

東洋大学 国際教育センター紀要

Bulletin of the Toyo University Center for Global Education and Exchange

第四号

No.4

March 2026



目 次

巻頭言	荒 卷 俊 也 …………… 1
研究論文 (Academic Papers)	
From Résumé to Rhetoric: A Genre-Based Analysis and Pedagogical Application of English Job Application Letters	Ian Hart …………… 3
研究ノート (Research Notes)	
From Textbook to Test: Vocabulary Frequency Analysis to AI-Generated Quizzes	Timothy Ang and Katsuichiro 'Ken' Ohashi ……………31
報告 (Reports)	
交換留学生用日本語科目における「訳語自動化単語リスト」導入の試み	大 澤 恵 利 ……………55
難民日本語教育の研究動向	片 山 奈緒美 ……………65
高度外国人材としての留学生のインターンシップの課題と展望	渡 部 裕 子 ……………81
編集後記	……………95

国際教育センター紀要第4号巻頭言

国際教育センター長

荒 卷 俊 也

東洋大学国際教育センターは、2014年の「スーパーグローバル大学創成支援」事業（SGU）への採択に伴い2015年度から新組織として改組され、本学の国際化への取り組みを先導、支援してきました。外国語力強化への様々な取り組み、留学生に対する日本語教育やキャリア形成の支援、各種留学プログラムおよび海外インターンシップ及び海外ボランティア活動の実施、協定校の開拓と国際学術交流の支援などです。

このような活動を進めていく中で得られた学術的な知見の公表を目的として、2022年度より本紀要の刊行を開始し、今回第4号が刊行される運びとなりました。第4号では、査読付き研究論文を含めた5編が掲載されております。編集、査読にご協力いただいた皆様、そして本紀要を利用して研究成果の公表をいただいた著者の方々には改めて御礼を申し上げます。

本学では、SGU事業終了後も「SGU × SDGsによる国際教育の推進」を教学中長期計画の柱の一つに掲げ、部局ごとに目標を設定し、その部局の特徴に応じた国際化の取り組みを展開しております。また国際教育センターにおいても、正課内外での外国力強化プログラム、留学生向けの多様な日本語教育プログラム、留学生のキャリア形成支援、そして、留学やインターン、ボランティアなどの海外派遣のプログラムを展開し、多文化共生グローバル人材の育成に力を入れております。本紀要ではこれらの活動に関係して、英語教育および日本語教育に関する報文とともに、留学生のキャリア形成支援のためのインターンシップに関する研究成果も紹介されております。

本紀要と通じたさまざまな学術的な成果の公表と対話が、より質が高く、効果的な教育・研修プログラムの設計や実施に波及していくことを期待しております。

2026年3月

Foreword for the Fourth Issue of the Bulletin of the Center for Global Education and Exchange

Director of the Center for Global Education and Exchange
Toshiya Aramaki

With the University's selection for the Top Global University (SGU) Project in 2014, the Toyo University Center for Global Education and Exchange was reorganized in AY 2015 and has been leading and supporting initiatives for the internationalization of the University since then. The initiatives include various efforts to improve students' foreign language skills; support for Japanese language education and career development for international students; implementation of various study abroad programs, overseas internships and overseas volunteer activities; increasing the number of partner universities; and support for international academic exchange.

With the aim of publicizing the academic knowledge obtained through these activities, we began to publish the bulletin in AY 2022, and this is the fourth issue. The fourth issue includes five articles, such as peer-reviewed research papers. I would like to take this opportunity to once again express my gratitude to everyone who cooperated with editing and peer review, as well as to the authors who have used this journal to promote their research results.

Toyo University has set goals for each department and has been developing initiatives for internationalization tailored to the characteristics of the departments, positioning the "Promotion of International Education through SGU x SDGs" as one of the pillars of its medium- to long-term academic affairs plan since the completion of the SGU Project. In addition, the Center for Global Education and Exchange has been focusing on fostering global-minded human resources who can contribute to a multicultural society by promoting programs to strengthen foreign language skills both inside and outside of the regular curriculum, various Japanese language education programs for international students, career development support for international students and overseas dispatch programs such as study abroad, internships and volunteer work. In connection with these activities, this bulletin introduces research achievements on internships to support the career development of international students, as well as reports on English and Japanese language education.

I hope that the publication of various academic achievements and discussions through this journal will have an impact on the design and implementation of higher-quality and more effective education and training programs.

March 2026

From Résumé to Rhetoric: A Genre-Based Analysis and Pedagogical Application of English Job Application Letters

Ian Hart

Center for Global Education and Exchange
Toyo University

Abstract

This article explores the application of a genre-based approach to the analysis of English Job Application Letters (EJALs), with a particular focus on Japanese university students and graduates preparing for international studies or business careers. The aim of the study is to examine the move-step structures and linguistic features of EJALs in order to provide pedagogical insights for English for Specific Purposes (ESP). Using a corpus of 15 authentic solicited letters, the analysis has revealed that most of the identified moves are commonly employed, and that grammatical features such as tense usage and register are strongly move-specific. Building on these findings, the study also reports on a classroom application in which explicit instruction in moves and language features helped students produce more coherent and professional cover letters. These results underscore the importance of genre-based pedagogy for equipping non-native English learners with the skills necessary for effective business writing in globalized professional environments. This study is believed to contribute to ESP by demonstrating how genre analysis can be directly translated into pedagogical practice for Japanese learners.

Keywords: genre analysis, English for Specific Purposes (ESP), job application letters, move-step analysis, Japanese EFL learners

Introduction

In recent years, the need for English business communication has grown rapidly within Japanese universities. Many institutions have established new language departments aimed at preparing graduates for participation in international business contexts. Increasing numbers of students now aspire to work overseas or for foreign-affiliated companies, while domestic firms engaged in global markets are also adopting English as their working language. Notably, large Japanese corporations such as Rakuten and Uniqlo have designated English as their official language of communication, requiring graduates to be proficient in both spoken and written

business discourse.

As more graduates seek to enter such organizations, it has become essential for them to understand international job application practices. One important distinction is the English Job Application Letter (EJAL), or cover letter, which remains uncommon in Japan but is widely expected elsewhere. As Baugh (1991) explains, the EJAL serves both as an introduction of yourself and a strategy to interest employers sufficiently to read your resume and to grant you an interview. Mastery of this genre, therefore, is a crucial skill for Japanese graduates aiming to compete in globalized employment markets.

Genre analysis has been widely applied in English for Specific Purposes (ESP) research to better understand the communicative purposes, structural moves, and linguistic patterns of workplace texts (Bhatia, 1993; Henry & Roseberry, 2001; Swales, 1990; Wang, 2005; Wijayanti, 2017). Building on this tradition, the present study employs an established move-step framework to investigate 15 authentic EJALs. By examining both structural organization and linguistic features such as tense and register, this paper aims to provide insights that are not only descriptive but also pedagogically valuable for the teaching of business writing to non-native English learners. While many studies have analyzed EJALs in other contexts, few have examined their pedagogical application for Japanese learners.

Genre-based Approach to Analysis

Since the 1980s, applied linguists working in English for Specific Purposes (ESP) have increasingly turned to genre analysis as a way of understanding written discourse, according to Shamsudin et al. (2012). Swales (1990) defines genres as communicative events that are goal-directed and shaped by shared communicative purposes. These events typically exhibit schematic structures with identifiable beginnings, middles, and endings. Building on this view, Dudley-Evans (1987) described genre analysis as “a system of analysis that can reveal something of the patterns of organization of a genre and the language used to express those patterns in the light of communicative purposes” (p.229), highlighting its value in uncovering form–function relationships.

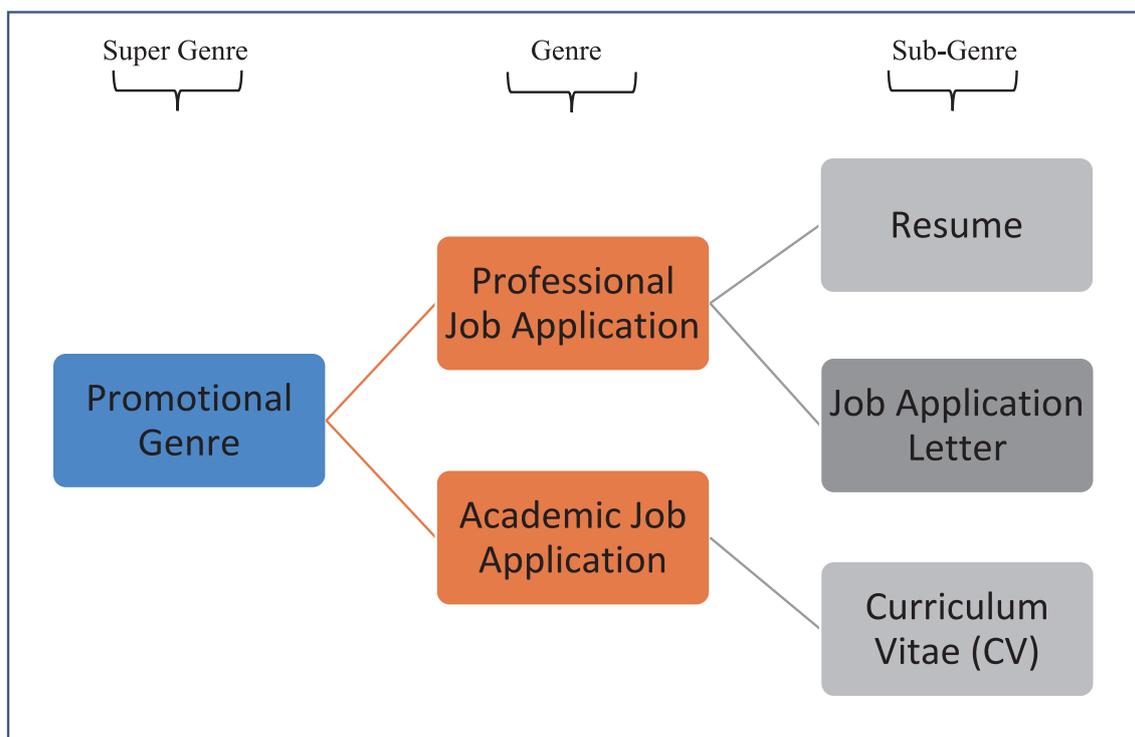
In the case of English Job Application Letters (EJALs), adherence to these formal conventions is essential to fulfilling the communicative purpose of persuading an employer. At the same time, standards of organization and language use are not universal; they are shaped by socio-cultural influences that vary across language communities. As Connor et al. (2002) show in their comparative study of business English writing in Belgium, Finland, and the United States, differences in rhetorical expectations can pose significant challenges for writers attempting to transfer their skills across contexts. This is particularly relevant for Japanese learners, who must adapt from domestic practices, where cover letters are rarely used, to international norms that

demand genre-specific rhetorical strategies.

EJALs as a Genre

Previous studies of English Job Application Letters (EJALs) have examined both their generic structures and linguistic features (Bhatia, 1989; Henry & Roseberry, 2001). Bhatia’s (2004) framework of genre generalization identifies the job application letter as a sub-genre of both academic and professional applications (*Figure. 1*). These are forms of promotional discourse that share the same communicative purpose: to persuade the reader and, ultimately, “to elicit a specific response from its reader, in this case a call for interview” (Bhatia, 1993, p. 45).

Figure 1 Levels of genre generalization



(Bhatia, 2004, p. 59)

As a promotional genre, the job application letter functions not only to introduce the candidate but also to summarize qualifications, highlight strengths, and spark the interest of the employer. This requires writers to meet audience expectations regarding both structure and language use. Early genre studies of EJALs provided foundational frameworks for analyzing these rhetorical strategies, particularly through move-step analysis, which reveals how applicants organize content and deploy language to achieve persuasive effect. Such frameworks are essential for understanding

the conventions of EJALs and for guiding pedagogical practice in teaching business writing.

This is especially relevant in the Japanese context, where traditional application practices differ considerably. Japanese companies typically require a standardized résumé (*irekisho*), which emphasizes factual, chronological information with limited space for self-promotion. Unlike EJALs, the *irekisho* does not provide an opportunity to highlight individuality, persuasive self-evaluation, or rhetorical strategies designed to impress the employer. As a result, Japanese graduates seeking employment in international companies, or domestic firms that have adopted globalized hiring practices, must learn to adapt to the rhetorical expectations of EJALs, a task that can be challenging without explicit instruction in genre conventions.

Move and Step Analysis

The most well-known example of the analysis of moves and steps in genre was created by Swales (1990). Swales created the Create a Research Space (CARS) model (*Table 1*), which consists of three “moves” with several possible steps within them. Ding (2007) defines a move as a “functional unit in a text used for some identifiable purpose” and “it is often used to identify the textual regularities in certain genres of writing” (p. 369). It can also be used to “describe the functions (or communicative intentions) which particular portions of text realize in the relationship to the overall task” (Connor, Davis, & De Rycker, 1995, p. 463). A step is a smaller unit than a move and it is placed on a lower level than a move. The length and size of a move and step can vary. Sometimes they may be one sentence while other times they can be several paragraphs long. However, they must contain a least one proposition (Connor & Mauranen, 1999). According to Swales, moves/steps can vary in sequence, can be repeated, and may be embedded within another (1990). Swales used this model to identify the schematic structures of introductions in 48 academic research articles from three different disciplinary domains, with the hope of reflecting a common written academic genre. Since its first publication, the CARS model has been widely studied, revealing its recursive nature, and validating its structure across various disciplines (Bruce, 2009; Cross & Oppenheim, 2006; Martin, 2003). The CARS model has influenced genre analysis in ESP, especially concerning the teaching of academic writing. Since then, many researchers have created their own moves and step patterns that can be used in different disciplines. This includes work by Bhatia (1993), Henry and Roseberry (2001), Wang (2005), and Furka (2008). In these variations, moves can be omitted, changed, or rearranged to suit the writers’ rhetorical response (Dudley-Evans, 2000).

Table 1 Swales' (1990) CARS model for research article introductions

Move 1: Establishing a Territory
Step 1: Claiming centrality, and/or
Step 2: Making topic generalizations(s), and/or
Step 3: Reviewing items of previous research
Move 2: Establishing a niche
Step 1A: Counter-claiming, or
Step 1B: Indicating a gap, or
Step 1C: Question-raising, or
Step 1D: Continuing a tradition
Move 3: Occupying a niche
Step 1A: Outlining purposes
Step 1B: Announcing present research
Step 2: Announcing principal findings
Step 3: Indicating RA structure

Bhatia (1993) pioneered the study of genre analysis on job application letters. Bhatia presented a seven-move model, where each move functions differently (Table 2). According to Bhatia (1993), the seven-move structure provides a rhetorical organization that writers of a job application letter can follow. With Bhatia's pattern, the writer does not need to use all of the moves. This was evident in Fung's (2014) study, where only five of the moves were necessary to analyze sample academic job application letters. Fung found that Move 3 (*Offering incentives*) and Move 6 (*Using pressure tactics*) are not employed in academic job application letters. Reasons were given such as the

Table 2 Bhatia's (1993) seven-move structural model of a job application letter

Move	Descriptors
Move 1: Establishing credentials	The writer establishes his credentials as someone who can fulfill the needs of the addressee.
Move 2: Introducing candidature <i>Step 1: Offering candidature</i>	The writer introduces himself The writer mentions the offer that he is referring to.
<i>Step 2: Essential detailing of candidature</i>	The writer presents his essential details for the job applied.
<i>Step 3: Indicating value of candidature</i>	The writer evaluates himself for the job applied.
Move 3: Offering incentives	The writer offers incentives to the addressee.
Move 4: Enclosing documents	The writer encloses documents to complement the letter.
Move 5: Using pressure tactics	The writer puts pressure on the addressee to solicit desired responses.
Move 6: Soliciting response	The writer solicits further contact with the addressee.
Move 7: Ending politely	The writer ends the letter politely.

academic discourse community emphasizing “greater value on the applicant’s educational background, research profiles and teaching experiences which are largely conveyed in Move 2”, and the use of pressure tactics not being necessary when only “niche skills and qualifications are highly sought after” (p. 37). In addition, the order of some of the moves can be rearranged. This provides more freedom for the writer to organize the moves in a way that best suits their communicative function. However, Bhatia (1993) explains that some of the moves are “obligatory”, as they are present in all promotional letters. These moves include *Introducing the offer*, *Soliciting response*, and *Ending politely*.

Bhatia’s (1993) move structure has been utilized as a framework for a variety of studies in various locations around the world (Khan & Tan, 2012; Mustaffa et al, 2007; Wang, 2005). However, Henry & Roseberry (2001) believed that Bhatia’s model neglected the idea of analysis in the letters of native speakers. This led to Henry and Roseberry identifying eleven moves in their move analysis of 40 job application letters written by native speakers (*Table 3*). Of the eleven moves, Henry and Roseberry explained that two of them can be accomplished by different strategies (*Table 4*). However, unlike Swales’ (1990) and Bhatia’s (1993) models, these were not considered to be steps as none of the strategies “appeared to be part of a series in a specific order” (p. 159).

Table 3 Henry & Roseberry’s (2001) eleven-move structural model of a job application letter

Move	Definition
Opening (O)	The writer identifies the target and invites the target to read the letter.
Referring to a job advertisement (AD)	The writer refers to the advertisement in which the position was named and described.
Offering candidature (CA)	The writer states an interest in applying for the position.
Stating reasons for applying (RA)	The writer gives reasons for wanting the position.
Stating availability (A)	The writer indicates when he or she would be able to take up the position.
Promoting the candidate (P)	The writer presents selected information demonstrating qualifications and abilities relevant to the desired position.
Stipulating terms and conditions of employment (TC)	The writer indicates expectations regarding salary, working hours, and other relevant contractual matters.
Naming referees (R)	The writer names referees who will support the candidature.
Enclosing documents (EN)	The writer lists documents enclosed with the letter
Polite ending (PE)	The writer ends the letter in a conventional manner.
Signing off (SO)	The writer signs his or her name in a respectful manner, thus claiming ownership of the letter

Table 4 Henry & Roseberry's (2001) list of moves that can be accomplished by different moves

Move	Strategies
Promoting the candidate (P)	Listing relevant skills, and abilities. Stating how skills and abilities were obtained. Listing qualifications. Naming present job. Predicting success. Listing publications. Giving reasons for leaving present job. Demonstrating knowledge of target position.
Polite Ending (PE)	Welcoming response. Inviting favorable consideration. Thanking. Offering to provide further information.

Differences between Bhatia's (1993) and Henry and Roseberry's (2001) move structure include the combination of *establishing credentials* and *offering incentives* into a new moved name *Promoting candidature*, as Henry and Roseberry believed that they held the same communicative purpose. In addition, four additional moves were found, including *referring to a job advertisement*, *stating reasons for applying*, *stipulating terms and conditions of employment*, and *naming referees*. Finally, the move, *using pressure*, proposed by Bhatia, was not included in Henry and Roseberry's structure as they believed "I look forward to..." is used to signal the end of a letter, rather than put pressure on the reader to reply.

This insight led to other researchers examining move structures that better-suited writers of other languages. Wang (2005) examined the schematic structure of job application letters by Chinese applicants. Their research led to the addition of four new moves, including *establishing self*, *offering self*, *referring to enclosed materials*, and *inviting further action*. In an examination of job application letters written by Hungarian EFL learners, Furka (2008) found that self-appraisal was not valued positively due to cultural differences. This was highlighted in Upton & Connor's (2001) study, which took a cross-cultural concept into account, where move interpretation differences were found between Americans, Belgians, and Finns. Al-Ali (2004) also carried out a contrastive analysis between Arabic and English job application letters. Due to different moves being used on either side, moves were emitted to suit both Americans and Jordanians. Move analysis by these researchers supports the belief that different languages use the same genre in different ways.

Job applicants from different cultural backgrounds use and arrange rhetorical moves of a job application letter similarly in general, but differently in specific. However, the main communicative purpose of this genre is always to "elicit a specific response from its reader" (Bhatia, 1993, p. 45)

or to “create a desire on the part of the employer to hire the job applicants” (Al-Ali, 2004, p. 5). This has important implications for Japanese learners. Whereas domestic hiring practices emphasize the *rirekisho* (a factual, standardized résumé with minimal rhetorical elaboration), international employers expect applicants to craft persuasive EJALs that highlight individual qualifications through carefully structured moves. Understanding these rhetorical frameworks is therefore essential for Japanese graduates aiming to succeed in globalized employment markets.

Grammatical and Lexical Features

For a job application letter to be successful, it is important for the applicants’ writing to be free of grammatical and word errors. In addition, Henry and Roseberry (2001) observed a correlation between the moves of a job application letter and the tenses used. In Allison, Townend, and Emmerson’s (2008) study, they found that it is not necessary for applicants to understand all English tenses, as long as they understood past-simple, present-simple, present-perfect, and present progressive. Faiza (2012) states that the present progressive tense is often used by the applicant to show recognition of what kind of individual the employer is looking for and also to refer to the source of the job advertisement. Wang (2005) also suggested that the present progressive is used to politely end the letter. The present perfect tense is more apparent at the start of the letter during the introductory stages, where the applicant details their candidature and highlights their work experience (Bhatia, 1993).

The choice of appropriate register is also an important factor as it indicates a degree of formality in a business letter. Butt et al. (2000, as cited in Young, Fitzgerald, & Fitzgerald, 2018) explain the notion of genre as follows: “When texts share the same context of the situation to a greater or lesser extent, they will share the same experiential, interpersonal and textual meanings and so they belong to the same register” (p. 9). Register refers to the variety of language that varies according to the way it is used in different situations and can be defined as “a configuration of meanings that are typically associated with a particular situational configuration of field, mode, and tenor” (Young, Fitzgerald, & Fitzgerald, 2018). Henry and Roseberry (2001) also highlight register in their study and refer to it as “the totality of all lexico-grammatical realizations of a genre” (p. 154).

