「ライブ会場」(26%)、「公共交通機関」(20%)を抑えて第1位となったのは「カラオケ」(36%)であった。すなわち、コロナ禍以前は、カラオケは「歌う」だけでなく、同席した者による歌唱を「聴く」場でもあったのだ。2021年の調査においても、「生活内で減った事(複数回答)」として、「カラオケの利用」(41%)がトップとなっている。つまり、カラオケに行けなくなったことにより、未知の曲を対面で知る主要な場が消失してしまったのである。

「コロナ禍で増えた事(複数回答)」では、「オンデマンドサービスの利用」(51%)、「有料音楽サービス」(29%)と、リモートによる音楽との接触が増大した。対面での音楽体験の代わりに学生が選択したのは、放送への回帰ではなく、インターネットメディアへのさらなる傾斜だったのだ。担当した学生は「コロナ禍が、メディアの転換に大きな関連性があると指摘できる」「YouTubeやTikTokなどの短い動画に慣れている若者にとって、テレビの特徴である長時間の制約が、マイナスに働いているのではないか」と考察した。

### Ⅲ サウンドの一部としての日本語歌詞に関する考察

明治大学国際日本学部は、知識や経験を積み重ねて、自分だけの「日本と世界をつなぐ力」を身につける舞台である<sup>(12)</sup>。この調査も、留学生を含む受講生の音楽視聴の解析だけに留まらず、日本の音楽を介して、世界との繋がりを見出せるようなインテリジェンスを提示するものでなくてはならない。

そのような目的で2021年に新たに設けたのが、母国語以外の音楽視聴に関する質問である。筆者は、世界中を取材する中で、「日本語の響きはエレガントである」という声を複数の国や地域で聞いた。このような意見は、日本語を母国語として用い、サウンドではなく意味で理解してしまう筆者には分からないことだ。しかし、逆の立場になれば、フランス語は潇洒、ドイツ語は硬い感じというように、意味があまり分からない言語については、語感をサウンドとして感じ取っている。

考察の一例として、K-PopグループのBTSが、韓国語の歌詞で歌った『Life Goes On』を挙げてみたい<sup>(13)</sup>。この曲は2020年12月に、全米ビルボードのメインチャート「HOT100」で首位に輝いた。だが、アメリカのファン全員が、この曲の歌詞の意味を理解していたとは、到底、思えない。しかしながら、サウンドの一部としての韓国語歌詞の響きが、多くのアメリカ人にとって耳心地良いものであったと推定できる。

学生は、「あなたは普段、母国語以外の音楽を聴きますか(単一回答)」という問いに対して、87%が「はい」と答えた。続いて、「母国語以外の歌詞を初めて聴く時、どんな点を最も重視していますか(複数回答)」と尋ねると、91%が「サウンド」と答え、「歌詞の意味」と返答したのはわずか9%だった。

「母国語以外の歌の歌詞の意味を知ろうと思いますか（単一回答）」という質問には、「知ろうと思う」が65%、「知ろうと思わない」が25%と答えた。過半数が意味を知りたいと思っているが、4分の1の学生は、何を歌っているかを理解するの必要性を感じていないのだ。

「あなたが『心地良い』と思える母国語以外の歌詞の楽曲は、どんな言語のものですか（複数回答）」という設問では、英語が58%、韓国語が28%となり、この2つの言語だけで86%に達した。英語を選択した学生は「リズムカルである」「韻を踏んでいる」など、英語がポップ・ミュージックのリズムに載りやすいという理由を述べる者が多かった。韓国語については「ダンスポップに合うから」「日本語と音が似ている」「響きがかわいい」というような意見が述べられた。さらに、「どんな言語でも、曲調とリズム感、ビートなどによって心地よく感じる」という、特定の言語にこだわらず、サウンドとして快適であれば良いという声もあった。なお、留学生は全員が日本語を選択した。

この結果から読み取れるのは、ファーストコンタクトに際しては、外国語楽曲の歌詞は、意味や内容ではなく、サウンドの一部として捉える者が多いということだ。日本語を母国語としない者の立場で考えてみれば、日本語を理解できなくても、その歌詞の響きが音楽の要素として心地良いと感じられるなら、文字通り、耳目を集める可能性があるのだ。

母国語以外の楽曲を聴く学生が多くなっているのは、主たる音楽聴取の方法がストリーミングサービスとなり、世界中の音楽を容易に聴くことができる環境に起因すると考えられる。同じように、日本語の楽曲も、いまでは全世界の人々が聴くことができる。「シニフィアン（記号表現）としての日本語歌詞」の世界的調査は、これからの主要な研究課題である。