Table 5 *Definitions of Field, Tenor and Mode (Young, Fitzgerald, Fitzgerald, 2018, p.230)*

1. Field	This accounts for what language is being used to talk about, accounting for the content of what is being said in different contexts.
2. Tenor	This explains the types of interactions between interactants carried out through language in the situation.
3. Mode	This has to do with the nature of language itself, whether it is spoken or written, or spontaneous or planned discourse. Each mode – spoken or written – and its subcategories have certain characteristics that determine discursive choices in different situations.

Literature Gap and Research Questions

Although prior studies have established the structural and linguistic characteristics of English Job Application Letters across various cultural contexts, limited research has examined how genre-analytic findings can be systematically translated into classroom pedagogy for Japanese learners. In particular, there is a lack of empirical work linking move analysis directly to instructional design and observable improvements in student writing. To address this gap, the present study investigates both the rhetorical features of authentic EJALs and their pedagogical application in a Japanese university setting. The study is guided by the following research questions:

1. What rhetorical moves and linguistic features characterize English Job Application Letters in a corpus of authentic solicited applications?
2. To what extent are these moves and features consistently realized across professional fields?
3. How does explicit genre-based instruction influence Japanese university students' ability to produce coherent and professional EJALs?

Methodology

Data Collection

A total of 15 solicited English Job Application Letters (EJALs) were selected for analysis. Solicited letters are written in response to a specific, advertised vacancy and are addressed to a known employer (Bhatia, 1993). The sample texts were collected from a range of publicly available online sources and represented different professional fields, written by successful job applicants. To protect anonymity, all personal details, such as names, company names, and addresses, were

removed. Letterhead details were also excluded, as they are not considered rhetorical moves within the genre.

Analytical Framework

The study employed a genre-based approach, focusing on rhetorical moves and linguistic realizations. Henry and Roseberry's (2001) move-structure model was adopted as the primary analytical framework. This model was chosen because, unlike Swales' (1990) framework, it was derived specifically from job application letters and provides insights into the conventions of native speaker writing within an ESP context.

Grammatical and Lexical Analysis

In addition to identifying rhetorical moves, the analysis examined grammatical and lexical features in order to highlight how language is used to realize communicative purposes. Particular attention was given to tense usage, such as the present perfect in introductions or the present progressive in closings, and to register, which reflects the level of formality and professionalism expected in business communication. Together, these dimensions provide a more comprehensive picture of how EJALs achieve their persuasive function.

Results and Discussion

Structural Features

Moves

Out of the eleven moves, only nine were identified in this study (*Table 6*). *Stipulating terms and conditions of employment* and *naming referees* did not occur. Also, *stating availability* only occurred once. Generally, the terms and conditions of employment and the start date of the position tend to be decided by the employer and presented in the job advertisement. Therefore, it would not be necessary for the applicant to discuss these in the application, in most cases. It is common for referees to be listed in the applicants' CV/resume, meaning the move *enclosed documents* would satisfy the function of this move. In Henry and Roseberry's (2001) study, similar findings were identified, with *stating availability* and *naming referees* only occurring twice, and *stipulating terms and conditions of employment* only taking place four times over a total of 40 EJALS. Henry and Roseberry explain that these moves are optional and that "context plays a crucial role in deciding which moves are optional and which obligatory" (p. 159). This study supports Henry and Roseberry's belief that O, CA, P, PE, and SO are considered obligatory moves.

Table 6 Occurrence frequency of moves in the current study

Moves	Occurrences	Percentage %
Opening (O)	15	100%
Referring to a job advertisement (AD)	12	73.3%
Offering candidature (CA)	14	93.3%
Stating reasons for applying (RA)	11	73.3%
Stating availability (A)	1	6.7%
Promoting the candidate (P)	15	100%
Stipulating terms and conditions of employment (TC)	0	0%
Naming referees (R)	0	0%
Enclosing documents (EN)	9	60%
Polite ending (PE)	15	100%
Signing off (SO)	15	100%

Opening occurred in all of the samples by referring to the recipient's title and surname or position (e.g., Dear Mr. Hart, Dear Hiring Manager). This move is common in all business letters (Henry & Roseberry, 2001).

Referring to a job advertisement and *offering candidature* often happened within the same sentence, with twelve cases of *offering candidature* taking place first (e.g., "I am writing to apply for the position of Media Relations Assistant Manager, which I saw advertised on MediaJobs.com"). In the letter related to Journalism, *offering candidature* did not explicitly occur, although it is indirectly implied when a reason is stated referring to a job advertisement in the following sentence: "Your news reporter opening advertised on Monster is an excellent match to my qualifications". In four samples, no reference is given to the job advertisement, supporting Henry and Roseberry's (2001) idea that it depends on the context. In Letters 12 (Marketing) and 13 (Management), the applicants have been working in the same industry for an extended period of time. It is possible that they learned about the new positions by word-of-mouth or through other connections/relations.

In eleven of the letters, *stating reasons for applying* was used to introduce positive self-evaluation. This was done with the use of verbs, in particular, "believe". For example, "I believe that my extensive experience in media makes me an ideal candidate". In ten of the letters, this move is used as a transition into *promoting the candidate*.

The move, *promoting candidate*, allows the writer to present "selected information demonstrating qualifications and abilities relevant to the desired position" (Henry & Roseberry, 2001, p. 159). In this study, the move accounted for the majority of the letters. For example, it accounted for 79.1% of the letter related to Marketing, with the pattern of discourse ranging from expository to narrative. This is Henry and Roseberry's (2001) introduction of strategies, rather than

steps that follow a set pattern. *listing of qualifications* and *listing of publications* did not occur often. Also, *giving reasons for leaving present job* was not evident in any of the samples. One explanation is that the majority of applicants were university graduates, so they may not have had jobs in the same field before.

The move *enclosing documents*, also proved to be optional, as it only occurred in nine of the letters. Important documents tend to be attached to job application letters, so the applicants may not have needed to address this. Supportive documents are also a requirement for most applications, so a cover letter would not be read without them being submitted.

The final two moves, *polite ending* and *signing off* were used to close the letters and occurred in all samples. One of the strategies noted by Henry and Roseberry (2001), *welcoming response* was evident in most of the samples. This was used to suggest further contact (e.g., interview, telephone call). However, in six of the letters, a more direct strategy was implemented to guarantee a response. Direct speech was used using the verb “will” in sentences such as:

- *I will call your office next week to inquire...*
- *I will call your office next Wednesday to follow up...*
- *...and will email you within the next week to arrange a time to speak further.*

In these cases, I feel Bhatia’s (1993) steps *using pressure tactics* and *soliciting response* a better suited to describe the strategy being used. Finally, the writer “signs his or her name in a respectful manner, thus claiming ownership of the letter” (Henry & Roseberry, 2001, p. 159).

The following fields had the same move patterns: Design and Engineering, Marketing and Management (see *Table 7*). Although, there is no strong evidence to explain why this happened. *Opening* always occurred first, and *polite ending* and *signing off* always occurred last. The remaining moves did not always follow the same pattern, with some moves (RA and PE) occurring more than once in some of the letters.

Linguistic Features

Words and Register

As in Henry and Roseberry’s (2001) study, the most frequently used word was the pronoun ‘I’ (see *Table 8*), used as a cohesive device that “enables the genre as a whole to lack the wide range of cohesive ties found in other genres” (p. 160). The words my and your were frequently used, especially in the opening and closing stages (O, CA, EN, PE) to highlight relationships between the applicant and the position available. The conditional, would, was used 23 times, which can be seen as a politeness strategy, as it is considered more polite than want. There was significant use of the

Table 7 Organization of moves in each individual EJAL

No.	Field	Move pattern	Freq.
1	Science	O + CA + AD + EN + P + RA + P + A + PE + SO	1
2	Publishing	O + CA + RA + P + EN + PE + SO	1
3	Media	O + CA + AD + RA + P + AD + P + PE + SO	1
4	Healthcare	O + AD + CA + P + EN + PE + SO	1
5	Design	O + CA + AD + RA + P + EN + PE + SO	2
6	Sports Coaching	O + CA + AD + RA + P + RA + PE + SO	1
7	Journalism	O + RA + AD + P + RA + P + RA + P + EN + PE + SO	1
8	Retail	O + CA + AD + P + EN + PE + SO	1
9	Programming	O + CA + AD + EN + RA + P + PE + SO	1
10	Engineering	O + CA + AD + RA + P + EN + PE + SO	2
11	Social Work	O + CA + AD + P + PE + SO	1
12	Marketing	O + CA + RA + P + PE + SO	2
13	Management	O + CA + RA + P + PE + SO	2
14	Education	O + CA + AD + P + RA + PE + SO	1
15	Human Resources	O + AD + CA + P + PE + EN + PE + SO	1

Table 8 Frequently used bigrams and trigrams in total sample (15 EJALs)

Order of Frequency	Bigrams	Trigrams
1	I am	I am a
2	I have	for your time
3	as a	your time and
4	in the	thank you for
5	for your	you for your

word and, especially when highlighting skills and abilities, *promoting the candidate*. It was used in binary phrases, mainly with two nouns, but also with adjectives and verbs.

Table 9 lists the binary phrases used. Writers “most likely feel that binary expressions of this kind seem more complete, and provide an additional promotional opportunity”, making them common in promotional genres (Henry & Roseberry, 2001, p. 16). This means that from a pedagogical perspective, it is important to teach the use of binary expressions to language learners, to conform to the norms of EJALs, and provide validity to a prospective employer.

Table 10 lists other genre-specific words that appeared throughout the samples. Words such as position, experience, skills, and company may be considered genre-specific, highlighting the importance for teaching of such vocabularies to conform to the conventions of the genre. These words also tend to be move-specific, such as position being used when *offering candidature*, and experience occurring more when *promoting the candidate*. In addition, Table 11 shows examples of field-specific words.

Table 9 *Pairs of nouns, verbs and adjectives*

Nouns	Verbs	Adjectives
Studies and research	Developed and pitched	Exciting and challenging
Research and experience	Writing and editing	Critical and technical
Institution and pioneer	Play and grow	Strong and immediate
Skills and attention	Report and work	Deep and varied
Time and consideration	Publish and promote	News and feature
Teams and families	Developed and supported	Local and national
Questions and answers	Learning and applying	Dependable and dedicated
Skills and knowledge		Outgoing and friendly
Lesson and practices		Experienced and professional
Issues and events		Unique and enriching
Expertise and knowledge		Enthusiastic and strong
Awards and recognition		Trained and prepared
Organizer and planner		
Experience and education		
Education and career		
Systems and design		
Techniques and strategies		
Creativity and enthusiasm		
Content and ads		
Restaurants and bars		

Table 10 *Genre-specific words in the total sample (15 EJALS)*

Genre specific words	Occurrences in sample (15 EJALS)
Position	26
Experience	24
Would	23
Time	19
Work	15
Writing	14
Skills	14
Company	13
Team	13
College	12
Media	12
Strong	11
Social	10
Thank	11
Graduate	9
Opportunity	9
Resume	8
Apply	8

Table 11 Examples of field-specific words

Journalism	Retail	Marketing	Education
Communications	Volume	Social media	Teacher
Newspaper	Revenue	Industry	Qualifications
Report	Merchandising	Activation	Certificate
Advertised	Inventory	Research	Practicum
News	Cash	Conversion	Extra-curricular
Articles	Customer	Campaigns	Academically
Editor	Staffing	Technological	Interpersonal
Blogger	Satisfaction	Action plan	Student
Political	Business	Markets	Faculty
Newscasts	Employees	Platforms	Classroom
Website	Turnover	Strategies	
Covering	Shrinkage	Creative	
Readers	Budget	Analyzing	
Media		Consumer	
Digital			

Tense

As previously stated by Henry and Roseberry (2001), tense use was also move-specific. There was a clear correlation between the common linguistic features in each individual move. *Table 12* shows the tenses most used with each move, and the keywords used. This suggests that the function of each move determines the tense used to achieve it.

Pedagogical Priorities for Teaching EJALS to Japanese Learners

The genre analysis of English Job Application Letters highlights several areas that require focused instructional attention when teaching Japanese EFL learners. These areas reflect differences between Japanese domestic application practices and the rhetorical expectations of international business communication rather than general linguistic deficiencies.

The move *promoting the candidate* emerged as the most prominent and rhetorically flexible component of EJALS, accounting for the largest proportion of text in the corpus. This move poses a particular challenge for Japanese learners, whose prior experience with standardized résumé formats encourages factual listing and discourages explicit self-promotion. Instruction should therefore emphasize evidence-based promotion, guiding learners to connect skills and experience directly to job requirements and to articulate value through concise narrative elaboration.

The analysis also showed that *stating reasons for applying* frequently functions as a transition into promotional content, establishing alignment between the applicant and the organization. Learner writing, however, often relies on generic expressions of desire that fail to demonstrate audience awareness. Pedagogical focus is needed to help students frame motivation as rhetorical

Table 12 Common tense usage during selected moves

Moves	Common tenses	Keywords
Offering candidature (CA)	Present progressive	Interest, writing, express
<i>Referring to a job advertisement</i> (AD)	Simple past tense	Advertised, posted in, position
Stating reasons for applying (RA)	Present tenses	Apply, career, interest
Stating availability (A)	Future Progressive	Available, will be, a specific date
Promoting the candidate (P) - Listing relevant skills, and abilities. - Stating how skills and abilities were obtained. - Listing qualifications. - Naming present job. - Predicting success. - Demonstrating knowledge of target position.	Present simple tense Past tense narrative Simple past tense Present tenses Present tenses Present tenses	Experience, team, strong, have College, student, work, university Skills, graduate, have, completed Position, company, work Know, will, feel, enjoy, love Know, hear, company
Enclosing documents (EN)	Present perfect	Enclosed, resume, CV
Polite ending (PE) - Welcoming response. - Inviting favorable consideration. - Thanking. - Offering to provide further information.	Present tenses Present tenses Present tenses Present tenses	Join, excited, and look forward Will, arrange, consider, beneficial Thank you, consideration Email, phone, contact, discuss, speak, meet

alignment by referencing organizational characteristics and linking them to personal qualifications.

From a linguistic perspective, tense usage in EJALs was found to be strongly move-specific. For example, the present progressive commonly realizes *offering candidature*, while past tense and present perfect forms are used to describe prior experience when *promoting the candidate*. Japanese learners often approach tense as a purely grammatical issue, resulting in inconsistent usage. Teaching tense as a rhetorical resource tied to communicative function can improve both accuracy and appropriacy.

Register control represents another key instructional priority. The corpus revealed consistent use of formal lexico-grammatical patterns and indirect politeness strategies, particularly in closing moves. In contrast, learner texts often contain overly direct or colloquial expressions. Instruction should therefore focus on how politeness and professionalism in English business writing are achieved through modality and indirectness rather than honorific forms.

Finally, the move analysis demonstrated that EJALs allow flexibility in move selection and sequencing, with certain moves absent or repeated depending on context. Learners' reliance on rigid templates can inhibit effective writing. Explicit instruction in move optionality can help students understand genre conventions as adaptable resources rather than fixed rules, facilitating a shift from résumé-style reporting to rhetorical organization.

In summary, the genre analysis suggests that instruction in EJAL writing for Japanese learners should prioritize promotional self-presentation, audience alignment, move-specific tense usage, register control, and genre flexibility. Addressing these areas explicitly can support learners in adapting domestic application practices to international professional writing conventions.

Pedagogical Application

Building upon the findings, pilot study was used to investigate whether explicit genre-based instruction, informed by the move-step analysis of English Job Application Letters (EJALs), could support Japanese university students in producing more coherent and professionally appropriate cover letters. Specifically, the study aimed to:

1. Examine how students apply rhetorical moves and linguistic features (tense, register) after targeted instruction.
2. Explore the feasibility of integrating genre analysis findings into a Business English curriculum.
3. Provide preliminary evidence to inform larger-scale classroom interventions in the future.

Teaching English Job Application Letters Through Genre-Based Instruction

The intervention was conducted as an action research project in a Business English Communication class over one semester. A total of twelve undergraduate students, who belonged to various faculties, participated in weekly 90-minute lessons. The instructor, informed by the prior genre analysis, explicitly taught students the obligatory and optional moves (e.g., *Opening, offering candidature, promoting the candidate, polite ending*), along with associated tense usage and register. Instruction included scaffolded exercises: drafting introductions, practicing promotional language, and producing polite closings.

Due to ethical considerations, all students received the same treatment; therefore, no control group was used. Data for evaluation included classroom observations, written drafts, and final EJAL submissions.

Drawing on the findings of the genre analysis presented earlier, this section proposes a structured, genre-based instructional approach for teaching EJALs in a Business English or English

for Specific Purposes (ESP) context. The pedagogical framework is informed by established genre theory (Bhatia, 1993; Swales, 1990) and empirical studies of job application letters (Henry and Roseberry, 2001; Wang, 2005), and it aims to bridge the gap between descriptive genre analysis and classroom practice.

The instructional sequence is organized into five stages: genre awareness, explicit move instruction, linguistic realization, guided production, and independent performance. This allows learners to gradually internalize genre conventions while retaining flexibility in rhetorical choice.

Genre Awareness and Communicative Purpose

Instruction begins with explicit awareness-raising activities designed to foreground the communicative purpose of EJALs as a promotional genre. As demonstrated in the genre analysis, EJALs function primarily to persuade employers and elicit a specific response, typically an invitation to interview (Bhatia, 1993). This purpose contrasts sharply with Japanese application conventions, such as the *rirekisho*, which emphasize factual reporting and standardized formatting rather than rhetorical persuasion.

To address this contrast, learners are guided through comparative analysis of Japanese application documents and authentic English job application letters. This activity helps students recognize differences in audience orientation, degrees of self-promotion, and rhetorical expectations. Such contrastive awareness is particularly important in the Japanese EFL context, where cultural norms may discourage explicit self-evaluation (Furka, 2008; Upton and Connor, 2001). By framing EJALs as a genre governed by shared professional conventions rather than personal stylistic preference, students are better positioned to adopt unfamiliar rhetorical strategies.

Classroom Activities

<p>Contrastive Warm-Up</p> <p>Show:</p> <ul style="list-style-type: none">• A Japanese <i>rirekisho</i>• A professional English EJAL <p>Ask students:</p> <ul style="list-style-type: none">• <i>Which one persuades?</i>• <i>Which one only reports facts?</i> <p>Elicit the idea that EJALs are promotional genres.</p> <p>Outcome: Students understand why self-promotion is required.</p>

Explicit Instruction in Rhetorical Moves

Once students understand the communicative purpose of EJALs, instruction shifts to explicit teaching of rhetorical move structure. Students are introduced to both obligatory and optional

moves identified in the corpus, including opening, *offering candidature*, *promoting the candidate*, *polite ending*, and *signing off* (Henry and Roseberry, 2001).

Rather than presenting these moves as fixed templates, instruction emphasizes their functional roles. Students analyze authentic EJALs to identify move boundaries and sequencing, through labeling or color-coding activities. This analytical work reflects the recursive and flexible nature of move realization observed in the data (see *Table 7*) and reinforces the idea that moves may vary in length, order, and frequency depending on context (Dudley-Evans, 2000; Swales, 1990).

Through this process, learners develop an understanding of genre structure as purposeful and adaptive rather than rigid, an insight that is essential for effective professional writing.

Table 13 *Teaching Tool: Move Function Table*

Move	What this move does	Key question for students
Opening	Establishes relationship	Who am I writing to?
Referring to ad	Shows legitimacy	Where did I find this job?
Offering candidature	States purpose	Why am I writing?
Stating reasons	Aligns with company	Why this company?
Promoting candidate	Persuades employer	Why should they choose me?
Enclosing documents	Signals completeness	What else am I sending?
Polite ending	Invites response	What do I want next?
Signing off	Closes relationship	How do I end professionally?

Classroom Activities

1. Move Identification Task

- Give students an authentic EJAL.
- Color-code each move.
- Compare answers in pairs.

2. Move Scramble

- Provide a scrambled letter (moves out of order).
- Students reorder based on communicative logic.

Outcome: Students internalize rhetorical logic, not memorized formulas

Linking Moves to Linguistic Realizations

Building on move awareness, instruction then focuses on the relationship between rhetorical moves and linguistic features. As shown in the grammatical and lexical analysis, language choices in EJALs are strongly move-specific. For example, the present progressive tense frequently realizes the move *offering candidature*, while past tense narratives and present perfect constructions are commonly used when *promoting the candidate* through prior experience (Henry and Roseberry, 2001; Wang, 2005).

These findings are incorporated into classroom instruction through guided sentence analysis

and rewriting tasks, enabling students to observe how grammatical choices contribute to persuasion and professionalism. Attention is paid to register, as the corpus analysis demonstrated consistent use of formal lexico-grammatical patterns associated with business communication (Halliday and Hasan, 1989; Young, Fitzgerald, and Fitzgerald, 2018).

This stage encourages learners to move beyond a narrow focus on grammatical correctness and toward a functional understanding of how language realizes communicative intent within specific rhetorical contexts.

Classroom Activities

1. **Bad to Better Rewriting**

- Students improve sentences by:
 - adjusting tense
 - increasing specificity
 - improving register

2. **Promotion Sentence Builder**

- **Scaffold:** Skill + where/how developed + result
(e.g., *I developed strong analytical skills during my internship, where I assisted with market research and reporting.*)

Outcome: Students see how grammar serves persuasion, not correctness alone.

Guided Writing and Scaffolded Production

Following explicit instruction, students engage in scaffolded writing activities that mirror the staged development of genre competence. Rather than drafting complete EJALs immediately, learners compose individual moves sequentially, beginning with opening and *offering candidature* before progressing to promotional and closing moves.

This approach reflects the move-by-move analytical framework used in the study and allows instructors to provide targeted feedback on discourse function rather than surface-level errors. Peer review activities are structured around move checklists derived from the analytical model, encouraging students to evaluate whether each move fulfills its communicative purpose. Such scaffolding has been shown to be particularly effective for learners who struggle with self-promotion or rhetorical elaboration (Hyland, 2007).

Classroom Activities

Move-by-Move Practice

Practice structure:

- **Task 1:** *Opening + Offering candidature*
- **Task 2:** *Stating reasons + Promoting candidate*
- **Task 3:** *Polite ending + Full draft*

Each task includes:

- Model
- Checklist
- Peer feedback focused on moves, not grammar only

Checklist for *promoting candidate*:

- *Did I give evidence, not adjectives?*
- *Did I connect skills to job requirements?*
- *Did I use appropriate tense?*

Outcome: Students gain confidence and control, especially regarding self-promotion.

Independent Performance and Reflective Practice

In the final stage, students independently produce complete EJALs in response to authentic or simulated job advertisements. This task requires learners to select, sequence, and linguistically realize moves without explicit guidance, demonstrating internalization of genre knowledge.

Reflective activities accompany this stage, prompting students to consider how genre awareness influenced their writing choices and how EJAL conventions differ from domestic résumé practices. Reflection reinforces the pedagogical goal of developing transferable genre competence rather than formulaic reproduction.

Overall, this instructional sequence demonstrates how genre analysis can be operationalized in ESP classrooms. By making rhetorical expectations explicit and linking structural moves to linguistic realizations, genre-based pedagogy enables learners to develop both pragmatic competence and confidence in professional writing. As illustrated by the classroom application reported in this study, such an approach is particularly well suited to Japanese EFL contexts, where differences between domestic and international application practices are often implicit and require explicit pedagogical mediation.

Final Task

Create
Write a complete EJAL for:

- a real job ad
- or a simulated international company

Reflection
Short guided reflection:

- *What felt uncomfortable?*
- *How is this different from Japanese style?*
- *How did the move structure help?*

Outcome: Students understand EJAL writing as strategic communication.

Learner Outcome

Students were able to internalize the rhetorical moves and apply them to their own writing with relative ease. Most produced cover letters that displayed clear structural organization, appropriate use of promotional language, and awareness of business register. The framework simplified lesson delivery, allowing the teacher to give precise feedback on target language, grammar, and discourse functions. Informal evaluation of student texts showed that the majority successfully employed obligatory moves and genre-specific vocabulary identified in the original study.

Classroom Observations

During the drafting stage, several students initially struggled with *Stating reasons for applying*, often producing vague statements such as “*I want to work at your company*”. However, after explicit instruction and examples, students began to elaborate more persuasively, writing phrases like “*I am particularly motivated to join your company because of its strong record in global marketing and its commitment to innovation.*” In addition, many drafts reflected *rirekisho*-style factual listing, whereas final versions integrated qualifications and skills with persuasive detail.

In peer-review sessions, students noticed each other’s misuse of tense. One student commented: “*You wrote ‘I work as an intern last year,’ but it should be past tense, ‘I worked.’*” This demonstrated growing awareness of the move-specific tense choices highlighted in the lessons.

In addition to classroom observations, informal student responses were collected after classroom activities as a supplementary means of monitoring students’ comfort with the instructional approach. The students appeared to be satisfied with the instruction, and no negative comments were given.

Sample Student Responses

- “Before this class, I didn’t know how to start or finish a cover letter. Now I can use expressions like ‘I look forward to discussing my application with you,’ which makes my letter more professional.”
- I realized the cover letter is not only about me, but about showing I understand what the company wants.”
- “It was difficult to ‘promote myself’ because in Japanese style we don’t write like that, but using the moves made it easier to organize.”

Improvements in Student Writing

Comparison of early drafts and final submissions showed clear development. For example:

- Early drafts often lacked *polite ending*, but all final versions included it.
- The frequency of vague self-descriptions (“I am hardworking”) decreased, replaced by more specific claims (“I specialized in international marketing and developed analytical and teamwork skills”).
- Register improved: colloquial phrases such as “I really want this job” were replaced by more formal alternatives like “I am eager to contribute my skills to your international team.”

Table 14 Study Example: Student A - Draft vs. Final Version

Moves	Draft Excerpt	Final Excerpt	Outcome
Opening / Referring to Job Advertisement	“I write this letter to apply job at your company.”	“I am writing to apply for the position of Marketing Assistant, as advertised on your company website.”	Improved accuracy, clearer reference to solicited position, adoption of present progressive.
Promoting Candidature	“I am hardworking and like to talk with people.”	“I have developed strong communication and problem-solving skills during my internship, where I assisted with client outreach and project coordination.”	Shift from vague self-description to concrete, specific qualifications.
Closing / Polite Ending	“Please hire me. I wait your answer.”	“I look forward to the opportunity to discuss how my skills and experiences can contribute to your team. Thank you for considering my application.”	Register upgraded, polite and professional tone achieved.

Closer examination of the revisions between early drafts and final submissions reveals that the observed improvements were not merely linguistic but fundamentally rhetorical in nature. In particular, students demonstrated increased control over genre-specific moves, greater awareness of audience expectations, and more appropriate deployment of linguistic resources to fulfill communicative purposes.

At the structural level, the consistent inclusion of a *polite ending* in final drafts indicates improved awareness of obligatory genre moves. While early drafts frequently omitted this move, its systematic presence in final versions suggests that students had internalized the schematic structure of EJALs and recognized the importance of explicitly inviting further contact. This shift reflects growing genre awareness rather than rote imitation, as students employed varied yet functionally appropriate realizations of the move.

Improvements were also evident in the realization of the *promoting the candidate* move. Early drafts tended to rely on vague, evaluative self-descriptions (e.g., *I am hardworking*), which provide limited persuasive value. In contrast, final submissions showed a clear move toward evidence-based promotion, with students grounding claims in concrete experience and contextual detail. This change suggests that learners had begun to reconceptualize self-promotion as a genre-specific communicative requirement rather than a personal assertion, an adjustment that is particularly significant in the Japanese context, where explicit self-evaluation is often discouraged.

In addition, the revisions demonstrate enhanced control over register and grammatical choices. Colloquial or overly direct expressions present in early drafts were replaced with more formal, indirect constructions typical of professional business correspondence. Notably, students also demonstrated more accurate use of move-specific tense patterns, such as the adoption of the present progressive in openings (e.g., *I am writing to apply...*) and the use of past tense narratives when describing prior experience. These changes indicate that students were beginning to align grammatical form with rhetorical function, rather than treating grammar as an isolated concern.

Taken together, these improvements suggest that genre-based instruction supported students in developing a more sophisticated understanding of EJALs as persuasive texts. Rather than simply correcting surface-level errors, students learned to select and organize content strategically, adapt language to audience expectations, and realize rhetorical moves more effectively. Although the scope of the study does not allow for strong generalization, the observed revisions provide qualitative evidence that explicit instruction in genre structure and linguistic realization can lead to meaningful improvements in professional writing performance.

Conclusion

This paper has demonstrated both the analytical and pedagogical value of a genre-based

approach to English Job Application Letters (EJALs). The genre analysis revealed recurring rhetorical moves, step structures, and linguistic features that characterize effective cover letters. These insights not only contribute to the broader body of research in English for Specific Purposes (ESP) but also provide practical guidance for teaching writing in professional contexts.

The classroom application further illustrates the utility of such findings. By explicitly teaching moves, steps, and language features derived from the analysis, students were able to produce coherent, persuasive cover letters that aligned with the conventions of international business communication. Importantly, the study highlights how genre analysis can bridge the gap between theoretical research and classroom practice, equipping students with concrete strategies for producing texts that meet real-world professional expectations.

However, the reflection on the teaching context also underlines certain limitations. The absence of a control group and reliance on a small sample size ($N=12$) restricted the ability to generalize these results. In addition, the study relied primarily on teacher observation and student writing samples rather than standardized assessment tools. Future research should aim to incorporate larger and more diverse cohorts, along with pre- and post-assessment measures, in order to provide stronger empirical evidence of learning gains. At the same time, this study raises an ethical consideration often overlooked in applied linguistics research: the fairness of instruction. By teaching all students the same strategies rather than withholding them from a “control group,” the teacher ensured equitable learning opportunities, reinforcing the practical and moral responsibilities of ESP practitioners.

Despite the limitations, the pilot study provides preliminary evidence that genre-based pedagogy is effective in Business English instruction for Japanese learners. It demonstrates how research findings from genre analysis can be directly translated into classroom practice, helping students bridge the gap between domestic résumé conventions (*rirekisho*) and international expectations for EJALs. For practitioners, this suggests that even small-scale interventions can yield significant benefits in students’ ability to produce professional business writing.

Ultimately, this study reinforces the idea that genre-based pedagogy is not merely an abstract analytical tool but a highly effective teaching framework. For Japanese graduates preparing to enter globalized work environments, such explicit instruction helps demystify the expectations of international employers, builds confidence in professional writing, and improves employability. More broadly, it shows that genre analysis, when directly applied to pedagogy, has the potential to transform ESP classrooms into spaces where students can develop both linguistic accuracy and pragmatic competence. By connecting research-driven insights with classroom practice, practitioners can foster deeper learner engagement, greater clarity in writing, and more authentic preparation for the communicative demands of the global workplace.

References

- Al-Ali, M. N. (2004). How to get yourself on the door of a job: A cross-cultural contrastive study of Arabic and English job application letters. *Journal of Multilingual and Multicultural Development*, 25(1), 1–23.
- Allison, J., Townsend, J., & Emmerson, P. (2008). *The business*. Macmillan Education.
- Baugh, L. S. (1991). *Handbook for practical letter writing*. NTC Publishing Group.
- Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Longman.
- Bruce, I. (2009). Results sections in sociology and organic chemistry articles: A genre analysis. *English for Specific Purposes*, 28, 105–124.
- Butt, D., Fahey, R., Feez, R., Sprinks, S., & Yallop, C. (2000). *Using functional grammar: An explorer's guide*. Macquarie University.
- Connor, U., Davis, K., & De Rycker, T. (1995). Correctness and clarity in applying for overseas jobs: A cross-cultural analysis of U.S. and Flemish applications. *Text*, 15(4), 457–476.
- Connor, U., & Mauranen, A. (1999). Linguistic analysis of grant proposals: European Union research grants. *English for Specific Purposes*, 18(1), 47–62.
- Connor, U., Precht, K., & Upton, T. (2002). Business English: Learner data from Belgium, Finland, and the US. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition, and foreign language teaching* (pp. 175–194). John Benjamins.
- Cross, C., & Oppenheim, C. (2006). A genre analysis of scientific abstracts. *Journal of Documentation*, 62, 428–446.
- Ding, H. (2007). Genre analysis of personal statements. *English for Specific Purposes*, 26, 368–392.
- Dudley-Evans, T. (1987). Genre analysis and ESP. *ELR Journal*, 1.
- Dudley-Evans, T. (2000). Genre analysis: A key to a theory of ESP? *Ibérica*, 2, 3–11.
- Faiza, S. (2012). *Raising ESP students' awareness of the generic structures and linguistic features of job application letters through genre-based instruction at the tertiary level: The case of third-year management students*. Djillali Liabes University.
- Fung, S. G. (2014). Genre analysis: Analysis of moves in academic job application letters. *Singapore Management Journal*, 3(1), 32–49.
- Furka, I. (2008). The curriculum vitae and the motivational letter: A rhetorical and cultural analysis. *WoPaLP*, 2, 18–37.
- Halliday, M. A. K., & Hasan, R. (1989). *Introduction to functional grammar*. Edward Arnold.
- Henry, A., & Roseberry, R. L. (2001). A narrow-angled corpus analysis of moves and strategies of the genre “letter of application.” *English for Specific Purposes*, 20, 153–167.
- Khan, A., & Tan, B. T. (2012). Generic patterns in application letters: The context of Pakistan. *RELC Journal*, 43(3), 393–410.
- Martin, P. M. (2003). A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes*, 22, 25–43.
- Mustaffa, R., Stapa, S. D., Darus, S., & Maasum, T. (2007). Organizational structure of job application letters by Malaysian graduates. In *Proceedings of the SOLLs INTEC 2007 Conference* (pp. 489–496).
- Shamsudin, S., Sarjit Singh, M. K., & Zaid, Y. H. (2012). Revisiting genre analysis: Applying Vijay

- Bhatia's approach. *Procedia – Social and Behavioral Sciences*, 66, 370–379.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Upton, T. A., & Connor, U. (2001). Using computerized corpus analysis to investigate the text-linguistic discourse moves of a genre. *English for Specific Purposes*, 20(4), 313–329.
- Wang, H. (2005). A pragmatic genre analysis of job application letters. *Peer Review Journal*, 2(12), 76–81.
- Wijayanti, S. H. (2017). The moves of Indonesian application letters. *Lingua Cultura*, 11(1), 39–45.
- Young, L., Fitzgerald, M., & Fitzgerald, S. (2018). *The power of language*. Equinox Publishing.

From Textbook to Test: Vocabulary Frequency Analysis to AI-Generated Quizzes

Timothy Ang

Toyo University, Center for Global Education and Exchange

Katsuichiro 'Ken' Ohashi

Rikkyo University, Center for Foreign Language and Research (FLER)

(At time of research Toyo University, Center for Global Education and Exchange)

Abstract

This paper outlines a two-stage approach to vocabulary selection and assessment for an International English Language Testing System (IELTS) preparation course within a Japanese university context. In stage 1, target words from twelve textbook passages were identified by profiling vocabulary against the frequency based lists New General Service List (NGSL) and New Academic Word List (NAWL) using LEXTUTOR. In stage 2, multiple choice quizzes were created using the identified target words from stage 1 and through use of a large language model (LLM). The study investigates both the effectiveness and limitations of using vocabulary profilers and LLMs. Our initial findings suggest that profiling vocabulary offers a theoretical foundation and framework to which AI-assisted methods can reduce relative teacher workload for vocabulary assessment.

Introduction

Vocabulary is fundamental for success in standardized English proficiency tests such as the IELTS. Language learners are aware of this, but due to academic demands of both English and non-English coursework, it is understandable that the learners encounter issues with study strategies such as time management and test preparation. The English language teacher can play an important role in reducing the language learner's time needed to study a language by guiding the learners' attention to words that the learners are most likely to encounter in multiple situations.

For the English language teacher in higher education that lacks theoretical knowledge in second language vocabulary studies, making decisions on what vocabulary they should guide their learners' attention to can be challenging. This is especially true when designing courses around materials that do not provide a list of target vocabulary. The words should be relevant to the context

that is being taught, while not being so content specific that the words will not be very useful in context outside of a given lesson. As such, for low to intermediate proficiency level learners, the words should be on the NGSL and NAWL for efficiently learning words that the learners are most likely to encounter (Browne, 2013). The teacher should also consider words that appear multiple times within the course since repeated exposure is known to be effective for vocabulary acquisition (Rott, 1999; Waring and Takaki, 2003).

Another issue for teachers, whether well versed in second language vocabulary theories or not, is that they may not be aware of openly available tools that can be used to create high-quality vocabulary lists and assessments. Vocabulary profiling technology such as LEXTUTOR *Complete Web-VP v.2.6* (Cobb), combined with the use of formulas on spreadsheet software such as Google Sheets or Microsoft Excel can be a viable solution for teachers to choose and assign target vocabulary for a given course over a semester.

High-quality assessment design in ESL contexts often faces challenges, particularly when instructors must develop additional materials without research-informed templates or textbook supplement material to draw from.

Research has shown that assessments can be prone to various deficiencies. These include potential fairness issues that arise when test tasks introduce factors unrelated to the skill being tested (Xi, 2010) or uneven skill profiles that reflect gaps between instructional focus and test demands (Sasaki, 2022). Gottlieb (2016) cautions that assessments may become linguistically inappropriate for English learners when test items contain complex structures (i.e., unnecessarily complicated instructions or overly dense question wording), thereby obscuring what is actually being measured. Our own departmental faculty experiences revealed issues with biased word selection, questions whose themes were distant or irrelevant in an IELTS context, and misaligned or inconsistent difficulty level of assessment questions.

This tenuous standardization in instructor-created quizzes reinforced the importance of the need for systematic design and planning to create assessments. This study reports on a dual-stage initiative conducted in the LEAP (Language Education for Academic Purposes) program at Toyo University, targeting Japanese university students whose English is at beginner to lower-intermediate levels, corresponding roughly to the Common European Framework of Reference for Languages (CEFR) A2–B1 (IELTS 3.5–5.0). Combining vocabulary profiling and LLMs remains underexplored, especially in test preparation contexts such as IELTS, where lexical range, frequency, and depth must align with exam expectations while maintaining balance with the wider aims of the course or program.

The first stage employed vocabulary profiling tools to construct targeted word lists from IELTS Trainer 2 (Cambridge University Press, 2019), the main textbook used in the course. The second stage involved generating quizzes using a LLM, with implementation supported by Google

Forms and Google Classroom.

Literature Review

Identifying Target Vocabulary

Past studies (Laufer, 1989) have indicated that the language learner would need to know at least 95% of the words in a given text to comprehend it. Van Zeeland and Schmit (2013) suggest that listening comprehension can be achieved with varying levels of success from 90% word coverage of the content. Taking this into consideration, it is reasonable to understand why frequency-based word lists have been compiled and used to identify words that are commonly used by native English speakers.

Among these vocabulary compilations, The General Service List (West, 1953) and Academic Word List (Coxhead, 2000) have been used by many researchers and English language educators over the years to analyze texts, create textbooks and graded readers, create tests such as the Vocabulary Levels Test, Vocabulary Size Test (Nation and Beglar, 2007), and Computer Adaptive Test of Size and Strength (Laufer, & Goldstein, 2004). Researchers have used word frequency bands of 1,000s to categorize and identify which words language learners of English should focus on acquiring.

It is said that the language learner would be able to identify around 80% of the words in any given text if they know the words in the first and second frequency bands, that is, words that are a member of the 2000 most frequently used word families (Waring & Nation, 1997). Coxhead's (2000) academic word list is a collection of 570 word families that appear in academic texts at a higher rate than in non-academic texts. Mastering these words would raise the word recognition rate close to 90% when reading academic texts, which is very beneficial for students in higher education (Coxhead, 2000).

Despite its high usage by researchers, the General Service List has been criticized for its text coverage rate for modern literature and using a corpus that can be considered limited in size compared to more modern ones (Browne, 2014). In 2013, Brown announced the New General Service List (NGSL) and New Academic Word List (NAWL) as an updated version of the GSL (West, 1953) and AWL (Coxhead, 2000). The NGSL and NAWL together provide close to a 90% coverage for abstracts in academic texts across 12 subject matter divisions of the American Educational Research Association (Hendry & Sheepy, 2018).

Vocabulary profilers have been used by researchers to critique EFL textbooks. Klinger (2004) examined six textbooks used at Japanese elementary and junior high schools and compared the vocabulary used in these books against the JACET8000 wordlist and found that 95% of the words used in these textbooks can be read by learning the first 3,000 words on the list. Nakayama (2022a,

2022b) looked at textbooks used at junior and senior high schools in Japan. He points out that roughly 95 % of the tokens used in the textbooks he examined were covered by the NGSL while the coverage is lower than 92% for senior high school textbooks. Despite its use in vocabulary research, documentation on the use of vocabulary profilers to identify vocabulary for instruction, especially studies that use NGSL and NAWL as the benchmark for vocabulary difficulty are difficult to locate.

Using Large Language Models to Create Tests

Large language models (LLMs) are artificial intelligence systems trained on large datasets to generate human-like language. In this study, we used ChatGPT (GPT-4o; OpenAI, 2024), a publicly available LLM released in 2022. GPT-4o is available in both free and paid tiers and was selected for its accessibility and capacity to generate structured assessment materials. LLMs are typically accessed through a graphical chatbot-style interface on a website or application, which allows users to interact with the model by entering prompts and receiving generated responses.

LLMs in relation to artificial intelligence can be a confusing connection as it is merely utilizing computing power on probabilistic outcomes of existing data. They do not possess understanding in a human sense but instead produce output by predicting likely word sequences based on statistical patterns in language. Therefore the quality and relevance of LLM-generated content depend heavily on two factors: the existing data the LLM has accumulated and user-provided instructions known as prompts. The careful design and testing is referred to as prompt engineering.

The rise and prevalence of large language models (LLMs), introduced new opportunities for teaching materials in education. Teaching materials refer to instructor-selected or created resources used to support learning, including quizzes, worksheets, lesson slides, and class activities. The underlying assumption is that LLMs can generate curriculum-relevant content, thereby supporting curriculum development. Early classroom applications suggest that while LLMs can assist in prompt creation, quiz writing, and feedback generation, they also pose practical and ethical challenges that require careful oversight (Yan et al., 2024). For this paper we focused on the practical uses and challenges.

Although the IELTS does not include a dedicated vocabulary section, vocabulary knowledge is an essential learning component to answering reading and listening multiple-choice tasks. Students who cannot recognize variations in vocabulary are prone to choosing the wrong option even when they understand the main idea of the reading passages, which questions are based on. Chen (2019) found that less proficient learners made higher rates of usage errors when dealing with paraphrases. Vocabulary quizzes can help this weakness by training learners to recognize synonyms, collocations, and paraphrases commonly found in IELTS questions.

It is worth noting that vocabulary assessments represent only one component of lexical retention. Stewart (2024) cautions that recognition-based vocabulary tasks, such as multiple-choice items, may not always provide a complete measure of productive word knowledge. However, in the context of IELTS preparation, recognition tasks remain pedagogically relevant. IELTS reading and listening tasks often require learners to discriminate between near-synonyms, attend to nuanced lexical distinctions, and identify and eliminate distractors under time pressure. In this sense, well-designed multiple-choice vocabulary quizzes may help students practice essential word discrimination skills required in IELTS.