#### IV おわりに

2021年に20歳を迎える者は、2001年生まれ、つまり、21世紀の初年に誕生した。この調査は、そのような世代を中心とする学生たちが、自分たちの音楽視聴について考察したものである。そんな彼らが音楽を聴いてきた期間である、ここ10数年の世界の音楽マーケットは、どのように変化してきたのだろうか。

図4を参照すると、世界の音楽市場は、2014年に底を打って以来、成長を続けている。コロナ禍のために伸びは鈍化したが、2020年もプラス成長を遂げた。その原動力となっているのは、学生たちが音楽を聴取する際に最も利用している、ストリーミングサービスだ。そして、2021年度の調査結果にあった通り、学生の回答者の9割近くは、普段から外国語楽曲を聴いている。彼らと世界のマーケットは、ダイレクトに繋がっているのだ。

これに対し、図5を参照すると、日本の音楽市場は過去10年間に渡ってプラス成長しておらず、むしろ全体的にはマイナス傾向にある。また、いまだに中心的な役割を果たしているのは、世界的には退潮にあるオーディオレコードである。日本のフィジカルな録音成果物を海外で入手するのは、容易ではない。日本の音楽産業は、世界の動きと密接にリンクしているとはいえないのである。

リサーチを行ったこの4年間は、「放送から通信へ」「フィジカルからデジタルへ」といったメディアの転換や、コロナ禍というパンデミックの発生など、さまざまな事柄が地球的規模で変化した期間でもあった。クリエイタービジネス論は、毎年平均60人ほどの音楽好きの学生が受講する小規模な講座だ。しかしながら、その受講生らから寄せられたデータを解析し、考察したものは、狭い教室だけで起こっていた事象ではなく、現代における世界の音楽視聴の状況が凝縮された調査報告だったことが分かる。

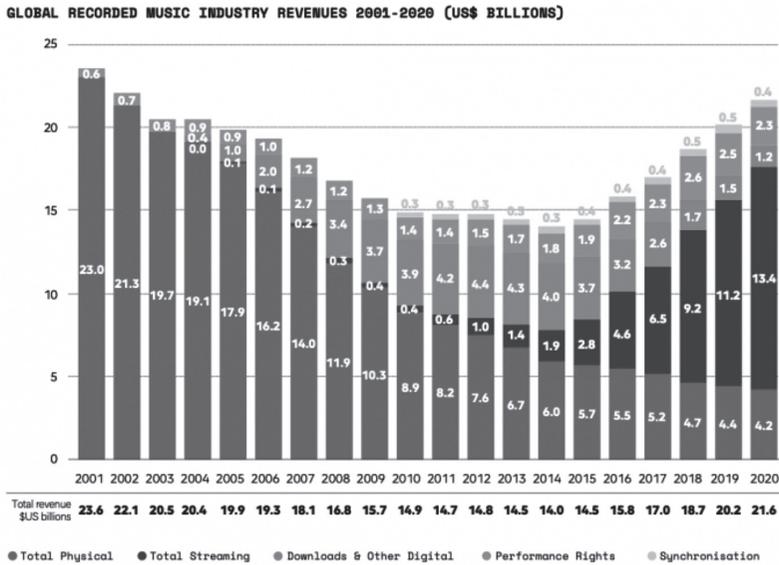


図4 全世界における年間売上レポート  
出所 IFPI (国際レコード産業連盟) ウェブサイト

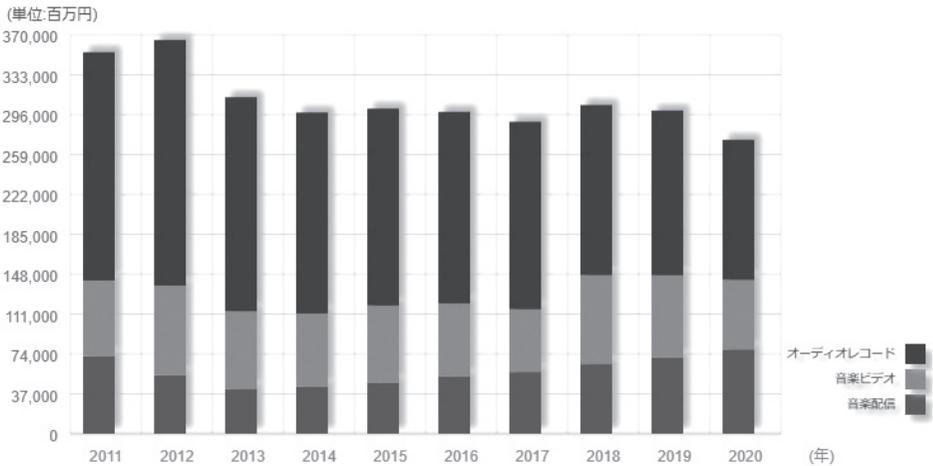


図5 生産実績・音楽配信売上 過去10年間合計  
出所 日本レコード協会 ウェブサイト

今後の課題は、海外の日本学研究者らと連携した、日本音楽に関するリサーチの世界的共有だと考える。日本音楽のさらなる世界伝播のためには、「クリエイタービジネス論のクラス/世界の音楽マーケット」のような密接な関係性を、「海外の学校のクラス/日本の音楽マーケット」でも構築するヒントを、世界中の音楽ファンから得なければならない。各国の大学や研究機関などでも同じような調査を行うためには、報告書の多言語化、とりわけ英訳を行うことが必須だ。

筆者は、引き続き、学生たちと共に、音楽に関する視聴調査を続けていきたい。未来において、「2010年代から20年代にかけて、あなた方の先輩たちは、このように音楽と関わっていたのだ」と、後世の現役学生と語り合えたらと願っている。

最後に、任意参加にもかかわらず、データの収集や解析を行った調査担当者に、心より謝意を表したい。

調査担当者（所属及び学年は実施当時）

（2018年）

伏見 宏司（国際日本学部3年）

白須 弘起（国際日本学部1年）

森下 涼帆（国際日本学部1年）

（2019年）

楊 帆（国際日本学部4年）

宮本 真緒（国際日本学部3年）

今川 満帆（国際日本学部2年）

木村 有希（国際日本学部2年）

稲村 俊佑（国際日本学部1年）

千葉 瑞斗（国際日本学部1年）

三輪 響（国際日本学部1年）

（2020年）

豊永 梨乃（国際日本学部4年）

青木 優花（情報コミュニケーション学部4年）

早坂 亮太郎（国際日本学部3年）

丹羽 智香（国際日本学部3年）

松島 早希（国際日本学部3年）

狩野 将大（情報コミュニケーション学部4年）

木立 友里香（国際日本学部2年）

中島 悠（国際日本学部2年）

西本 朱里 (国際日本学部 1 年)

大川 萌生 (国際日本学部 1 年)

匿名 (総合数理学部)

(2021 年)

粕谷 夏実 (国際日本学部 4 年)

金澤 英貴 (国際日本学部 3 年)

安部 玲音 (総合数理学部 3 年)

伊藤 希実香 (国際日本学部 2 年)

西上 友理 (国際日本学部 2 年)

岩崎 梨沙 (国際日本学部 2 年)

JEONG YUSHIN (国際日本学部 1 年)

三枝 伶紗 (国際日本学部 1 年)

中島 堯太 (国際日本学部 1 年)

清水 晃希 (国際日本学部 1 年)

丸橋 恵那 (国際日本学部 1 年)

伊藤 想 (国際日本学部 1 年)

吉川 拓海 (文学部 1 年)

(調査協力 / 2020・21 年)

木下 陽介 (博報堂)

谷口 由貴 (博報堂)

各年度の質問項目

(2018 年)

61 名が回答した。

1 あなた自身のことを教えてください。(学年、年齢、性別、出身地等)

2 音楽について、どこで情報を得ていますか？

3 音楽をどのように聴いていますか？

4 「魅力的な店舗や商品」を選ぶ基準はなんですか？

5 どのような音楽関連商品 / イベントに興味がありますか？

6 あなたは音楽を聴くとき、音質を気にしていますか？

7 あなたが音楽を聴きたくなるのはどのようなときですか？

8 あなたは音楽を聴いて泣いたことがありますか？あるなら、なぜですか？

9 あなたは楽器を弾けますか？なんの楽器を弾くことができますか？

10 あなたが最も好きなアーティストと、その理由を教えてください。

(2019年)