Research Questions:

RQ1: How can vocabulary profilers be used to identify vocabulary that should be prioritized in an IELTS preparation course?

RQ2: How effective are large language models (LLMs) in vocabulary assessment preparation work in terms of productivity gains and test quality output?

RQ3: What are the limitations of using lexical profilers and LLMs for creating vocabulary quizzes?

Methodology

Stage 1: Vocabulary Selection

This first stage concerns the procedures that were implemented when selecting the target vocabulary for instruction for the IELTS preparation course. How LEXTUTOR, Microsoft Word and Excel was used to select the 180 target words from the 10,320 words that appeared in the 12 reading passages from *IELTS Trainer Academic* (Cambridge University Press, 2019) is described.

Phase 1: Preparing Data

As part of the course update, *IELTS Trainer Academic* (Cambridge University Press, 2019) was chosen as the new textbook for the level 1 courses at LEAP. This book is a test book containing six full practice tests. Each test contains three reading passages for a total of 18. IELTS reading tests are designed so passage 1 is the easiest and passage 3 is the most difficult within the test. Level 1 being the lower level within the LEAP program with a good number of students' English proficiencies being at the CERF A1-A2 level, a decision was made to use Passage 1 and 2 from each test as teaching material to be covered during the 15-week semester.

The 12 passages were prepared so the words could be run through a vocabulary profiler. The publisher was generous to provide a digital copy of the textbook that could be used by the teachers. Texts from each of the passages of this digital version were copied onto separate Microsoft Word documents. Microsoft Word was used to speed up the process of locating words misidentified by

the optical character recognition (OCR) technology. Each word that was marked with a red wavy line that indicates a spelling error was checked and corrected.

Phase 2: Identifying Frequency and Range

In order to identify which words from the textbook the teachers should bring to their students' attention, measures were taken to identify words of the appropriate difficulty and importance. Appropriate difficulty was considered as words that have a frequency higher than 1000 on the NGSL and words found on the NAWL. The 1000 most frequently used words were excluded since these words were considered as words all university students should know already. Importance was determined as words that appear in multiple texts within the 12 chosen passages so students will receive exposure to the target words multiple times over the semester to maximize opportunities for uptake.

The 12 passages were run through LEXTUTOR (Cobb) separately to be compared against the NGSL and NAWL. Words that were categorized as 2 (frequency of 1,001-2,000 on the NGSL), 3 (frequency of 2,001-2,800 on the NGSL) and A (New Academic Word List) were extracted onto spreadsheets on Microsoft Excel. The words were placed in individual cells with 10 words in a row for easy counting. Different color fonts were used for each frequency band for easy recognition of the word's frequency. Words from each passage were placed on separate tabs that were renamed to easily understand which passage the words were from (e.g. 'T1P1' for test 1, passage 1). Each of the 12 tabs contained roughly 90-120 candidate words.

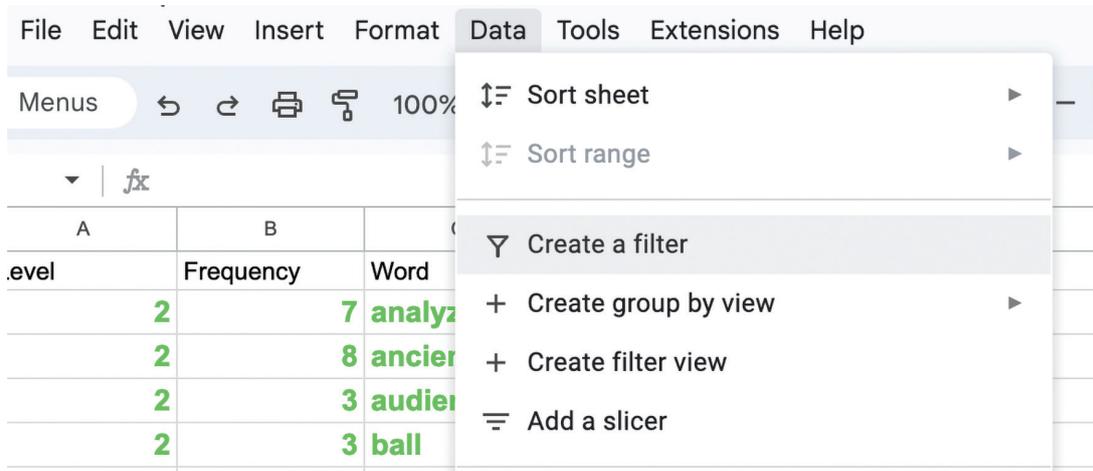
A combined list of all words from the 12 tabs were run through LEXTUTOR (Cobb) together to grasp the range of each word. If LEXTUTOR identified a word's frequency as 4 with this combined wordlist, this would translate as the word appearing in 4 different passages. The words were placed on a separate tab labeled Overall in the same Microsoft Excel file mentioned previously. All words were placed in a single column with the range listed in the cell directly to the right of the word. Each word was color coded to match the colors used on the 12 other tabs for individual passages. The accumulated frequency of each word was also added to this tab, on the same row as the word in a column labeled total frequency. This statistic was also acquired when running the combined file of all 12-passages, a corpus of 10,340 words, through LEXTUTOR.

Phase 3: Finalizing Word Lists Using Conditional Formatting

Candidate words on each tab were shortlisted using Microsoft Excel's conditional formatting function through a two-step process. The first step of isolating words with a high range score was conducted on the *Overall* tab. Words that appear on more than 3 passages were isolated using the filter function (Figure 1). All words that remained after filtering, along with the information for range and frequency, were copied to each of the 12 tabs at a location that did not interfere with the 90-120 candidate words on each tab.

Figure 1

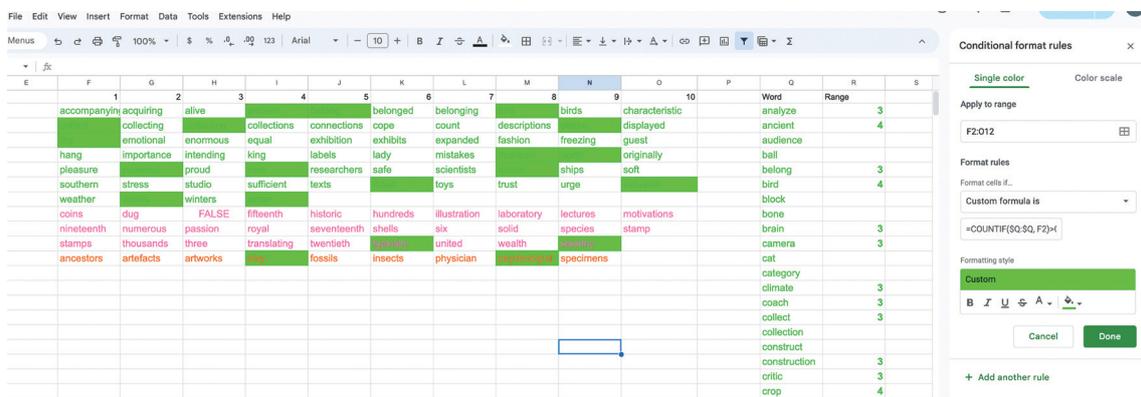
Screenshot of filter creating on Microsoft Excel



The second step was to locate the high frequency words that were identified in the previous step within the 90-120 candidate words on each tab by using a custom formula, “=COUNTIF(\$Q:\$Q, F2)>1” as the conditional format rule. This formula identifies duplicate words within a set area. “\$Q\$Q” specifies the column where the filtered words from the previous step are copied. “F2” is the first cell of the candidate words. As a result, approximately 20 words were highlighted for each passage (Figure 2). It is worthy to note that since words with a range score of 3 were used, each word was highlighted on more than three tabs.

Figure 2

Screenshot of highlighted candidate words after applying conditional formatting on Microsoft Excel



To standardize the difficulty of the 12 lists of 15-words, a decision was made to include 9 words from band 2, 3 words from band 3 and 3 words from band A, of the LEXTUTOR (Cobb) vocabulary profile results, on each list. Words with high range scores within each band were

prioritized in the selection process. Total frequency was the second criteria that was consulted in the selection process.

The shortlisted words on each tab were placed in a single table on the overall tab across 12 columns, one column of roughly 20 words for each passage. Microsoft Excel's conditional formatting formula “=COUNTIF(\$C\$2:C2,C2)=1” was used to identify the first appearance of each word. C2 represents the first cell of the shortlisted words.

As a general rule, each word was to be placed on a list at the earliest possible time in the semester. For example, the word *ancient* appears in four passages, test 1 passage 1, test 3 passage 2, test 4 passage 1, and test 5 passage 1. Out of the four passages, test 1 passage 1 is the passage scheduled to be used first within the semester, therefore the word *ancient* would be placed on the vocabulary list of 15 words for test 1 passage 1. This is to maximize the opportunities for students to encounter the target words consciously. An exception was made to this rule for some words that appeared in more than 5 passages. These were not placed on the list of first appearances to disperse the candidate words more evenly over the semester and still provide students with multiple opportunities of exposure to the word in different passages later in the semester.

Stage 2: Vocabulary Assessment

The second stage involved creating vocabulary assessments from the refined vocabulary list using both manual / human and LLM-assisted methods. Structured prompts were generated through ChatGPT (GPT-4). Twelve vocabulary quizzes were developed to correspond to the twelve textbook passages in *IELTS Trainer 2* (Cambridge University Press, 2019), with the first six created manually by one instructor and the remaining six produced with LLM assistance. This distribution was determined in advance to enable a systematic comparison between human-authored and LLM-assisted quiz construction.

Phase 1: Generating Quiz Questions Using LLM

The first step involved sourcing vocabulary from the master word list file compiled in Stage 1 (see Figure 3). This file served as the data source for creating 15-word subsets aligned with each passage from the course textbook. The subset list of vocabulary for Passage T1P1 was used and labelled as Quiz 1. These subset lists were then uploaded to Google Classroom separately so that students were expected to study all 15 words for each passage.

Figure 3

Screenshot of master word list file

	Quiz 1	Quiz 2	Quiz 3	Quiz 4	Quiz 5	Quiz 6	Quiz 7	Quiz 8	Quiz 9	Quiz 10	Quiz 11	Quiz 12
	T1P1	T1P2	T2P1	T2P2	T3P1	T3P2	T4P1	T4P2	T5P1	T5P2	T6P1	T6P2
1	ancient	audience	bone	brain	experiment	critic	climate	distance	block	coach	decade	crop
2	belong	camera	construction	cycle	global	eventually	directly	hardly	column	intelligence	hole	dry
3	collection	category	left	obvious	highly	frequently	upper	none	heat	implication	lake	engineering
4	museum	ideal	native	powerful	increasingly	professor	warm	outcome	practical	predict	consequence	fuel
5	secret	trend	river	rare	largely	scientific	burn	notion	proportion	objective	decline	output
6	writer	truth	stone	regular	ordinary	spread	farmer	suspect	component	analyze	observe	plenty
7	lady	platform	wood	survive	valuable	ticket	previously	apparently	construct	context	ought	tough
8	text	supporter	researcher	advice	widely	emotion	visitor	convince	engineer	enhance	threat	cheap
9	wealthy	digital	insight	device	emphasize	scientist	concrete	desire	expose	fan	sink	fruit
10	dig	festival	craft	recommend	reject	remarkable	layer	expert	heavy	football	agriculture	seed
11	stamp	historian	impact	motivation	technique	genuine	wooden	progress	sheet	speed	irrigation	vehicle
12	clay	maker	weave	alarm	flexible	laughter	everyday	found	apartment	league	drain	innovation
13	specimen	transform	tribe	rapid	stream	humor	roof	typically	steel	statistic	kilometer	extract
14	artifact	cinema	dye	hormone	influential	psychologist	pipe	honest	tower	algorithm	plug	fossil
15	artwork	depict	fiber	incredibly	limb	contradictory	outer	interestingly	architect	artificial	shallow	crude

From each 15-word list, ten target vocabulary items were manually selected as the correct answers for the quiz. Five of these came from the current passage’s list, while the remaining five were drawn from earlier quizzes to reinforce retention. This design required students to recall previously studied vocabulary rather than relying only on the most recent items, which provided additional challenge.

The finalized target words, together with their answer choices, were then fed into an LLM prompt. For each quiz, the correct answers were manually selected from the relevant 15-word subset list (e.g., *decade*, *decline*, *sink*, *drain*, *shallow*, *statistic*, *expose*, *outer*, *spread*, *incredibly*). Distractor options were drawn from previously studied subset lists that had already been quizzed on, ensuring cumulative review. As the semester progressed, this distractor pool expanded naturally, allowing later quizzes to incorporate a broader range of prior vocabulary items.

The standardized prompt format used to generate all LLM-based quizzes is shown below. In the reproduced prompt, square brackets (e.g., [manually selected target vocabulary list], [manually selected distractor pool]) indicate where the researcher inserted the vocabulary selected in the procedures previously described.

LLM Prompt

1. Task Overview

- Create a 10-question multiple-choice fill-in-the-blank quiz using the provided word lists
- Each question must have one correct answer and three distractors.

2. Correct Answers

- The 10 correct answers must be exactly these words: [manually selected vocabulary list]
- Each correct answer must be used once only.
- Correct answers must be inserted in a randomized order, not matching the sequence in which they are listed here, and must not appear consecutively or in a block (e.g., Q1–Q5). Ensure they are spread across the full set of 10 questions.

3. Distractors

- All distractors must come from the provided word list: [manually selected distractor pool]
- Distractors cannot repeat within the same quiz.
- Distractors should normally match the part of speech of the correct answer, but up to three per quiz may deliberately differ if they remain grammatically plausible in the sentence.
- Distractors must be grammatically correct and semantically possible in the sentence, but clearly wrong because they shift meaning along a different dimension (e.g., speed vs. significance, frequency vs. degree).
- Distractors should be close in meaning, academic in tone, or thematically related, creating an IELTS-style challenge.
- Avoid distractors that are clearly irrelevant (e.g., “ticket” in a science context) or overused examples.
- Each quiz should contain at least one “productive trap” distractor (such as *audience vs. critic* or *trend vs. progress*), but no more than three.

Special Parts of Speech Rules

- Adverbs: When the correct answer is an adverb, all distractors must also be adverbs from the provided list (e.g., *highly, frequently, widely, largely, increasingly*).
- Adjectives: When the correct answer is an adjective, distractors should also be adjectives (e.g., *ordinary, remarkable, rare*) and should test subtle differences in degree, intensity, or evaluation.
- Nouns: When the correct answer is a noun, distractors should preferably come from the same semantic category (e.g., *trend vs. progress, audience vs. critic, truth vs. notion*). However, distractors may also include adjectives or other descriptors (e.g., *native, wealthy*) if they are grammatically plausible in the sentence but represent a wrong dimension of meaning.
- Verbs: When the correct answer is a verb, distractors should be verbs with related but incorrect meanings (e.g., *belong, spread, survive*), making them plausible in structure but inaccurate in meaning.

4. Question Format

- Each sentence must reflect IELTS-style content (themes from Reading, Listening, Writing, or Speaking).
- Use authentic, natural English suitable for IELTS Band 6–7 level.
- Each item must allow only one grammatically and semantically correct answer.

5. Output

- Provide clean, unformatted plain text suitable for Google Forms.
- Do not use bullet points or letter labels for the choices.

- For each question, present the stem followed by four answer options listed one per line.
- At the end, provide a clearly labeled Answer Key in this format:

Answer Key

1. correct answer

...

10. correct answer

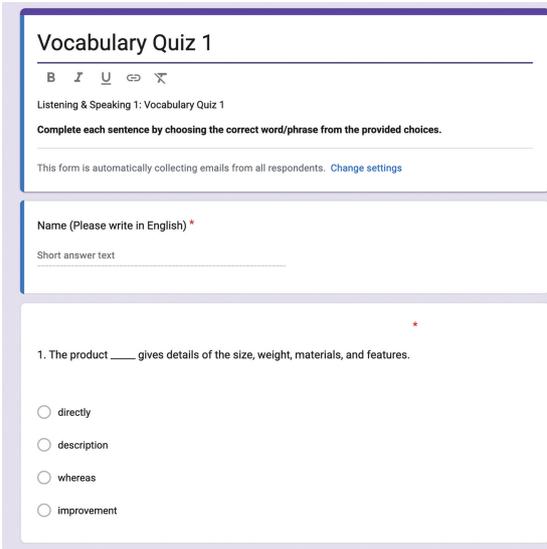
Phase 2: Google Forms Creation

The quiz questions generated in Phase 1 were transferred into Google Forms. During this phase, some answer choices were refined, with weaker distractors replaced by alternatives selected from the master word list. Once finalized, the multiple-choice fill-in-the-blank quizzes were published and shared via Google Drive for distribution to instructors, who then linked the quizzes to their individual Google Classroom pages (Figure 4), ensuring that students could access the material directly. This workflow also enabled a direct comparison of human and LLM-based quiz construction in terms of preparation time.

In the manually created vocabulary quizzes, decisions about distractor refinement were based on instructor judgment rather than formal statistical criteria. At the time of quiz construction, distractors were evaluated intuitively for grammatical fit, semantic plausibility, and perceived difficulty relative to the target vocabulary item. The length and complexity of the fill-in-the-blank stems were not standardized and were determined by instructor judgment during item construction.

Figure 4

Screenshot of Google Form integration



The screenshot shows a Google Form titled "Vocabulary Quiz 1". At the top, there are formatting icons (B, I, U, link, unlink) and the text "Listening & Speaking 1: Vocabulary Quiz 1". Below this, it says "Complete each sentence by choosing the correct word/phrase from the provided choices." and "This form is automatically collecting emails from all respondents. Change settings". There is a required text field for "Name (Please write in English)". Below that is a question: "1. The product ____ gives details of the size, weight, materials, and features." with four radio button options: "directly", "description", "whereas", and "improvement".

Phase 3: Administering Quiz

A one to two-week gap was implemented between the pre-quiz exposure (announcement of quiz and reminder to students to study) and the actual quiz to minimize short-term recall effects and ensure that performance reflected meaningful vocabulary acquisition. Quizzes were administered every other lesson across the 15-week semester consisting of 30 lessons (two per week). This approach aligns with principles of retrieval practice and spaced repetition, both of which are recognized to enhance long-term vocabulary acquisition by requiring learners to actively recall words across multiple intervals (Roediger & Butler, 2011).

Each quiz had an allotted time of 5–10 minutes. Quizzes were auto-scored through the Google Forms platform, and results were made immediately available to students, allowing them to reflect on their performance. Each vocabulary quiz was worth 1%, for a total of 10% of the overall class grade.

Data Collection

The dataset consisted of responses from two IELTS Reading and Writing courses in the LEAP program at Toyo University during the Spring 2025 semester, with 20 and 18 registered students, respectively. Across both classes, the number of submissions per quiz ranged from 16 to 20, depending on attendance. Quiz 12 had no responses at the time of analysis due to scheduling constraints and was therefore excluded from the dataset.

To address RQ2 (How effective large language models are in vocabulary assessment preparation work in terms of time saved and quality of output), effectiveness was operationalized along two dimensions: (a) productivity and (b) test-item quality.

To measure productivity, the instructor recorded the time required to complete key stages of quiz production for both manually created and LLM-generated quizzes. For manual / human creation, time segments included item construction (drafting fill-in-the-blank questions and selecting correct answers, distractor selection), and entry and formatting of items in Google Forms. For LLM-generated quizzes, timing included prompt submission and model response (automatic production of quiz items and answer options by the LLM), post-editing of output (correcting deviations, refining distractors, and adjusting phrasing), and entry into Google Forms.

It should be noted that substantial upfront time (approximately 30–45 minutes) was required to design (designing task instructions and constraints for quiz generation) and refine the initial prompt used for quiz generation. However, this prompt engineering constituted a one-time setup and was therefore excluded from the per-quiz timing, as the same prompt structure was reused across all subsequent quizzes.

To evaluate test-item quality, we examined student performance outcomes and distractor quality. Student response data were exported from Google Forms for each quiz, including total

scores per item (performance outcomes) and item-level response frequencies for all multiple-choice options. Distractor analysis was conducted for all 110 questions. A distractor was classified as ‘functioning’ if it attracted at least 5% of student responses, a threshold commonly used in item analysis to distinguish meaningful distractors from infrequently selected options. These metrics were computed separately for manually created quizzes (Quizzes 1–6) and LLM-generated quizzes (Quizzes 7–11) in order to compare test quality across conditions.

All productivity and test-item quality analyses were conducted using Microsoft Excel for data organization, calculation of descriptive statistics, and aggregation of response-level data. Limitations related to the use of lexical profilers and LLMs for quiz creation (RQ3) were identified through qualitative examination of item-generation errors, prompt deviations, and constraints observed during the productivity and test-quality analyses.

Results

Productivity

As shown in Figure 5, for manual / human created quizzes, total item-development time ranged from approximately 11–14 minutes per 10-item quiz. For the instructor, item construction and Google Forms entry occurred largely simultaneously, as sentence stems, correct answers, and distractors were composed and typed directly into the Google Forms. Approximately 8–10 minutes was spent on item construction (the cognitive task of formulating sentence stems and selecting correct answers / distractors) and 3–4 minutes was given to manual entry into Google Forms, including typing fill-in-the-blank sentences and entering multiple-choice options. It should be noted that the test creation times recorded were an average of the author’s own and variations may occur depending on the instructor.