53名が回答した。

- 1 あなた自身のことを教えてください。
- 2 音楽について、どこで新しい情報を得ていますか？
- 3 音楽をどのように、あるいはどこで聴いていますか？
- 4 魅力的な店舗や商品を選ぶ基準
- 5 どのジャンルの音楽が好きですか？
- 6 音楽にかける支出
- 7 楽曲を好きになるとき、最初に心を惹かれる部分
- 8 あなたが最も好きなアーティストと、その理由を教えてください。
- 9 あなたが音楽を聴きたくなるのは、どんなときですか？
- 10 あなたが初めて「好き」になった音楽はどの曲ですか？
- 11 あなたにとって音楽とはどんな存在ですか？

(2020年)

60名が回答した。うち52名に対しては、追加調査も行った。

アンケート回答者のデータ<sup>(14)</sup>

- 1 あなたはどの程度、音楽が好きですか？とても好きなら10、全く好きではないのであれば0としてお答えください。
- 2 あなたが1のように答えた理由を、音楽にまつわるエピソードも踏まえて、なるべく具体的に説明してください。
- 3 あなたの好きな日本のアーティスト（歌手やミュージシャン）を教えてください。
- 4 あなたがときめく存在のアーティストや俳優、芸人などを教えてください。
- 5 あなたは、音楽について、この1年での平均で、1日あたり平日・休日で、どの程度、時間を使っていますか？
- 6 あなたが、音楽を視聴する時に利用している機器は何ですか？
- 7 音楽作品に関する行動（商品・サービスの利用など）について、あなたの行動としてあてはまるものをお答えください。
- 8 あなたはこの1年間に、音楽に関する商品・サービスについて、お金を使って利用したことがありますか？
- 9 あなたはどのような時に音楽を視聴しますか？
- 10 あなたが、普段、アーティストや楽曲、イベントなどについて知る情報源としているものは何ですか。

- 11 あなたが新しいアーティストや楽曲を知るきっかけになる情報源は何ですか。
  - 12 あなたが、お金を支払って利用している音楽サービスは何ですか。
  - 13 あなたが普段、音楽と接する際に、「ここが不満だ、これは使いにくい」と思うことを教えてください。
  - 14 COVID-19（新型コロナウイルス感染症）の流行により、あなたの音楽生活に変化があれば教えてください。
- 追加1 ステイホーム期間中に始めた生活習慣、熱中した趣味や習い事などで、いまでも継続していることはありますか？それは、どんなことですか？
- 追加2 追加1 であげたものを継続することができている理由は、どんなことに理由があると思いますか？

(2021年)

82名が回答した。うち22名は「クリエイタービジネス論」を受講していない明大生であった。

- 1 性別
- 2 年齢
- 3 出身地、出生地
- 4 国籍
- 5 母国語
- 6 1か月に自由に使える金額
- 7 使用しているメディア
- 8 好きなアーティスト
- 9 好きな楽曲
- 10 偶然の発見
- 11 キュンについて
- 12 セレンディピティ
- 13 聴く時間と頻度
- 14 音楽と関連する行動について
- 15 母国語以外の音楽視聴の有無
- 16 母国語以外の音楽について
- 17 母国語以外の歌詞の意味について
- 18 母国語以外の「耳心地良い」音楽
- 19 コロナ禍による生活の変化
- 20 コロナ禍による音楽視聴の変化
- 21 コロナ禍で増えた事・減った事
- 22 コロナ禍を経て、今後、増えるもの・減るものは何だと思うか

- 23 音楽を視聴するタイミング
- 24 音楽に関する情報源
- 25 新しい音楽を知るきっかけ
- 26 音楽を聴くときに使用している機器
- 27 音楽を聴くときに使用しているサービス
- 28 この一年間の音楽視聴に関する行動
- 29 音楽に関する課金の有無
- 30 使用している有料サービス

### 注

- (1) 「J-MELO リサーチ」  
NHK ワールド TV  
2010 年から 2016 年までの調査報告書が掲載されている。  
<https://www.nhk.or.jp/j-melo/nhkworld/english/research/2010.html>
- (2) 「明大中野リサーチ 2018 ～明大生の音楽視聴に関する諸調査～」  
クリエイタービジネス論リサーチ班：伏見宏司、白須弘起、森下涼帆  
<https://www.meiji.ac.jp/nippon/info/2018/6t5h7p00000sdihk.html>
- (3) 「クリエイタービジネス論リサーチ 2019」  
<https://www.meiji.ac.jp/nippon/info/2019/6t5h7p00001p77nq.html>
- (4) 「クリエイタービジネス論リサーチプロジェクト 2020」  
明治大学クリエイタービジネス論 リサーチプロジェクトメンバー  
<https://www.meiji.ac.jp/nippon/info/2020/6t5h7p000036kbtn.html>
- (5) 「クリエイタービジネス論リサーチプロジェクト 2021」  
明治大学クリエイタービジネス論 リサーチプロジェクトメンバー  
<https://www.meiji.ac.jp/nippon/info/2021/6t5h7p00003coblj.html>
- (6) 「ビデオリサーチ 週間高世帯視聴率番組」  
1977 年 9 月 26 日以降に放送された、関東地区の番組平均世帯視聴率。番組タイトルが同じものは同一番組とし、最高視聴率を記録したものが抽出されている。放送分数 15 分未満の番組は除く。  
[https://www.videor.co.jp/tvrating/past\\_tvrating/music/01/post-14.html](https://www.videor.co.jp/tvrating/past_tvrating/music/01/post-14.html)
- (7) 「NHK 紅白歌合戦ヒストリー」NHK  
[https://www.nhk.or.jp/kouhaku/history/history\\_35.html](https://www.nhk.or.jp/kouhaku/history/history_35.html)
- (8) 「関東地区の視聴率」ビデオリサーチ  
<https://www.videor.co.jp/tvrating/>
- (9) 「関東地区の視聴率」ビデオリサーチ

なお、個人視聴率では、第一部 25.0%、第二部 29.6% と、さらに低減する。

[https://www.videor.co.jp/tvrating/past\\_tvrating/music/02/nhk-2.html](https://www.videor.co.jp/tvrating/past_tvrating/music/02/nhk-2.html)

- (10) 『日本で国民的ヒット曲を生み出すための「最も有効な仕掛け」とは』

現代ビジネス 柴那典 2017年12月25日掲載

<https://gendai.ismedia.jp/articles/-/53891?page=4>

- (11) 「YouTube チャンネル THE FIRST TAKE」再生回数 1 ~ 10

[https://www.youtube.com/playlist?list=PLjHW8hmA8BqF1YV\\_OudSrq9GR4BADPpgC](https://www.youtube.com/playlist?list=PLjHW8hmA8BqF1YV_OudSrq9GR4BADPpgC)

- (12) 「国際日本学部とは」

明治大学国際日本学部

<https://www.meiji.ac.jp/nippon/about/outline.html>

- (13) 『BTS、韓国語歌詞の「Life Goes On」がビルボードのメインシングルチャート「HOT100」で1位を獲得!』

M-ON! MUSIC 2020年12月3日掲載

<https://www.m-on-music.jp/0000447538/>

- (14) 2020年度においては、性別、年齢、出身地などは質問項目には含めず、基礎データとして扱った

ウェブサイトは全て 2021年10月1日閲覧

### 参考文献

小泉文夫 (1994) 『日本の音 世界のなかの日本音楽』 平凡社

町田健 (2004) 『ソシユールと言語学 コトバはなぜ通じるのか』 講談社

川原繁人 (2015) 『音とことばの不思議な世界 メイド声から英語の達人まで』 岩波書店

坂口博樹著 櫻井進数学監修 (2016) 『数と音楽 美しさの源への旅』 大月書店

柴那典 (2016) 『ヒットの崩壊』 講談社

ベン・パー著 依田卓巳、依田光江、茂木靖枝 訳 (2016) 『アテンション』 飛鳥新社

デレク・トンプソン著 高橋由紀子訳 (2018) 『ヒットの設計図』 早川書房

ロール・クレア、レイエノブノワ・レイエ著 根来龍之監訳 門脇弘典訳 (2019) 『プラットフォーム勝利の法則』 日本経済新聞出版

早稲田大学坪内博士記念演劇博物館監修 後藤隆基編 (2021) 『ロスト・イン・パンデミック 失われた演劇と新たな表現の地平』 春陽堂書店



## 『明治大学国際日本学研究』刊行に関する内規

国際日本学部

### 1. 有資格者

- (1) 明治大学国際日本学部専任教員、特任教員、客員教員、兼任教員（国際日本学部授業科目を担当している兼任のみ）、定年により専任教員を退職した者（共著論文の場合、筆頭著者が上記の者であれば投稿できることとする）
- (2) 上記（1）のほか、教授会が特に認めた者

### 2. 内容

学術研究に関わる論考（「研究論文」「翻訳」「資料紹介」「研究ノート」と依頼論文）。退任記念号では、退任教員の履歴・業績、記念文章（退任者の希望による）なども掲載する。