Figure 5

Table of Manual vs. LLM-assisted Vocabulary Quiz Creation

Stage	Manual / Human	LLM-assisted
Item construction	8–10 min	~0 (LLM does it)
Prompt submission + model response	-	~0.5 min
Post-editing	-	2–4 min
Entering into Google Forms	3–4 min	2–4 min (copy-paste)
Total preparation time per quiz	11–14 min	5-8 min

Note. For manually created quizzes, item construction and Google Forms entry were performed concurrently by the instructor; time estimates are presented as conceptual components rather than strictly sequential stages.

For LLM-generated quizzes, total per-quiz preparation time after prompt development ranged

from approximately 5–8 minutes. Prompt submission and model response time using GPT-4o averaged approximately 20–40 seconds per quiz, during which the system generated the full set of ten questions and answer options. On average, approximately 2–4 minutes per quiz were spent on post-editing to correct prompt deviations (instances where the generated output did not fully align with the intended format or constraints), refine distractors (to improve plausibility and alignment with the target vocabulary), or adjust phrasing (to enhance clarity and appropriateness for the learner level), while the remaining time was used for Google Forms entry. As with the manual condition, all items still had to be transferred into Google Forms; however, this step was faster due to copy-and-paste insertion of complete question sets generated by the LLM.

A comparative analysis of the two workflows revealed that LLM-assisted quiz construction yielded a 45–50% reduction in total preparation time relative to manual methods. This efficiency gain was primarily driven by the automation of initial item drafting and distractor suggestion.

Test-Item Quality

Student Performance Outcomes

The score distributions for both conditions showed similar results (Figure 6). Both sets of quizzes exhibited scores on the upper end of the scale, with the majority of students achieving high marks, clustering between 8 and 10.

Figure 6

Item Difficulty and Quiz Results

Condition	Quiz Range	Total Student Responses (N)	Mean Score / Facility (p)	Standard Deviation (SD)
Manually Created	Quizzes 1–6	211	0.825	1.454
LLM-Generated	Quizzes 7–11	146	0.830	1.507

The transition from manually authored assessments (Quizzes 1–6) to LLM-generated quizzes (Quizzes 7–11) resulted in fairly stable outcomes across the dataset. Statistics for the human-authored quizzes (N = 211) revealed a mean score facility (p) of 0.825 with a standard deviation (SD) of 1.454, indicating a generally competent performance of vocabulary mastery among students.

Following the implementation of the LLM-assisted workflow, the LLM-generated quizzes (N = 146) demonstrated nearly identical performance characteristics, with a mean facility of 0.830 and an SD of 1.507. Despite a different number of total responses, score distributions remained

consistent across conditions, with results clustering toward the upper end of the scale.

Distractor Analysis

Distractor quality was evaluated using a standard applied criterion. A distractor was classified as *functioning* if it attracted at least 5% of student responses. The metrics in Figure 7 summarize distractor efficiency across manually created quizzes (Quizzes 1–6) and LLM-generated quizzes (Quizzes 7–11).

Figure 7

Distractor Efficiency Comparison

Metric	Quizzes 1–6 (Human)	Quizzes 7–11 (LLM)
Mean Functioning Distractors per Item	1.46	1.64
Items with 3 Functioning Distractors	7 / 60 (11.6%)	10 / 50 (20.0%)
Total Non-Functioning Distractors (NFDs)	91	68
Mean NFDs per Quiz (10 items)	15.17	13.60

Although the mean number of functioning distractors per item increased from 1.46 to 1.64 in the LLM-generated quizzes, this represents a modest change in absolute terms (approximately 1–2 additional functioning distractors per 10-item quiz). Importantly, this increase did not correspond to a meaningful change in overall item difficulty, as reflected in the stable facility values across conditions. Across all items, manually created quizzes contained a higher total number of non-functioning distractors (NFDs; $n = 91$) than LLM-generated quizzes ($n = 68$), indicating fewer unused distractor options in the LLM condition. In proportional terms, non-functioning distractors accounted for 50.6% (91/180) of distractors in the human-authored quizzes and 45.3% (68/150) in the LLM-generated quizzes, representing a modest reduction of approximately five percentage points in the LLM condition.

While both conditions contained a substantial number of distractors that attracted zero selections, the LLM-generated quizzes showed better distractor efficiency, with a higher mean number of functioning distractors per item (1.64 vs. 1.46) and a greater proportion of items where all three distractors functioned (20.0% vs. 11.9%). In practical terms, the LLM condition reduced the average number of non-functioning distractors per quiz (13.6 vs. 15.2), suggesting more consistent plausibility across the full set of items.

Discussion

Stage 1

RQ1: How can vocabulary profilers be used to identify vocabulary that should be prioritized in an IELTS preparation course?

The authors have described how LEXTUTOR (Cobb), a freely accessible vocabulary profiler, can be a valid tool for homing in on vocabulary for instruction when creating target vocabulary lists for an IELTS preparation course when used in tandem with spreadsheet software such as Microsoft Excel. For this case study, LEXTUTOR made it possible for the teacher to not only understand the word difficulty level of reading passages objectively but also identify high frequency words that appear multiple times over the course of a semester. Klinger (2024) suggests students should be exposed to words that are worthwhile to learn more than three times to be learned effectively. This was incorporated as one of the criteria when screening the textbook reading passage corpus down to 1.7% of its entirety.

The IELTS is a language proficiency test which does not focus on a specific content topic. This makes learning high frequency words in the English language an obvious necessity to achieve a high score on the test. Especially for students with an English proficiency of CEFR A2-B1. For students with a higher proficiency, including words that are identified as off list of the NGSL and NAWL is also possible.

Stage 2

RQ2: How effective are large language models (LLMs) in vocabulary assessment preparation work in terms of productivity gains and test quality output?

Productivity

The use of large language models (LLMs) to generate vocabulary quizzes demonstrated clear productivity gains in the assessment preparation process. A time-on-task comparison showed that LLM assistance substantially reduced human item-construction workload, particularly in the generation of sentence stems and distractor sets. While manual quiz creation required instructors to design, type, and format each item individually, LLM-assisted workflows shifted much of this burden to automated content generation followed by human review. Although post-editing was still necessary, particularly in relation to distractor refinement and adherence to prompt constraints, these adjustments functioned primarily as targeted calibration rather than substantive correction.

Although initial prompt engineering required a one-time investment of time, this cost was amortized across subsequent quizzes and did not affect per-quiz efficiency. As model response accuracy and prompt adherence continue to improve, it is likely that post-editing time will decrease

further, leading to additional efficiency gains in AI-assisted assessment development.

These time savings suggest that LLMs may allow instructors to reallocate effort away from clerical tasks, though how this reclaimed resource is ultimately used remains dependent on instructor practice.

Test Quality Output

Productivity represents only one component of effective assessment design; efficiency gains are meaningful only if resulting quizzes maintain appropriate difficulty levels and avoid introducing unintended complexity through distractor design.

The quantitative analysis of Quizzes 1–10 indicates that the transition from manual item generation to LLM-assisted generation resulted in no meaningful change in student performance. Mean item facility remained nearly identical across conditions ($p \approx .81$), and standard deviation values showed comparable score distributions. These findings suggest that LLM-assisted quiz construction replicated the intended difficulty level of the instructor-authored assessments without increasing score variability.

The data from Quizzes 1–6 (human-authored) illustrate challenges in distractor performance across items. The human-authored quizzes (Quizzes 1–6) contained a higher number of non-functioning distractors, reflecting less consistent distractor performance across items. In some cases, non-functioning distractors appeared thematically or contextually misaligned with the semantic constraints of the stem. For example, in the item “The book had a huge _____ on his life,” options such as *insight*, *secret*, and *left* were grammatically plausible but failed to satisfy the collocational and causal meaning required by the sentence, leaving *impact* as the only semantically appropriate choice. While human intuition can produce effective items, the LLM-assisted condition demonstrated slightly more consistent distractor functioning across the assessment set, as reflected in the lower number of non-functioning distractors.

Taken together, these results indicate that the reduction in preparation time did not come at the cost of assessment quality. The LLM-assisted workflow maintained overall difficulty and performance characteristics while improving distractor consistency. Beyond the present dataset, the scalability of LLM-assisted workflows warrants consideration. While the current study involved approximately 180 target vocabulary items, instructional contexts involving substantially larger lexical sets would require proportionally greater manual item-construction effort. Under such conditions, the relative efficiency advantage of LLM-based generation may become more pronounced, particularly in large-scale or cumulative vocabulary programs.

While prior research has documented substantial item-writing flaws in some LLM-generated assessments, the present findings suggest that outcomes may be context-dependent. When compared with manually authored quizzes in this study, the LLM-generated items demonstrated

slightly greater consistency in distractor functioning, including a modest proportional reduction in non-functioning distractors. Moreover, many distractors classified as non-functioning were grammatically plausible yet unselected by students, indicating that non-functioning status may reflect cohort response behavior and overall performance level rather than inherent irrelevance alone.

RQ3: What are the limitations of using lexical profilers and LLMs for creating vocabulary quizzes?

Limitations During Stage 1

The limitations observed in this stage can be grouped into two broad categories: word misrecognition by software and lack of candidate words fully meeting criteria. Word misrecognition by the different software used added procedures to the workflow that were not originally anticipated. The phenomenon was a good reminder to not put too much trust in technology and to manually double check if the output makes sense. Lack of candidate words that fulfill the set criteria for creating 12 wordlists of 15 words each from a corpus of 10,340 can be a challenge for any teacher. Details of the limitations during phase 1 and how they were coped with are described.

Word Misrecognition

Word misrecognition is a phenomenon that can occur when using OCR software to recognize the text on documents that were digitized with a scanner. It can also occur when copying and pasting texts from one format to another. For example, copying texts from a PDF to Microsoft Word or other online platforms. In the case of this study, word misrecognition occurred when; copying text from a PDF to LEXTUTOR, copying text to Microsoft Word, and using OCR on a scanned document.

In the case of copying the reading passages from a PDF to LEXTUTOR directly, unnecessary spaces appeared in the middle of words such as “mention”. This created an unusually high count of the word “ion” in the LEXTUTOR output, which would be very unlikely since the 12 passages covered a variety of topics including history, psychology, and engineering. As a method to check if the error was due to directly copying and pasting texts from a PDF to the webpage and checking for any other type of mistakes in the text, the reading passages were copied to a Microsoft Word document for checking.

When the documents were transferred to Microsoft Word, four types of errors were identified: letters morphing into Chinese characters, letters missing, addition of unnecessary spaces mid-word, and letter shuffling. On 6 out of the 12 passages four sets of Chinese characters were found. The letter combination “fi” was represented as 昫椀 94 times, “fi” as 昫氈 7 times, and “ff” as 昫昫 4 times. On four passages some letters of words were not recognized leading to words with missing

letters. The letter combination “Th” was missing from the words “this” and “the” 3 times, “tt” in “cutting” once, “fl” in words such as “flow” 9 times, and hyphens in hyphenated words such as “short-distance” twice. Unnecessary spaces mid-word such as “plat form” were located 15 times in two passages. In one passage the letters were scrambled for 345 words out of 945 making it impossible to read. For this unreadable passage, the textbook pages containing this reading were scanned.

A CZUR Shine digital scanner was used to digitize the passage from the physical textbook. Unfortunately for reasons unknown the OCR software that is included in the scanner driver could not convert the scanned documents into a searchable PDF or a Microsoft Word document successfully. The Microsoft Word document version of the problematic passage was cleaned using ChatGPT and manually compared against the original passage in the textbook to assure there were no mistakes then run through LEXTUTOR.

Lack of Candidate Words Fully Meeting Criteria

Due to the nature of the shortlisted words, lists that were scheduled to be used later in the semester did not have enough candidate words to be placed on the final vocabulary list of 15-words. For these lists, a second criterion was used to decide which words to add. Words that appear in less than two passages but appear multiple times in a single passage were considered to be added to the list for that passage. Frequency of appearance in a single passage was determined by running all words from each passage separately through LEXTUTOR (Cobb) a second time. When there were not enough words to complete the lists of 15 using the two previously mentioned criteria, an executive decision was made to include words that made an appearance in only one passage and belonged in bands 2, 3, or A in the LEXTUTOR vocabulary profile output. A total of 8 words fit this description.

In relation to the nature of shortlisted words, not all vocabulary lists were able to fully follow the word difficulty standard of 9 words from band 2, 3 words from band 3, and 3 words from band A of the LEXTUTOR output.

Limitations During Stage 2

LLM-Generated Output

The limitations observed primarily concerned inconsistencies in model-generated output. While LLMs substantially accelerated quiz creation, adherence to prompt constraints was not always reliable and required human review. Even when instructed to use IELTS Band 6–7-level English, some generated items exhibited phrasing that was overly formal or textbook-like rather than reflective of authentic usage. In several instances, the model failed to follow explicit prompt

constraints, such as the requirement that all distractors be drawn from a provided vocabulary list; for example, items occasionally used distractors that also appeared in the pool of eligible correct answers (e.g., *trend* or *artifact*), necessitating post-generation editing.

Despite improved distractor functioning overall, some challenges were observed in distractor construction within the LLM-generated items. In some instances, although distractors were intended to be plausible but incorrect, some options overlapped too closely with the correct answer, creating ambiguity, while others were clearly implausible or unrelated, reducing item quality and discriminatory power. These inconsistencies underscore the need for post-editing and instructor judgment when deploying LLM-generated assessment materials.

These limitations align with findings reported in prior research on LLM-generated assessment items. Camarata et al. (2025) reported that nearly half of ChatGPT-generated medical multiple-choice items contained item-writing flaws, with over one-fifth exhibiting factual or conceptual errors. Similarly, Bitew et al. (2023) found that 25% of stems and 33% of distractors in LLM-generated vocabulary items were inappropriate, while Putri et al. (2025) reported that 41.27% of generated distractors were irrelevant.

Sample Size and Analytic Approach

Several limitations should be considered when interpreting the findings of stage 2. First, class sizes were relatively small, with between 16 and 24 responses per quiz, which may limit the generalizability of the results to larger or more diverse learner populations. In addition, counts of non-functioning distractors (NFDs) are sensitive to sample size, as distractors may remain unselected due to chance rather than poor design, particularly in high-performing cohorts.

An additional limitation concerns the operationalization of item difficulty in the present study. Difficulty was inferred primarily from student response behavior and distractor performance, without accounting for other dimensions that may contribute to perceived item difficulty, such as stem length, syntactic complexity, or overall reading load. As a result, the analysis reflects difficulty at the level of response options rather than the full cognitive demands of each item.

The analysis relied primarily on descriptive statistics rather than inferential testing, reflecting the process-based and exploratory nature of the study. While this approach was appropriate given the bounded timing estimates and repeated-measures structure of the data, it limits the strength of statistical claims that can be made regarding observed differences between conditions.

Instructional Context and Task Scope

The study was conducted within a single instructional context and involved quizzes created and reviewed by one instructor, which may constrain the applicability of the findings to other teaching settings or levels of instructor expertise. In the LLM-assisted condition, all generated quiz

items were post-edited by the instructor, meaning that the observed outcomes reflect a human–LLM collaborative workflow rather than fully automated item generation.

Furthermore, the 5% threshold used to classify functioning distractors represents a discretionary analytic decision rather than a fixed standard, and alternative thresholds could produce slightly different classifications. Finally, the analysis focused exclusively on vocabulary quizzes, and the findings may not extend to other assessment formats or language skill domains.

Conclusion

This preliminary study demonstrates that vocabulary profilers and LLMs can be used in combination to support IELTS test preparation. The use of lexical profilers such as LEXTUTOR with NGSL/NAWL ensured level-appropriate vocabulary selection, while LLM assistance substantially reduced the time required for quiz construction. Together, this approach offers an efficient work flow to vocabulary assessment design within a test-preparation context.

The observed limitations of LLM-generated items did not translate into reduced assessment quality when compared with manually authored quizzes. Item-analysis metrics indicated comparable difficulty levels and slightly improved distractor functioning in the LLM-assisted condition. Within this dataset, model-generated quizzes were at least as stable as instructor-created items.

While prior studies have documented substantial rates of item-writing flaws in LLM-generated assessments, the present findings underscore the importance of contextual variables, including prompt structure and analytic thresholds such as the 5% criterion used to classify functioning distractors. In this study, many distractors deemed non-functioning were grammatically plausible yet unselected, suggesting that error rates in LLM-assisted assessment design should be interpreted with methodological caution rather than treated as inherent instability. A pragmatic approach rather than categorical acceptance or rejection may therefore be more productive for instructional contexts.

Finally, the findings help clarify the pedagogical role of LLMs in assessment design: not as replacements for instructor expertise, but as structured generative partners whose value lies in accelerating draft production while preserving human supervision. The results therefore suggest that LLMs operate most effectively as collaborative tools rather than fully autonomous assessment generators, preserving the need for human editorial oversight.

Implications and Future Research

Future research on curated vocabulary lists should examine their integration with additional language skills, particularly writing and speaking, as well as their impact across longer instructional periods and more diverse learner populations. Expanding the scope with these skills would allow

for a more comprehensive evaluation of how vocabulary-focused interventions contribute to broader language development within IELTS-oriented curricula. In addition, learner surveys and reflective instruments could be employed to investigate student perceptions of curated vocabulary lists, including their perceived usefulness in developing awareness of meaning, usage, and paraphrasing, and whether such knowledge transfers to authentic IELTS reading and listening tasks.

Unlike prior studies that systematically coded item-writing flaws or calculated formal error rates in LLM-generated assessments, the present study focused primarily on performance-based indicators such as facility values and distractor functioning. Future investigations could therefore incorporate structured coding of item-writing errors and formal error-rate analysis alongside multidimensional measures of item difficulty, including stem length and overall reading load, in order to better capture the full cognitive demands of assessment items.

LLMs may also be explored in the post-assessment processes, such as generating individualized feedback, identifying common error patterns, and tailoring follow-up exercises based on student responses. Lastly, future studies should also examine educator receptivity to LLM generated assessments and post-assessment processes, as the pedagogical value of these tools ultimately depends on how teachers and administrators engage, interact, and form policy with these technologies available.

References

- Bitew, S. K., Deleu, J., Develder, C., & Demeester, T. (2023, July 30). *Distractor generation for multiple-choice questions with predictive prompting and large language models*. arXiv. <https://doi.org/10.48550/arXiv.2307.16338>
- Browne, C. (2013). The New General Service List: Celebrating 60 years of vocabulary learning, *The Language Teacher*, 37:4, 13-15.
- Browne, C. (2014). A new general service list: The better mousetrap we've been looking for?. *Vocabulary learning and Instruction*, 3(2), 1-10.
- Camarata, T., McCoy, L., Rosenberg, R. L., Temprine Grellinger, K. R., Brettschneider, K., & Berman, J. (2025). LLM-generated multiple choice practice quizzes for pre-clinical medical students: Prevalence of item writing flaws. *Advances in Physiology Education*. <https://doi.org/10.1152/advan.00106.2024>
- Chen, Q. (2019). Phrasal paraphrase learning: The role of memory, analogy, and input frequency. *English Language Teaching*, 12(5), 68–80. <https://files.eric.ed.gov/fulltext/EJ1215106.pdf>
- Cobb, T. *Complete Web-VP v.2.6* [computer program]. Accessed 11 Jan 2025 at <https://www.lex tutor.ca/vp/comp/>.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Gottlieb, M. (2016). *Assessing English language learners: Bridges to educational equity: Connecting academic language proficiency to student achievement* (2nd ed.). Thousand Oaks, CA: Corwin.
- Hendry, C., & Sheepy, E. (2018). How much vocabulary is needed for comprehension of research publications in education?. *Future-proof CALL: language learning as exploration and encounters*, 94.