### 3. 発行

- (1) 発行回数 原則として、年1回
- (2) 発行部数 500部

### 4. 雑誌形態

- (1) 用紙 B5版
- (2) 文字 MS明朝, 10.5pt
- (3) 縦書きは右頁から、横書きは左頁からおこす。
- (4) 頁数は左右から付ける。

### 5. 原稿募集

7月中旬を締切りとして原稿を募集する。執筆希望者は応募に際して、原稿の種類についての希望を紀要編集委員会に伝える。募集原稿の種類は、次の4つとする。

- (1) 研究論文 (Research Article)
- (2) 翻訳 (Translation)
- (3) 資料紹介 (Introduction of Research Materials)
- (4) 研究ノート (Research Notes)

### 6. 提出原稿

- (1) 和文または欧文
- (2) 字数 和文 32,000字以内、欧文 8,000ワード以内（表題・図版・表・スペース等を含む）
- (3) 電子テキスト（メール添付による word、フロッピーディスク、CD など）およびプリントアウト原稿
- (4) 手書きの場合は完全清書原稿

### 7. 提出期日

10月初旬（厳守）

### 8. 査読と掲載原稿の決定

掲載原稿は、別に定める基準に従って査読をおこない、査読結果を参考にして紀要編集委員会で掲載の可否を決定する。その際、「研究論文」としての掲載を希望した執筆者に対して「研究ノート」としての掲載を提案することがある。査読は、原則として学部所属教員がおこなう。

### 9. 校正

執筆者校正は再校までとする。なお、原則として、校正時点で行を超えるような加筆は認めない。念校は、紀要編集委員会にて行う。

### 10. 抜刷

執筆者には抜刷 50部を上限として無料配布する。但し、50部以上必要な場合は、執筆者の実費負担により、配布できることとする。

### 11. その他

- (1) 原稿料は支払わない。
- (2) 執筆者には紀要2部及び論文データ（CD-R1枚）を配布する。

#### 附則

この内規は、2010年6月11日より施行する。

附則（2010年度規則6, 7）

この内規は、2011年1月29日より施行する。

附則（2011年度規則8）

この内規は、2012年6月8日より施行する。

附則（2012年度規則2, 7）

この内規は、2013年3月4日より施行する。

附則（2013年度規則2）

この内規は、2013年10月18日より施行する。

附則（2014年度規則7）

この内規は、2014年6月20日より施行する。

附則（2015年度規則2, 5, 8, 10）

この内規は、2015年7月3日より施行する。

附則（2017年度規則1, 5, 8, 9）

この内規は、2017年6月16日より施行する。

附則（2017年度規則1）

この内規は、2018年2月24日より施行する。

附則（2021年度規則3, 8）

この内規は、2022年1月30日より施行する。

本号執筆者

白戸伸一	国際日本学部教授
尾関直子	国際日本学部教授
廣森友人	国際日本学部教授
小森和子	国際日本学部教授
大矢政徳	国際日本学部教授
田中絵麻	国際日本学部専任講師
安高紀子	国際日本学部特任講師
Garside Paul	国際日本学部特任講師
原田悦志	国際日本学部兼任講師
王ウエイトン	国際日本学部兼任講師
大塚清高	国際日本学部兼任講師
伊集院郁子	東京外国語大学教授
	国際日本学研究院教授
李在鎬	早稲田大学国際学術院教授
大江宏子	ボーンマス大学
	ビジネススクール教授
黄叢叢 /	明治大学国際日本学
黄秀智 /	研究科博士後期課程
吴梅	
金田慎吾 /	
大谷温理 /	
柴田亘	
新井智大 /	明治大学国際日本学
大石有紗	研究科博士前期課程

編集委員

○ 鈴木賢志  
白戸伸一  
瀬川裕司  
張競  
横田雅弘  
(○編集委員長)

表紙デザイン

森川嘉一郎

明治大学国際日本学研究 第14巻第1号  
白戸教授 退任記念号

Global Japanese Studies Review, Meiji University Vol. 14, No.1  
Special Edition in Honor of Professor Shirato

2022年3月31日 発行

発行 明治大学国際日本学部

School of Global Japanese Studies, Meiji University  
〒164-8525 東京都中野区中野4-21-1

印刷 共立速記印刷株式会社

〒102-0072 東京都千代田区飯田橋 3-11-24