- Klinger, R. (2024). Vocabulary frequency and dispersion in Japanese junior high school EFL textbooks. *Vocabulary Learning and Instruction*, 13(2), 1-18.
- Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Multilingual Matters.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language learning*, 54(3), 399-436.
- Nakayama, S. (2022a). A close examination of vocabulary in Japanese EFL textbooks. *Reflections and New Perspectives*, 209-216.
- Nakayama, S. (2022b). Vocabulary in Japanese EFL textbooks: A bidirectional coverage analysis. In *The IAFOR International Conference on Education–Hawaii 2022 Official Conference Proceedings* (pp. 157-169).
- Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Putri, A., Fauziati, S., & Rizqi, A. A. (2025). Syntagmatic distractor generation for multiple-choice language tests: A large language model-based approach. *Proceedings of the 4th International Conference on Electronics Representation and Algorithm (ICERA)*.
<https://doi.org/10.1109/ICERA66156.2025.11087340>
- Roediger, H. L., III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20–27. <https://doi.org/10.1016/j.tics.2010.09.003>
- Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading. *Studies in Second Language Acquisition* 21, 589-619.
- Sasaki, M. (2022). Skill profiles of Japanese EFL learners: Evidence from large-scale assessments. *Language Testing in Asia*, 12(1), 1–23. <https://doi.org/10.1186/s40468-022-00203-3>
- Stewart, J. (2024). Establishing meaning recall and meaning recognition vocabulary knowledge as distinct psychometric constructs in relation to reading proficiency. *Language Testing*.
<https://doi.org/10.1177/02655322231162853>
- Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension?. *Applied linguistics*, 34(4), 457-479.
- Waring, R. and Nation, I.S.P. (1997) Vocabulary size, text coverage, and word lists. In *Vocabulary: Description, Acquisition and Pedagogy* N. Schmitt and M. McCarthy (eds.). Cambridge University Press, Cambridge: 6-19.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130-163.
- West, M. (1953). A general service list of English words. London: Longman, Green & Co.
- Xi, X. (2010). How do we go about investigating test fairness? *Language Testing*, 27(2), 147–170. <https://doi.org/10.1177/0265532209349465>
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). *Practical and ethical challenges of large language models in education: A systematic scoping review*. arXiv. <https://arxiv.org/abs/2303.13379>

交換留学生用日本語科目における 「訳語自動化単語リスト」導入の試み

大澤 恵 利 (東洋大学国際教育センター)

要旨

本稿は、交換留学生対象の日本語スピーキング授業において導入した「訳語自動化単語リスト」の実践を報告するものである。日本語中級レベルの履修者が多様な文化的・言語的背景を持つ中で、かねてよりプレゼンテーションの理解や集中力の維持に課題があった。先行研究が示す語彙リストの有効性に基づき、本クラスでは Google スプレッドシート上で漢字の読みがなや多言語訳を自動入力できる単語リスト書式を導入した。単語リスト書式には「よみたん API」による漢字読みがなの自動入力と GOOGLETRANSLATE 関数による多言語訳を組み込んだ。これにより、履修生は日本語列を入力するのみで単語リストを完成させることが可能となった。この結果、発表者は聴衆に単語リストの参照を促し、聴衆は単語リストを参照しながら発表を理解する姿勢が定着した。また、多言語訳を通じて各自の言語的背景を意識し、相互理解を深める契機ともなった。一方で、自動翻訳の精度や訳語の自然さには課題が残った。今後も他の翻訳ツールの活用可能性を探りつつ、多文化共生を意識した授業設計を進める必要がある。

キーワード：プレゼンテーション、多言語・多文化主義、単語リスト、機械翻訳、自動翻訳、多文化共生社会

目次

1. 本クラスの概要
2. 問題意識
3. 先行研究
4. 「訳語自動化単語リスト」とは
5. 「訳語自動化単語リスト」書式の作成方法
6. 「訳語自動化単語リスト」使用の実際
7. 漢字の読みがなと訳語の検証
8. 今後の課題
9. まとめ

1. 本クラスの概要

筆者が2023年度秋学期より担当している「Japanese Speaking Ⅲ」は、交換留学生のためのスピーキングクラスである。本学の日本語能力指標（TJ）において3レベル（日本語をある程度学んだレベル）の学生を対象としている。本クラスの目標は、中級レベルの日本語スピーキング力を習得することである。

交換留学生は開講前にプレースメントテストを受験し、その結果によりクラスが振り分けられるが、交換留学生が実際に履修するクラスを決定する際には、交換留学生自身の希望が優先されることもある。そのため、本クラスの履修生の日本語レベルにはばらつきがある。本クラスは日本語中級レベルであるため、クラス内での使用言語は基本的に日本語としている。しかし、担当教師が履修生とやり取りする際、あるいは履修生同士がやり取りする際には、英語が使用されることもある。

本クラスでは、履修生には自分の興味や関心に沿った3つのプロジェクトワーク（調べ学習）を課し、履修生はフォーマルな語彙・表現を使って3回のプレゼンテーションを行う。プレゼンテーションは、単なる口頭表現練習ではなく、大学教育で必要とされる論理的説明力や根拠を示す力といったアカデミックスキルを養う機会となる。また、多様な文化的背景を持つ履修生が互いの知識や視点を共有することで、異文化理解や協働的な学びを促す点でも教育的意義が大きい。

本クラスでは、まず担当教員（筆者）が発表デモンストレーションを行う。そして、日本語プレゼンテーションの基本的な構成（序論、本論、結論）、作成時の注意点について説明する。さらに、発表スライドの作り方、プレゼンテーションに欠かせないフォーマルな挨拶・表現、質疑応答の指導を行う。

また、本クラスでは、プレゼンテーションのテーマ一覧表を Google ドキュメントで作成し、クラスで共有している。プレゼンテーション実施前には履修生が各自で目標の設定を行い、プレゼンテーション実施後には自己評価と相互評価を行っている。担当教員からは履修生にコメントを記入したプレゼンテーション振り返りシートを渡し、個別指導している。

なお、本クラスには数名の学部学生がボランティア登録している。このボランティア学生は、履修生のブレインストーミング（発表テーマ決定のためのグループディスカッション）を補助したり、プレゼンテーションの質疑セッションに参加したりすることもある。

2. 問題意識

本クラスのプレゼンテーションは、履修生を3～4人のグループに分けてグループ内で行う場合（2回）と、一人ずつ前に出てプレゼンテーションを行う場合（1回）とがある。

残念ながら、クラスメートのプレゼンテーションを聴かずに自分のプレゼンテーション

の準備をしていたり、スマートフォンで授業に関係ないことをしていたりする履修生が一定数いる。プレゼンテーションの準備の際、テーマを事前に共有することに加え、履修生には発表用スライドに訳語を含めるなどの工夫を促してきた。しかし、本クラスの履修生にとって、初見のスライドを見ながらクラスメートの発表を聴くというのは簡単なことではない。中級レベルとはいえ、発表者が正確に発音し、適切な単語、表現、文法を用いて話しているとは限らないからである。そのため、履修者がクラスメートのプレゼンテーションを十分に理解していない可能性がある。特に大教室での個人プレゼンテーションの際は、履修生が次第に集中力を失っていく様子が見て取れた。発表中、筆者が机間巡視し、何度注意しても態度を正さない履修生が少なくなかった。このことは担当教員として大きな悩みであった。

3. 先行研究

柴崎（2005）は、外国語学習者が語彙リストで未知語の母語訳を知ることによってテキストの難しさを軽減する効果があることを明らかにした。これは、語彙リストを学習者のレベルや授業活動と結びつけて活用することの重要性と、語彙リストが単なる語彙の暗記にとどまらず実際の言語使用にも生かせる学習につながることを示唆している。

また、山本（2020）は、発表者が聴衆用資料として「難しい単語・表現リスト」を作成し、クラスで事前に共有することで聴衆のプレゼンテーション理解を支援できたとし、授業設計上、事前に語彙リストを構築・共有する実践の有用性を示した。

近年、外国語学習における機械翻訳の活用に関する調査研究が増えているが、寺門（2023）、寺門ら（2024）は、日本語学習における機械翻訳の活用が語彙・文法の確認、カタカナ語や熟語の習得、日本語コンテンツの教材化などに有効であるとした。特に、機械翻訳は「書く」、「読む」の学習で効果が大きく、多言語環境における日本語教育支援の可能性を示している。その一方で、誤訳や不自然な表現もあるため批判的利用が必要であることも指摘している。

4. 「訳語自動化単語リスト」とは

上記の先行研究から知見を得た筆者は、24年度秋学期から本クラスの履修生に対してプレゼンテーションの発表スライドと併せて単語リストの作成と事前共有を課すことにした。

本学の交換留学生の出身国・地域は様々で、当然のことながら母語も様々である。前述した通り、交換留学生の日本語クラスにおいては、目標言語である日本語以外に英語が使用される場面が少なくない。本クラスのシラバスも英語で書かれたものが配布されている。

しかし、多言語・多文化主義の観点から考えると、どの履修者の母語に対しても公平・公正であるべきである。そこで、筆者は英語以外の訳語も単語リストに組み込めないかと考え、多言語訳が自動入力されるようにあらかじめプログラムをした「訳語自動化単語リスト」書式の作成に取り掛かった。

多言語訳の単語リストを一から作成することは、履修者にとって大きな負担である。しかし、あらかじめ多言語訳の自動入力がプログラムされた書式を使用するのであれば、履修者にとっての負担は大幅に軽減される。

5. 「訳語自動化単語リスト」書式の作成方法

筆者は、まず Google スプレッドシートで単語リスト書式を作成し、筆者の本学アカウントの Google ドライブにアップロードした。Google スプレッドシートを使うメリットとしては、インターネットに接続されていれば環境を問わず使えること、複数人での共有や共同編集作業がしやすいことがあげられる。筆者は、発表グループごとにファイルを分け、タブで各発表者のシートを切り替えられるようにした。履修生には本学アカウントでの編集権を付与した。

単語リストには、まず漢字の読みがなが必須である。マイクロソフト社製の表計算ソフト Excel には、漢字の読みがなを取得する PHONETIC 関数が用意されているが、Google 社製の Google スプレッドシートでは、ふりがなを取得する PHONETIC 関数が使用できない。

そこで、この単語リスト書式で漢字の読みがなを取得するために「よみたん API」というひらがな化 API を使用した。「よみたん API」は、非商用の利用で 1 日のリクエスト回数が 1 万回以下なら会員登録が不要で、無料利用が可能であるためだ。この単語リスト書式に「よみたん API」を連携させるための Google Apps Script コード（以下、GAS コード）を導入することで、漢字の読みがなの自動入力が可能となる。この GAS コードは、引数を指定したセルのデータを取得し、「よみたん API」にリクエストした結果が出力されるというものである。なお、この GAS コードは、Google Apps Script や生成 AI を活用した業務改善に取り組んでいるエンジニア・ヤマタケ氏が独自に作成したもので、氏が個人ブログで一般公開しているものを使用した。

Google スプレッドシートには、Google ニュートラル機械翻訳システム (GNMT) を利用した GOOGLETRANSLATE 関数という独自の関数がある。各国語への訳語の自動入力は、Google スプレッドシートにこの GOOGLETRANSLATE 関数を組み込むことで実現した。手順は以下の通りである。

構文：

GOOGLETRANSLATE (テキスト, ソース言語, ターゲット言語)

- ① 1つめの引数 [テキスト] で翻訳するテキストを指定 (筆者は日本語列を指定)
- ② 2つめの引数 [ソース言語] で翻訳するテキストの言語を Google Translation の言語コードで指定。
- ③ 3つめの引数 [ターゲット言語] で翻訳後の言語を言語コードで指定。言語コードは「」(ダブルクォーテーション) で囲む。

表1 : 「訳語自動化単語リスト」の例
(囲みが GOOGLETRANSLATE 関数と実際の訳語)

The screenshot shows a spreadsheet with the following data:

	B	C	D	E	F	G
1			参考: Google翻訳 言語コード一覧			
2	日本語	読み	英語	韓国語訳	中国語訳 (簡体字)	中国語訳 (繁体字)
3	例 電気バス	でんきばす	electric bus	전기 버스	电动公交车	電動公車
4	感染症	かんせんしょう	infectious disease	감염	传染病	傳染病

Google Translation の言語コードを入れ替えることによって、英語はもちろん、あらゆる言語への翻訳が可能になる。なにより自動翻訳・自動入力であれば、スペルミスの心配がない。これも利用する側にとっては大きな利点である。

このように単語リスト書式にあらかじめ関数を組み込んでおけば、履修生の作業は日本語列の入力のみとなる。こうして、履修生が日本語列に入力すれば、漢字の読みがなど多言語訳が自動的に入力されるという単語リストの仕組みが完成した。

表2 : 本クラスで使用した Google Translation 言語コード一覧
(「Google Translation 言語サポート」より筆者が抜粋、作成)

言語名	言語コード
英語	en
韓国語	ko
中国語 (簡体字)	zh-CN
中国語 (繁体字)	zh-TW
ベトナム語	vi
フランス語	fr
ドイツ語	de
オランダ語	nl
スペイン語	es
イタリア語	it
ウクライナ語	uk

6. 単語リスト使用の実際

発表者には、発表開始時に「単語リストがありますので、ご覧ください」と言って、クラスメートに単語リストの閲覧を促してから発表を始めるよう徹底させた。この声掛けにより、ほとんどの履修生がその場で発表者の単語リストを開き、単語リストを見ながらクラスメートの発表を聴くようになった。また、履修生はプレゼンテーションの準備作業に慣れてくると、自主的に単語リストを作成するようになった。そして、発表時に発表者が聴衆に声掛けしなくても、自主的に単語リストを開き、見ながら発表を聴く様子も見られるようになった。

クラスで使用した訳語は、2024年秋学期で11か国語、2025年春学期で7か国語にのぼった。履修生はほかの外国語訳（主に英語）と母語訳を比較しながら、未習語の意味を確認していた。履修生にはアニメやマンガなど共通の話題も多いため、多言語訳をおもしろがる様子も見られた。

7. 漢字の読みがなと訳語の検証

この単語リスト書式の問題点は、その正確性である。「よみたん API」の漢字の読みがなは、「よみたん API」作成者が言うように、検索された読みがなが必ずしも正しいものとは限らない。日本語列に送り仮名までを入力することが大前提であるが、それでも間違った読みがなが検索されることがある。そのため、プレゼンテーションの前にあらかじめ筆者がすべての漢字の読みがなに目を通し、加筆・修正を行った。

前述の先行研究が指摘した通り、機械翻訳が完全に正しいというわけではなく、誤訳や不自然な表現になることもある。そのため、プレゼンテーションを実施した後、履修生に自分のグループで使用した単語リストの訳語を検証させた。

英語母語話者の履修生によれば、GOOGLETRANSLATE 関数による英語訳はほとんど問題ないとのことだった。英語以外の訳語については「意味はわかるが、自然ではない訳語もある」との回答もあったが、おおむね問題ないという回答が多かった。そこで、スプレッドシートのコメント機能を使って、正しい訳語、より適切な訳語を書かせた。

8. 今後の課題

GOOGLETRANSLATE 関数を使った翻訳の正確性については、やはり課題が残る。履修生による検証結果を見てみると、GOOGLETRANSLATE 関数での翻訳は、単語や単文レベルの翻訳ではほとんど問題ないことがわかる。しかし、イディオムや比喩的表現の翻訳となると、やや精度が劣る印象がある。これは、GOOGLETRANSLATE 関数がセル内のテキストという限定的な情報に基づいて翻訳することが原因であると考えられる。

一方、一部の履修生からは、「Google 翻訳より DeepL 翻訳の方が正確で良い」という意見もあった。筆者も日常的に DeepL 社製の機械翻訳サービス「DeepL」を利用しているが、精度の高い翻訳であることもさることながら、単なる単語の置き換えではなく、自然で違和感のない翻訳を出力するという印象を受ける。そこで、筆者は「DeepL」の無料プランを使った「訳語自動化単語リスト」の作成も試みた。Google スプレッドシート上で DeepL 翻訳を実現するには、まず利用者名でアカウント登録をしたうえで「DeepL API」キーを取得する必要がある。その後、GAS コードを導入して「DeepL API」と連携させる。「DeepL」の無料プランの翻訳可能字数は、1 カ月 50 万字までで、リクエスト速度にも制限がある。小規模の翻訳であれば無料プランでも十分足りるだろう。しかし、利用者名でのアカウント登録を求められることから、履修生が複数人で利用し、一斉編集をする「訳語自動化単語リスト」には向かない。

なお、ランゲージソリューションを手掛ける株式会社ヒューマンサイエンスが DeepL、Google、Microsoft、Amazon の機械翻訳サービスの翻訳精度を評価したところ、DeepL が IT 日英、医療英日・日英、契約書英日・日英、特許英日での翻訳精度で最も優れていたという。しかし、分野、言語方向によっては Google、Microsoft などほかの機械翻訳サービスが優れているということも指摘している。

以上のことから、現時点では利用制限がなく、だれでも無料で使用できる GOOGLETRANSLATE 関数が利便性という点で優れていると言えるのではないだろうか。今後、Google 翻訳の精度向上を強く期待したい。

9. まとめ

本クラスで「訳語自動化単語リスト」を使用したことで、履修生が同じクラスに様々な国・地域出身の留学生がいること、そして各々に文化的・言語的背景があることを意識できるようになったのであれば、担当教員として大変喜ばしい。履修生同士が互いの文化的・言語的背景に敬意を払うことは、多文化共生社会におけるコミュニケーションの基本姿勢であり、円滑な授業運営の基本でもある。履修生には、今後も互いの文化的・言語的背景を踏まえた上での積極的な授業参加を促したい。担当教員としては、理想的な単語リストの実現を目指し、今後も引き続き「訳語自動化単語リスト」の検証と改良を重ねながら本クラスで活用していきたい。

参考文献・参考資料

- ・ 柴崎秀子 (2005) 「第二言語におけるテキスト要因と語彙リストの関係」『留学生日本語教育センター論集』 31, pp.141-241.
- ・ 寺門芽衣 (2023) 「AI 機械翻訳利用の影響と日本語学習での AI 機械翻訳の活動方法の検討」『2023 年度日本語教育学会グローバル人材奨励プログラム活動報告』

- ・ 寺門芽衣、菅谷克之（2024）「日本語学習における AI 機械翻訳の活用方法」『CIEC 春季カンファレンス論文集』 15,pp.15-21.
- ・ 山本由紀（2020）「大学英語教育においてプレゼンテーションがスピーキング能力向上に与える効果」『武蔵野大学教養教育リサーチセンター紀要』 Vol.10, pp.45-61.
- ・ 「機械翻訳とは？メリット・デメリットから最新動向、「DeepL」と「Google 翻訳」との比較まで解説！」
<https://www.science.co.jp/nmt/blog/32334/#link05>（2025 年 9 月 21 日アクセス）
- ・ 「スプレッドシートで漢字のよみがな（ふりがな）・ルビを取得する方法～ GAS によるオリジナル関数」『Auto Worker ～ Google Apps Script(GAS) と Sikuli で始める業務改善入門』
<https://auto-worker.com/blog/?p=3081>（2025 年 9 月 21 日アクセス）
- ・ よみがな検索システム「よみたん API v2」
<https://yomitan.harmonicom.jp/>（2025 年 9 月 21 日アクセス）
- ・ 「DeepL API キーの取得方法 DeepL とスプレッドシートを使った高精度な翻訳の準備」
<https://84office.jp/get-deepl-api-key/>（2025 年 9 月 21 日アクセス）
- ・ 「DeepL API を使ってスプレッドシートで自然な翻訳を実現する方法」<https://84office.jp/googlesheet-deepl-api-translation/>（2025 年 9 月 21 日アクセス）
- ・ 「GOOGLETRANSLATE」
<https://support.google.com/docs/answer/3093331?hl=ja>（2025 年 9 月 21 日アクセス）
- ・ 「Google Translation 言語サポート」
<https://cloud.google.com/translate/docs/languages?hl=ja>（2025 年 9 月 21 日アクセス）

An attempt to introduce an ‘automated translation word list’ in Japanese language courses for exchange students.

Eri Osawa

(Center for Global Education and Exchange, Toyo University)

Abstract

This paper reports on the implementation of an “automated translation word list” in Japanese-speaking classes for exchange students. Intermediate-level learners with diverse native language backgrounds faced challenges in comprehension and concentration. Prior research has shown that vocabulary lists are effective. Building on this, the course used a word list format in Google Spreadsheets that automatically inputs kanji readings and multilingual translations. The format included automatic kanji reading input via the “Yomitan API” and multilingual translations using the GOOGLETRANSLATE function. Students completed the word list by simply entering the Japanese words in the corresponding column. Presenters encouraged the audience to use the word list, and the audience developed the habit of consulting it to better understand the presentation. The multilingual translations also helped participants become aware of their own linguistic backgrounds and supported mutual understanding. Challenges remain regarding the accuracy and naturalness of automatic translations. Moving forward, it will be necessary to explore other translation tools and develop course designs that foster multicultural coexistence.

Keywords: Presentation, Multilingualism/Multiculturalism, Word lists, Machine translation, Automatic translation, Multicultural society

難民日本語教育の研究動向

片山 奈緒美 (東洋大学国際教育センター)

要旨

日本は難民認定率が低いことで知られるが、近年、難民認定申請数が増加する傾向にあり、ミャンマーやアフガニスタン等の政変などの影響で難民認定数も増加した。認定された難民には文化庁の「難民に対する日本語教育」プログラムが提供され、日本語支援のみならず、自立して生活していくためのさまざまな支援が行われている。難民が日本で生活していくために日本語習得は重要な役割を果たすが、文化庁のプログラムは難民保護を目的に非公開でプログラムが実施されているため、どのように日本語教育が行われているのかについては、これまであまり知られてこなかったという側面もある。本研究では論文検索サイト CiNii におけるキーワード検索で難民を対象にした日本語教育の先行研究を抽出し、抽出された先行研究のデータからテキストマイニングを用いて可視化分析を行った。さらに「日本語教育の俯瞰図」を用いて先行研究が扱ってきた分野を整理し、これまでの難民を対象にした日本語教育研究の動向を捉え、まだあまり扱っていないテーマを明らかにしたい。

キーワード：難民日本語教育、「難民に対する日本語教育」、テキストマイニング、ワードクラウド、「日本語教育学の俯瞰図」

目次

1. はじめに
2. 研究目的と方法
 - 2.1 研究目的
 - 2.2 研究方法
3. 研究結果
 - 3.1 CiNii の先行研究抽出とワードクラウド
 - 3.2 「日本語教育学の俯瞰図」
 - 3.3 難民日本語教育研究の分類
 - 3.4 年代別の研究動向
4. まとめ

1. はじめに

日本は世界的に難民認定率の低さで知られるが、コロナ禍にあった2020～2022年を除くと、2015年頃から日本における難民認定申請（以後、難民申請）の数が増加傾向にあり、それに伴って条約難民の認定数もこれまでにない高水準を維持しているⁱ。国は1951年の出入国管理及び難民認定法において難民の認定を受けた条約難民及びその家族に日本語教育支援や就労支援を行うことを2002年に閣議了解したⁱⁱ。この「条約難民に対する日本語教育事業」に加え、2014年の閣議了解等で、第三国定住難民に対しても「第三国定住難民に対する日本語教育事業」を行うこととし、条約難民及び第三国定住難民の定住促進を図ってきた。こうした難民を対象とした日本語教育は難民の定住支援事業と位置づけられて法人格を有する団体に業務委託され、文化庁が委託費を支出するかたちで行われているⁱⁱⁱ。その概要を示すのが図1である。

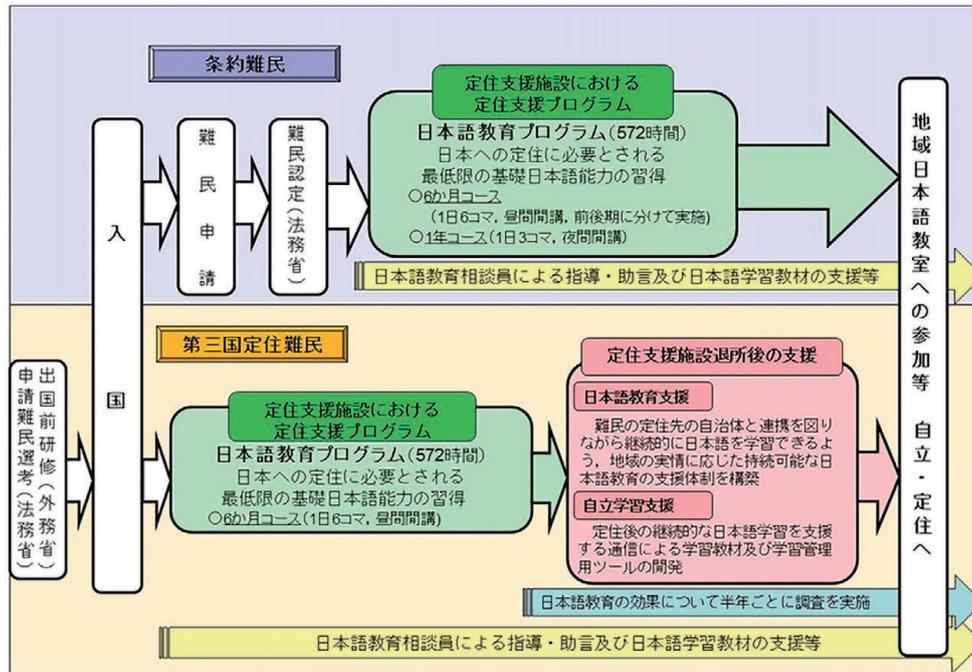


図1 条約難民及び第三国定住難民に対する日本語教育事業（文化庁）^{iv}

図1の条約難民を例に入国から自立・定住までの流れを見ていくと、入国後に難民申請を行い、難民として認定されたのちに日本での定住支援プログラムのひとつである日本語教育プログラムを受講できる。日本語教育プログラムは文化庁の業務委託を受けた法人格を有する団体がを行い、難民認定者は1日6コマの6か月コースか1日3コマの1年コース

ⁱ 法務省「我が国における難民保護の状況等」。

ⁱⁱ 平成14年8月7日閣議了解「難民対策について」。

ⁱⁱⁱ 平成22年2月19日「条約難民及び第三国定住難民に対する日本語教育事業委託実施要領」。

^{iv} 文化庁 https://www.bunka.go.jp/seisaku/kokugo_nihongo/kyoiku/nanmin_nihongokyoiku/

で合計 572 授業時間（1 授業時間は 45 分）のプログラムで「日本への定住に必要とされる最低限の基礎日本語能力」の習得を目指すことになる。この間、日本語教育相談員による指導・助言及び日本語学習教材の支援等」が提供され、プログラム修了後、難民は地域日本語教室に参加するなどして日本語のさらなる習得を進めながら自立・定住していくという設計である^v。

しかし、難民を対象とした日本語教育事業についての情報の開示は限定的であり、文化庁ウェブサイト内での図 1 を含めた「難民に対する日本語教育」等における事業の全体像の概略や使用している教科書や教材『はじめまして にほん』『にほんご えじてん』等の公開にとどまっている。日本語教育プログラムを受講する難民の日本語学習歴やプログラム修了後の習得度など難民自身に紐付けられる情報はもちろんのこと、『はじめまして にほん』や『にほんご えじてん』がプログラムで具体的にどのように使われているのか、日本語 4 技能をどのように伸ばす試みがなされているかなど、公開されない要素が多い。そのため、難民に対する日本語教育は研究の機会がきわめて限られており、日本語教育の中でも研究が進んでいない分野の一つと言えるだろう。これらを踏まえ、本研究はこれまでの難民に対する日本語教育について研究動向を整理し、まだあまり扱われていない研究課題を明らかにすることを研究目的とする。これにより、今後の難民の日本語教育の発展に貢献できる意義のある研究であると考えている。

ここで難民の定義について確認しておきたい。難民条約では、条約が適用される難民について次のように定義している。

人種、宗教、国籍もしくは特定の社会的集団の構成員であることまたは政治的意見を理由に迫害を受けるおそれがあるという十分に理由のある恐怖を有するために、国籍国の外にいる者であって、その国籍国の保護を受けることができない者またはそのような恐怖を有するためにその国籍国の保護を受けることを望まない者及びこれらの事件の結果として常居所を有していた国の外にいる無国籍者であって、当該常居所を有していた国に帰ることができない者またはそのような恐怖を有するために当該常居所を有していた国に帰ることを望まない者^{vi}。

日本は 1970 年代後半にインドシナ難民が大量に発生したことを受けて 1981 年に難民条約、1982 年に難民議定書に加入した。そして、それまでの法令を改正して新しい難民認定制度を導入した「出入国管理及び難民認定法（入管法）」を制定したことにより、難民申請者が難民として認定されると社会保障ほか一定の保護を受けられるようになった^{vii}。

^v プログラム受講が対面またはオンラインを選択できることなど、2025 年度のプログラム申し込みの詳細は日本語教育プログラムの委託を受けた公益財団法人アジア福祉教育財団難民事業本部 (RHQ) のウェブサイトで公開されている。

^{vi} 1951 年の難民条約第 1 章第 1 条 A(2) より。 <https://www.unhcr.org/jp/treaty-1951>

^{vii} 外務省「国内における難民の受け入れ」。 <https://www.mofa.go.jp/mofaj/gaiko/nanmin/main3.html>

しかし、居住地または宿泊地等がある地方出入国在留管理官署に難民申請のための書類が提出されたあとに行われる審査の基準は上記の難民条約における難民の定義に基づくとされているのみであり^{viii}、難民として認定されるか、不認定となるかの境界線は不透明である。そのため、難民として庇護を求める人々の立場は極めて不安定であると言える。また、難民認定されて定住のためのビザを取得し、さまざまな保護を受けられるようになっても、就労や日本語習得、地域での受け入れ状況などの面でやはり不安定な生活を免れないのが現状である。こうした状況を踏まえ、伴野（2013）は難民の日本語教育は不安定で不確実な現場で生身の人間である難民に対して行われ、理論と実践間の往還または統合を目指すものだとし、これを「難民日本語教育」とした。以後、本稿においても難民に対する日本語教育を「難民日本語教育」と呼ぶこととする。

2. 研究目的と方法

2.1 研究目的

先述したとおり、難民に対して実際に日本語教育を行うのは文化庁から業務委託された法人格を有する団体であるが、この日本語教育プログラムは文化庁の事業として日本において限定的に実施されている公的な日本語教育の一つであり、これまでの事業成果の検証と今後の発展のための調査研究が望まれる。だが、難民本人や家族の保護の目的から研究者などが当事者に接触するのは困難であり、2025年現在、難民日本語教育に関連する研究は限られている。しかし、今後不安定な世界情勢によって大量の難民が発生し、その一部を日本で受け入れたときのために、難民日本語教育について研究を進めておくことに意義があると考えられる。本研究は今後の難民日本語教育の発展のために、現時点での難民日本語教育の研究動向を捉え、これまで先行研究があまり扱ってこなかった研究課題を明らかにすることを研究目的とする。

2.2 研究方法

本研究は、まず論文検索サイト CiNii^{ix} でキーワードを入力して難民日本語教育に関連する先行研究を抽出した。次に抽出された論文のタイトル及び抄録について「AI テキストマイニング」^x でテキストマイニングを行って生成されたワードクラウドによる可視化分析を行った。ワードクラウド上ではワードの大きさで出現回数の多少を表し、色でワードの品詞が区別される。この可視化されたデータによって難民日本語教育の研究動向の全体像を可視化分析した。最後に CiNii で抽出された先行研究を「日本語教育学の俯瞰図」（日

^{viii} 出入国在留管理庁「難民認定手続・補完的保護対象者認定手続」<https://www.moj.go.jp/isa/refugee/procedures/16-6.html>

^{ix} CiNii（NII 学術情報ナビゲータ）<https://cir.nii.ac.jp>

^x AI テキストマイニング <https://textmining.userlocal.jp>

本語教育学会 2023)^{xi} に沿って分類した。以上のように CiNii から抽出された先行研究をワードクラウドで可視化分析し、また「日本語教育学の俯瞰図」を用いてこれまでの研究動向を分類・整理することで、難民日本語教育の先行研究があまり扱ってこなかった研究分野を明らかにしたい。

3. 研究結果

3.1 CiNii の先行研究抽出とワードクラウド

2025年6月23日、CiNii で「難民／日本語」「難民／日本語教育」「難民／日本語プログラム」「難民／言語／習得」「難民／言語教育」などの複数の組み合わせでキーワードを入力し、1980～2025年に発表された難民日本語教育にかんする先行研究を検索した。重複や外国における難民の言語教育制度のみを扱ったものを排除すると「難民 日本語」の入力で最も多くの先行研究を検索できたため、この検索結果 197 件を本研究の基礎データとした。その内訳は表 1 の通りである。

表 1 CiNii 「難民 日本語」 検索結果

研究データ	論文	本	博士論文	プロジェクト	合計
3	77	67	5	45	197

このうち CiNii 上で「論文」「本」「博士論文」に分類されている計 149 件から、タイトル及び抄録から重複しているものと外国の難民の事例のみを扱ったものは除外し、日本の難民に対する日本語教育や日本語習得、他国と日本の難民の言語教育の比較等、日本の難民の日本語教育に関連のある先行研究を抽出したところ、37 件が残った。この 37 件についてタイトルと抄録、抄録が無い場合はタイトルのみ、または論文本文の第 1 章「はじめに」などで論文の概要を記した部分を AI テキストマイニングに入力したところ、図 1 のようなワードクラウドが生成された。

まず、図 2 では、表示される語の大きさで出現回数の多少が表現され、名詞が青、動詞が赤、形容詞ほか緑で表されている。文字の大きさが目立つのは名詞に集中しており、先行研究のタイトルと抄録において「日本語教育」「日本」「インドシナ難民」「難民」「人材」「教育」「定住」「ベトナム」等の語の使用頻度が高かったことがわかる。尚、名詞が表示されている青字のなかに「インドシナ難民日本語教育」や「難民」などの語の一部と推察される「民日」「民」、動詞の赤字に「おる」、形容詞ほかの緑字のなかに「なるい」といった不自然な表示が見られる。これらは日本語は英語などと違って単語間にスペース

^{xi} 公益社団法人日本語教育学会において、日本語教育学の構造化を目指したワーキンググループが中心になって作成。日本語教育と日本語教育研究に携わる人の実際の活動や関心に注目して日本語教育学の姿を描いたと説明されている。
https://www.nkg.or.jp/musubu/contents/kenkyu/20230620_2396954.html

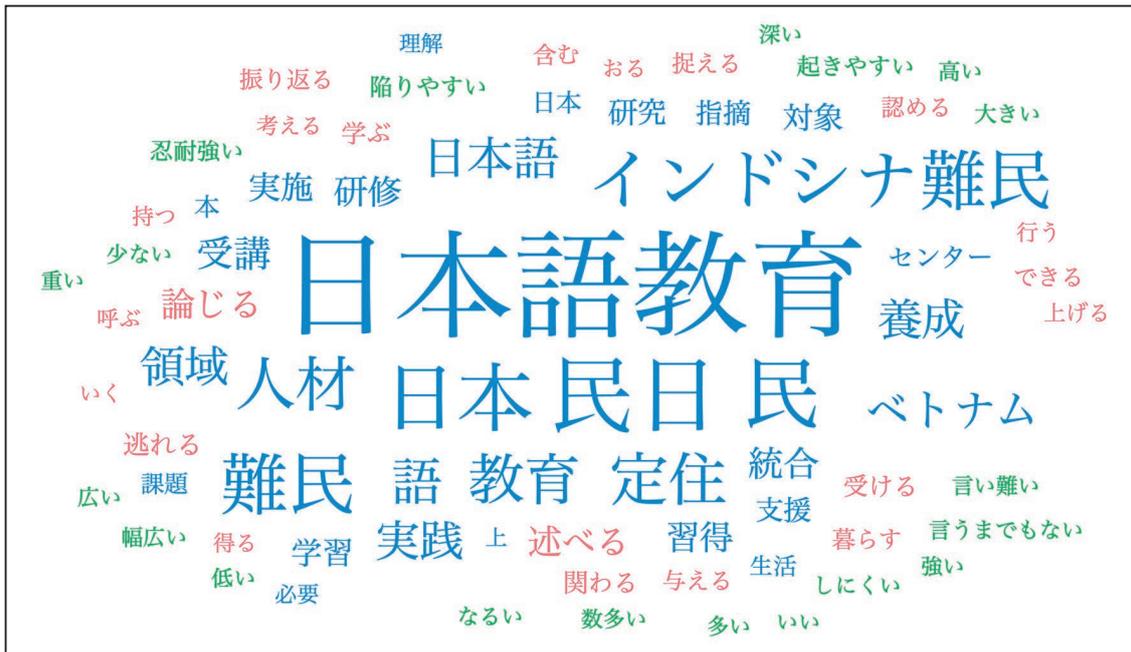


図2 CiNii から抽出した先行研究のタイトルと抄録のワードクラウド

がなく、ワードクラウドで不自然な単語分割が行われることがあるために表れたものと考えられ、分析の対象外とした。

比較的大きな文字で表れている「インドシナ難民」については、以下のような背景がある。日本では1960年代末からベトナム戦争によって生じた孤児などを救済する目的を持つベトナム孤児福祉教育財団が設立されたが、のちにベトナムのほかラオス、カンボジアのインドシナ三国で大量に発生した難民の救済や受け入れを行うようになって、アジア福祉教育財団と名称を変更した。2025年現在、同財団は難民認定者に対する日本語プログラムの運営を国から委託され、同財団の難民事業本部がその実務を担っている。インドシナ難民の受け入れ事業は2005年に終了している^{xii}ため、ワードクラウドで大きな文字で表れている「インドシナ難民」をタイトルや抄録に含む先行研究はインドシナ難民を対象とした日本語教育を論じた吉田・湯川（1983）や、定住促進センター^{xiii}での日本事情教育を扱った西尾（1988）など、検索対象とした1980～2025年の先行研究のうち事業継続中に発表されたものが多くを占めている。

伴野（2023b）が難民受け入れに大きな転換期を迎えたと指摘する2021年に発生したミャンマーの軍事クーデター、2021年のアフガニスタンの政権崩壊、2022年からいままなお続くロシアによるウクライナ侵攻などに端を発するこれまでにない数の難民及び避難民

^{xii} 公益財団法人アジア福祉教育財団「財団の歩み」より。https://www.fweap.or.jp/about/history

^{xiii} 日本での定住希望者への日本語教育、健康管理、就職斡旋を目的とした施設で、姫路定住促進センター（1979-1996）、大和定住促進センター（1980-1998）があったが、のちにインドシナ三国の政情が安定したことなどを理由に閉鎖された（https://www.rhq.gr.jp/outline/p01/）。ワードクラウドに表れた「センター」は定住促進センターを指すと思われる。

を日本でも受け入れてきた実態が検証され、日本語教育の文脈で語られる機会が極めて限定的だったと言えるだろう。

次に名詞に着目すると、AI テキストマイニングでは名詞出現頻度順の上位 15 位は表 2 のようになっている。

表 2 名詞出現頻度上位 15 位

	1	2	3	4	5	6	7
名詞	難民	日本語教育	教育	日本語	語	日本	本
出現頻度	100	73	68	64	50	42	34
	8	9	10	11	12	13	14
定住	実践	領域	研修	対象	研究	インドシナ 難民	ベトナム
出現頻度	33	28	27	27	26	23	22

テキストマイニングで多く表れた名詞は「難民」「日本語教育」などのほか「定住」「実践」「領域」「研修」などになっており、このことからこれまでの難民日本語教育の研究は、難民に対して日本語を教育するという視点が中心の研究や、定住促進に関連する実践研究が中心だったのではないかと考えられる。留学生や外国人就労者や生活者など難民以外の外国人を対象にした就労や進学、JLPT などの日本語試験関連、子供の教育、地域との結びつき、日本語教師、ボランティアなど、日本語教育研究でしばしば扱われるものが研究テーマになることがなかったか、または少なかったのではないだろうか。しかし、今後の難民日本語教育の発展のためには、これまで行われた研究を振り返って整理し、まだあまり研究されていない分野を捉えておくことは重要だと言える。何より、難民申請が認可されて日本に定住して、日本語を習得したり、日本での人生設計を考えたりする人々に関連する研究で得られる知見を積み重ね、新たな難民日本語教育研究はもちろん、他の日本語教育研究の分野にも応用、発展させていくことが必要だろう。

次節では「日本語教育学の俯瞰図」から難民日本語教育の先行研究を整理したい。

3.2 「日本語教育学の俯瞰図」

日本語教育学会では日本語教育学を捉えて全体像を示し、日本語教育の実践との関係を考える議論を踏まえて、2022年に日本語教育学の構造化に関するワーキンググループを組織し、2023年、日本語教育研究を整理して図式化することを目的に図3の「日本語教育学の俯瞰図」をまとめた。



図3 「日本語教育学の俯瞰図」(日本語教育学会 2023)

「日本語教育学の俯瞰図」(以下、俯瞰図)は、日本語教育学は「A. 日本語教育の諸分野」「B. 日本語教育の諸側面」「C. 日本語教育の研究的関心」の3層で構成される。俯瞰図の内側のA層に示された日本語教育の諸分野は、それぞれがB層に示された日本語教育の諸側面を有するとしている。さらに俯瞰図外側のC層に示された研究的関心により、これまで日本語教育の諸分野で多様な研究が行われてきたことが俯瞰できる。

俯瞰図のA層に難民日本語教育または難民の日本語教育等の語は見られない。難民日本語教育をA層に当てはめるとすれば、現在は文化庁の難民に対する日本語教育のみで行われているため、生活者や技能実習・特定技能、介護士候補者・看護師候補者等への日本語教育が含まれる〈A2 各種の分野の日本語教育〉に分類されると思われる。ただし、A層に難民日本語教育を意味する語が示されていないことは、研究分野として十分に確立されていないことを意味すると言える。そして、A層の外側のB層には、〈B1 教育の開発と改善〉〈B2 教育活動〉〈B3 評価〉〈B4 学習者〉〈B5 教師〉〈B6 学習段階〉〈B7 ICT活用の言語教育〉の7つの側面が記されている。以上の2層からなる日本語教育の諸分野・諸側面について、C層では研究的関心が分類されており、これまで〈C1 日本語〉〈C2 言語〉〈C3 コミュニケーション〉〈C4 言語教育〉〈C5 学習と教育〉〈C6 制度・社会・歴史〉〈C7 言語の行使と機能〉〈C8 哲学・思想〉の8分類の研究的関心によって研究が行われてきたことを示す。また、このC1～C8の研究的関心は表3のように4つの事項に関わるとされる。

表3 「日本語教育学の俯瞰図」C層4事項

C層（研究的関心）	関わる事項
〈C1 日本語〉〈C2 言語〉	(1) 日本語教育実践のあらゆる側面の意志決定において援用される基本的な認識に関わる
〈C3 コミュニケーション〉	
〈C4 言語教育〉〈C5 学習と教育〉	
〈C6 制度・社会・歴史〉	(2) 日本語教育というものの捉え方に関わる
〈C7 言語の行使と機能〉	(3) 言語活動の実態に関わる
〈C8 哲学・思想〉	(4) 教育や言語そのものの本質に関わる

以上のように、日本語教育研究は諸分野（A層）、諸側面（B層）に分けられ、多様な研究的関心（C層）に基づいて行われており、それぞれの研究的関心は表3の通り4つの「関わる事項」によって分類できる。これにしたがって難民日本語教育についても先行研究を分類し、これまでの研究動向を捉えることが可能であると考えられる。

3.3 難民日本語研究の分類

本節では3.1でテキストマイニングにかけた37件の先行研究を前節の「日本語教育学の俯瞰図」に沿って分類を行った。

前節で述べたとおり、難民日本語教育はA層の〈A2 各種の分野の日本語教育〉とし、先行研究をB層のB1～B7に分類した。さらに各先行研究についてC層C1～C8の分類を試みた。B層とC層の分類は複数の側面や研究的関心が見られるケースがあったが、分類しにくい場合は研究タイトルにある用語から判断した。その結果が表4である。

表4からはB層の〈B3 評価〉〈B6 学習段階〉〈B7 ICT活用の言語教育〉に分類される先行研究が見られなかったことがわかる。これまで難民日本語教育は文化庁の「難民に対する日本語教育」を除くと地域日本語教室や支援団体等による学習支援等が行われていたが、教育や支援が行われる期間や内容、環境等に制限があるため、B3の評価にかかわる研究やB6の学習段階別の教育の研究は困難であったと思われる。また、B7のICT活用の教育も同様の理由で研究テーマとしにくいのではないかと推察される。その他、B1では難民日本語教育で使用される教材の研究や調査報告が見当たらなかった。文化庁の「難民に対する日本語教育」プログラムで使用している教科書は公開されており、こうした教材の研究も難民日本語教育研究を一步前進させることになるのではないだろうか。表4を作成するまでは難民日本語教育は難民として保護されている人々が対象であるため、B4の学習者を扱った研究が少ないのではないかと考えていたが、B5の教師を扱った研究よりは数が多かった。しかし、B4とB5の先行研究の発表年を見ると、難民認定数が急増した2022年以降の研究で林（2024）が学習者を扱っているが、林が研究対象としたのは新たに増えた難民ではなく、すでに定住しているベトナム難民である。また、教師を研究対象とした2022年以降の研究5件のうち、4件が急増した難民に対する日本語教育の教育や支援を担う人材について論じた伴野（2022,2023a,2024,2025）である。この結果は、

表4 「日本語教育学俯瞰図」による難民日本語教育先行研究の分類

A層	B層	先行研究	C層	事項
A2	B1	伴野 (2023b)	C6	2
		宮下 (2023)	C4	1
		宗田 (2021)	C6	2
		木村・佐伯・人見 (2018)	C4	1
		小松 (2013)	C7	3
		伴野 (2013)	C6	2
		国際日本語普及協会 (2011)	C4	1
		アジア福祉教育財団 (2006)	C6	2
		B2	鈴木 (2025)	C4
	平岡 (2023)		C4	1
	小瀧 (2009)		C7	3
	内藤 (2009)		C4	1
	西尾 (1990)		C4	1
	西尾 (1988)		C4	1
	吉田・湯川 (1983)		C4	1
	B3			
	B4	林 (2024)	C4	1
		松尾 (2021)	C7	3
		長谷部 (2020)	C7	3
		長谷部 (2019)	C2	1
		林 (2018)	C7	3
		岩下・小西 (2017)	C4	1
		瀧尻・植本 (2015)	C7	3
		松田 (2012)	C7	3
		いしもと (2011)	C7	3
		山川 (2006b)	C7	3
		吉本 (2001)	C3	1
		福留・増井 (1997)	C3	1
		関口 (1994)	C4	1
		福留・山元 (1994)	C5	1
		福留・山元 (1993)	C2	1
		上条 (1989)	C2	1
	B5	伴野 (2025)	C4	1
		斉藤・森田・小林 (2024)	C4	1
		伴野 (2024)	C6	2
		伴野 (2023a)	C4	1
		伴野 (2022)	C4	1
山川 (2006a)		C8	4	
B6				
B7				

難民日本語教育に携わる日本語教師を対象とする調査がしにくいことを暗示しているのではないだろうか。

C層では、多文化共生や市民性教育、インクルーシブ教育、継承語、複言語などにかかわるC8に分類された先行研究が山川(2006a)のみであった。俯瞰図の中央には日本語教育学会が理念とする「ことばと共生」が書かれているが、ことばの教育によって難民を含めた外国人住民への社会の理解を促し、一市民として包摂する共生社会を目指すのであれば、C8は今後より研究を深めていかねばならない分野だろう。

3.4 年代別の研究動向

1990年代前後の研究で西尾(1988、1990)はインドシナ難民や中国帰国孤児及びその家族を対象とした日本語教育について調査している。インドシナ難民については大和定住促進センター他で、中国帰国孤児は定着促進センター(所沢)で行われていた日本語教育を詳細に記述し、日本語教育と社会生活の適応を目的としたトレーニングが別立てで行われていたことが記されている。それによると、一つの教室で学ぶ難民の母語がカンボジア語、ラオス語、ベトナム語など複数になることから、媒介語を使わずに日本語で日本語を教える教授法が採られていた。また、まったく異なる文化背景を持つ人々に社会適応のための指導を行う難しさについても触れており、「難民・帰国者の日本語教育が第2言語としての日本語教育の初めての現場となり、その経験と実績が新しい日本語教育の分野を確立することになった」(西尾1990:336)としている。

福留・山元(1993)はインドシナ難民に対する日本語教育と社会生活適応指導の終了後の日本語習得状況にいちやく着目し、インドシナ難民のうち数世帯について1年間追跡調査を行った。インタビュー調査やアンケート調査などから学習成果が定着していなかったり、改善されていなかったりする場面などを記録し、難民に対する日本語教育のカリキュラム・デザインに貢献することを目指した。福留・増井(1997)も、日本語教育と社旗生活適応指導を受けていた国際救援センター退所後のインドシナ難民の日本人とのコミュニケーションについて追跡調査を行っている。

2000年代に入るとベトナム難民を対象にした談話における相づちの研究もおこなわれたが(吉本2001)、研究対象となる難民の確保は容易でなかったことが推察され、同様のテーマの研究は広がっていない。その一方、山川(2006b)ではタイトルに「共生」という語が使用されており、難民条約により保護される立場である難民と日本社会及び日本人との関係性が語られるようになった。内藤(2009)も日本語は難民にとって自己実現のツールとなるものであることについて触れており、日本語を教えて社会生活適応指導を行う難民について、日本語を一人の人間として生きがいを持って生きるための手段と見なす傾向が観察される。2000年代に入ってから、難民と日本語についてこれまでと異なる視点が明確になってきたと言えるのではないだろうか。

こうした傾向は2010年代により明確になっている。松田(2012)は難民の自立に関し

て就労と日本語教育の関連を明示し、保護されてきた難民が保護期間後に必要な就労と日本について述べている。また、伴野（2013）は難民を対象にした公的日本語支援、ボランティアやNPO等による日本語教室などの実践が難民の日本語習得を支えている一方で、これまでの枠組みだけでは十分な支援とは言えないことを指摘した。また、難民の自己表現や居場所の確保、社会統合、周囲とのコミュニケーション等を目指す必要があるとしている。

このほか2010年代には難民の子どもへの言及が見られた（木村・佐伯・人見2018）。当然ながら若い世代の難民が来日後に日本で家族を持ったり、年少者を含む家族を伴って来日したりしたケースを考えると、今後は難民の子どもの教育についても日本語支援や学習支援、進学相談などを複合的に考えていかねばならないだろう。また、こうした支援のほかに「誰も取り残さない」社会（宗田2021）に関連した難民自身のアイデンティティに目を向けた母語教育や、日本語習得に対する日本人側からの評価ではなく、ベトナム難民の視点に立った日本語のあり方（林2024）についても議論された。このように、これまでの研究を概観すると、少なくとも研究者の難民へのまなざしに変化が確認できる。

他方、難民に日本語を教える教師や、ボランティアなど難民を支える人々はどうだろうか。伴野（2022・2023a）は難民日本語教育の実践者である自身を調査対象として、専門家としての自己の成長を記述・可視化することで人材育成などに活用するモデルについて検討し、難民日本語教育の実践者のための参考資料とすることを試みた。現在、国によって行われている難民の日本語プログラムは公開されていないため、実際に教育にあっている日本語教師にアクセスすることも容易ではない。だが、難民日本語教育の特殊性を考えると、こうした数少ないデータを少しずつ積み重ねていくことは必要であろう。

難民日本語研究の動向を概観すると、まだ少数である研究の中で難民自身以外に着目した研究がごく限られていることがわかった。難民に対する日本語その他の支援のあり方や支援プログラムの制度そのものだけでなく、そこに関わっている日本語教師やボランティア、難民と接する地域住民など、難民の周りに存在する「人」に対する研究も今後進めていくべきではないだろうか。

4. まとめ

本研究で対象とした研究を概観すると、難民日本語研究はインドシナ難民への日本語教育に始まり、インドシナ難民の受け入れ終了後は近年の不安定な世界情勢によって発生した多様な背景の難民への日本語教育へと研究の焦点が少しずつ移動してきたように見える。インドシナ難民の高齢化や多様な難民の受け入れによる年少者への教育問題、一人一人の難民の自己実現などの問題が積み重なり、可視化されてきたとも言えるが、それらを対象とした研究はまだ少数である。

今後、ますます複雑化すると予測される世界情勢を考えると、これまでインドシナ、

ミャンマー、アフガニスタン、ウクライナ等から受け入れた難民や避難民とは異なる背景の難民が、日本人の隣人になる可能性は充分にある。西尾（1990）が「新しい日本語教育」と表した難民日本語教育は、複雑に拡大していくものかもしれない。難民日本語教育研究を進めることは、こうした未知の難民への日本語教育にも貢献できるはずである。今後の難民日本語教育研究の課題としては、難民を保護し、そのプライバシーを守りながら、将来新たな難民を受け入れたときのために常に研究を続け、研究対象を広く深く捉えていくことが重要だと言えるだろう。

引用文献

- アジア福祉教育財団（2006）「国際救援センターにおけるインドシナ難民等に対する日本語教育調査研究：平成16・17年度文化庁日本語教育研究委嘱」アジア福祉教育財団難民事業
- いしもとあやこ（2011）「Special Interview 異文化に学ぶ（Vol.3）羽賀友信 長岡市国際交流センター『地球広場』センター長 無気力状態に陥っている難民に生きる希望を与えること。『自立支援』から、国際協力は始まる」『月刊日本語』24(6)、4-7、アルク
- 岩下智美・小西広明（2017）「2015年度第三国定住難民に対する渡日前日本語研修報告」『国際交流基金日本語教育紀要』13、117-124、国際交流基金
- 木村光伸・佐伯奈津子・人見泰弘（2018）「外国人・難民問題にどう取り組むか」『名古屋学院大学論集 社会科学編』55(1)、183-192、名古屋学院大学総合研究所
- 小瀧雅子（2009）「小学校日本語学級と交流活動の記録——難民クラスでの『生活ガイダンス』プログラム実践から」『AJALT』4、76-91、国際日本語普及協会
- 小瀧雅子（2021）「難民のための日本語教育初任教师養成研修講座：オンライン事始め」『AJALT』44、44-46、国際日本語普及協会
- 国際日本語普及協会第三国定住難民事前研修担当グループ（2011）「メラから日本へ 第三国定住難民のための日本語教育——出国前研修プログラムを中心に」『AJALT』34、16-21、国際日本語普及協会
- 小松由美（2013）「第三国定住での難民受け入れと定住支援としての研修についての一考察」『東京外国語大学留学生日本語教育センター論集』39(39)、105-112、東京外国語大学留学生日本語教育センター
- 斉藤知花・森田淳子・小林亜希子（2024）「ロシアによるウクライナ侵略後の日本語教育環境の変化と現状——日本語教師を対象としたアンケート調査を中心に——」『国際日本号研究』4、227-240、東京外国語大学大学院国際日本学研究院
- 鈴木美穂（2025）「日本における難民の現状と日本語学習支援——小規模日本語教室の取り組みと課題」『目白大学外国語学部開設20周年記念論集』341-359、目白大学外国語学部
- 関口明子（1994）「日本定住児童の日本語教育——インドシナ難民児童の多様な言語背景と日本語習得」『日本語教育』83、1-15、日本語教育学会
- 瀧尻明子・植本雅治（2015）「在日ベトナム人高齢者の生活と健康状態に関する研究」『大阪市立大学看護学雑誌』11、11-20、大阪市立大学大学院看護学研究科
- 伴野崇生（2013）「『難民日本語教育』の可能性と課題：難民の権利・尊厳のための日本語学習

- 支援の構想」『難民研究ジャーナル』3、26-43、難民研究フォーラム
- 伴野崇生（2022）「『難民日本語教育』実践者の自己形成と成長——オートエスノグラフィーとAuto-TEMを通じて——」『社会情報研究』3(2)、1-15、先端教育機構
- 伴野崇生（2023a）「難民を対象とした日本語教育実践者の自己形成・成長過程における自己内対話 Auto-TEM 分析結果を基にした『対話的自己』による考察」『東京医科歯科大学教養部研究紀要』53、95-112、東京医科歯科大学教養部
- 伴野崇生（2023b）「難民支援としての日本語教育・難民を対象とした日本語教育」『小出記念日本語教育学会論文集』31、165-191、小出記念日本語教育学会
- 伴野崇生（2024）「『難民等』への日本語教育人材養成の現状と課題：モードの往還による知識生産」立命館大学博士論文
- 伴野崇生（2025）「『難民日本語教育』におけるネガティブ・ケイパビリティの意義と課題」『待遇コミュニケーション研究』22(0)、213-213、待遇コミュニケーション学会
- 内藤真知子（2009）「自己実現の日本語へ—条約難民に対する日本語教育が目指すもの—『声』を受け止め、『声』を支える」『AJALT』32、23-27、国際日本語普及協会
- 西尾珪子（1988）「姫路・大和定住促進センターにおけるインドシナ難民に対する日本事情教育」『日本語教育』65、95-108、日本語教育学会
- 西尾珪子（1990）「難民・帰国者に対する日本語教育」『日本音響学会誌』46(4)、333-338、日本音響学会
- 長谷部美佳（2019）「インドシナ難民の日本語力：日本社会とのつながりから考える」『明治学院大学教養教育センター附属研究所年報』2018、14-14、明治学院大学教養教育センター附属研究所
- 長谷部美佳（2020）「インドシナ難民の定住状況調査をめぐる一考察」『駒沢社会学研究』55、25-49、駒沢大学文学部社会学科
- 林貴哉（2018）「ベトナム人集住地域における複数言語の使用と学習に関する研究」『言語文化教育研究』16(0)、136-156、言語文化教育研究学会
- 林貴哉（2024）「在日ベトナム難民のライフストーリー：職場での言語使用・習得についての語りに焦点を当てて」『日本オーラル・ヒストリー研究』20、256-275、日本オーラル・ヒストリー学会
- 平岡憲人（2023）「ウクライナ学生支援会（JSUS）の挑戦：ウクライナ避難民支援における日本語学校の役割と提案」『海外日本語教育研究』16、4-7、海外日本語教育学会
- 福留伸子・増井世紀子（1997）『筑波大学留学生センター日本語教育論集』12、171-196、筑波大学留学生センター
- 福留伸子・山元啓史（1993）「インドネシア難民の日本語習得の状況とその問題点（教育編）」『日本語教育方法研究会誌』1(1)、26-27、日本語教育方法研究会
- 福留伸子・山元啓史（1994）「日本語学習者のコミュニケーション能力習得の一考察：インドシナ難民の日本語学習環境に関するケーススタディ」『筑波大学留学生センター日本語教育論集』9、185-193、筑波大学留学生センター
- 松尾慎（2021）「対等なパートナーとしての学び合い：ビルマ（ミャンマー）難民との日本語活動」『部落解放』812、91-103

- 松田節子 (2012) 「難民定住者たちの自立への道：就労と日本語教育の三つの提言」『Work & Life：世界の労働』2012(5)、49-53、日本ILO協議会
- 宮下しのぶ (2023) 「理念を守り、前に進む：条約難民・第三国定住難民への日本語教育」『AJALT』46、40-44、国際日本語普及協会
- 宗田勝也 (2021) 「『誰も取り残さない』社会への手がかり」『ボランティア学研究』21(0)、33-38、国際ボランティア学会
- 山川徹 (2006a) 「今、共生の現場で一開かれた社会は来るのだろうか (第6回) インドシナ難民、その過去と現在」『月刊日本語』19(10)、44-47、アルク
- 山川徹 (2006b) 「今、共生の現場で一開かれた社会は来るのだろうか (第7回) 難民を支える立場から」『月刊日本語』19(9)、56-59、アルク
- 吉田弥寿夫・湯川純幸 (1983) 「インドシナ難民に対する日本語教育」『言語生活』376、44-54、筑摩書房
- 吉本優子 (2001) 「定住ベトナム難民における相づち習得の研究——談話展開の観点から」『日本語教育』110、92-100、日本語教育学会
- 外務省「国内における難民の受け入れ」 <https://www.mofa.go.jp/mofaj/gaiko/nanmin/main3.html> (2025年9月28日最終閲覧)
- 公益財団法人アジア福祉教育財団「財団の歩み」 <https://www.fweap.or.jp/about/history> (2025年9月28日最終閲覧)
- 公益財団法人アジア福祉教育財団難民事業本部 (RHQ) 「条約難民とその家族のための定住支援プログラム 2025年10月開講」 <http://www.rhq.gr.jp/wp/wp-content/uploads/2025/07/leaflet2025-oct-convention-refugees-jp.pdf> (2026年1月12日最終閲覧)
- 出入国在留管理庁「難民認定手続・補完的保護対象者認定手続」 <https://www.moj.go.jp/isa/refugee/procedures/16-6.html> (2026年1月13日最終閲覧)
- 日本語教育学会 (2023) 「『日本語教育学の俯瞰図』の解説と活用法—日本語教育と日本語教育研究の相互活性的なダイナミクスの促進をめざして—」 https://www.nkg.or.jp/musubu/contents/kenkyu/20230620_2396954.html (2026年1月13日最終閲覧)
- 文化庁「条約難民及び第三国定住難民に対する日本語教育事業委託実施要領」 https://www.bunka.go.jp/seisaku/kokugo_nihongo/kyoiku/nanmin_nihongokyoiku/pdf/r1392633_01.pdf (2025年9月26日最終閲覧)
- 文化庁「難民対策について」 https://www.bunka.go.jp/seisaku/kokugo_nihongo/kyoiku/nanmin_nihongokyoiku/pdf/nanmin_kaigi_1.pdf (2025年9月26日最終閲覧)
- 文化庁「難民に対する日本語教育」 https://www.bunka.go.jp/seisaku/kokugo_nihongo/kyoiku/nanmin_nihongokyoiku/ (2025年9月26日最終閲覧)
- 法務省「我が国における難民保護の状況等」 <https://www.moj.go.jp/isa/content/001434641.pdf> (2025年9月26日最終閲覧)
- UNHCR「難民の地位に関する1951年の条約」 <https://www.unhcr.org/jp/treaty-1951> (2025年9月26日最終閲覧)

Research Trends in Japanese Language Education for Refugees

Naomi KATAYAMA

(Center for Global Education and Exchange, Toyo University)

Abstract

In recent years, despite Japan's consistently low refugee recognition rate, the number of refugee applications has shown an upward trend. The number of recognized refugees has also increased, influenced by political upheavals in countries such as Myanmar and Afghanistan. Recognized refugees are provided with Japanese language programs and support to promote independent living; however, because Japanese language education is conducted in closed environments intended to protect refugees, research in this area has remained limited.

In this study, previous research on Japanese language education for refugees was extracted through keyword searches in CiNii (Scholarly and Academic Information Navigator). Using data from the selected studies, we conducted a visualization analysis employing text mining techniques.

Furthermore, by applying the "Panoramic Framework of Japanese Language Education Studies," we organized the research domains addressed in prior studies and sought to identify themes that have received limited attention in research on Japanese language education for refugees.

Key words :

Japanese language education for refugees, text mining, word cloud, panoramic framework of Japanese language education studies

高度外国人材としての留学生の インターンシップの課題と展望

渡 部 裕 子 (東洋大学国際教育センター)

要旨

留学生は将来的な高度外国人材として期待され、国内の就職・定着の促進のために文部科学省によって「留学生就職促進教育プログラム」認定制度が実施されている。しかし、プログラムの柱の1つであるインターンシップ実施には様々な課題があり、プログラム実施大学は2024年度までの5年間で大学全体の3%にとどまっている。

本稿は留学生のインターンシップに焦点を当て、インターンシップのあり方を考える一助にすることを目的として、先行研究を概観し、留学生のインターンシップの課題を明らかにした。また、留学生や関係機関に求められる具体的な到達目標や実施プロセスについて、準備、実施、振り返りの3つの段階から洗い出しを行った。日本人学生等の国内学生に比べると、キャリア探索の過程で留学生ならではの要素である異文化理解や日本語の課題が加わり、プログラムの開発・運営に関わる大学や企業が今後検討すべき内容が大枠で可視化できた。

キーワード：

高度外国人材、留学生、インターンシップ、留学生就職促進教育プログラム

目次

1. 研究の背景と目的
2. 大学生等のインターンシップの先行研究
 - 2.1 海外のインターンシップ事例
 - 2.2 インターンシップに関する国の方針
 - 2.3 大学生のインターンシップの効果に関わる先行研究
3. 留学生のインターンシップの先行研究
4. 留学生に特化したインターンシップの役割と課題
 - 4.1 留学生に特化したインターンシップの役割
 - 4.2 留学生に特化したインターンシップの課題
5. 留学生のインターンシップにおける到達目標と実施プロセスの枠組み
6. 今後の課題

1. 研究の背景と目的

高度外国人材の活用については、1999年の『第9次雇用対策基本計画』で経済社会の活性化と国際化を図るために、専門的・技術的分野の外国人労働者の受入れを積極的に推進することが示された。2016年の『日本再興戦略』では留学生の国内就職を3割から5割に上げるという具体的な数値目標が示され、そのための施策として、文部科学省は2017年から「キャリア教育」「日本語教育」「インターンシップ」の3つを柱にした『留学生就職促進プログラム』を開始し、選定された17大学によって先進的な取り組みが行なわれた。さらに2020年からはより多くの大学を対象に、『留学生就職促進プログラム』と同様に3つの教育内容を柱にした『留学生就職促進教育プログラム（以下「留学生就職促進教育P」）』認定制度が開始され、2024年度までに25大学が認定されている。

これらのプログラムの3つの柱である「キャリア教育」「日本語教育（ビジネス日本語含む）」「インターンシップ」の具体的な到達目標や内容の要点は何か、文部科学省の『留学生就職促進教育プログラム認定制度実施要項』を基に検討する。

まず、「日本語教育」については「取組に参画する各業界の需要や履修者の専攻分野、想定するキャリアパス等を踏まえ、履修者が在学中に身に付けるべきビジネスコミュニケーション能力や日本語の能力水準を明確に設定し、対外的に明示できるようにすること。」とある。つまり、ここでいう日本語教育は、前提として学生が自分のキャリア探索を行い、目指す業界等がある程度絞った上で、または検討しながら各業界における企業等が求めるコミュニケーションができるように、目標設定をしなければならない。また、大学はこれらの教育内容について、在校生はもちろん、対外的にもわかりやすく伝える必要があることが読み取れる。

次に、「キャリア教育」については、「企業と外国人留学生の間に生じるミスマッチを防ぐため、外国人留学生に企業文化を講義することが必要不可欠である。そのため、単に一般的な企業文化の講習にとどまらず、より実践的なキャリア教育を施すことにより、学生のキャリアプランをしっかりとイメージさせる取組であること。」とされており、これが企業文化の講義を含めた留学生に特化したキャリア教育を指し、キャリア教育を通じて、日本での就労をイメージしたキャリアプランを描けることが目的だということが理解できる。ただし、企業文化とされる内容が抽象的で幅広く、具体的にどんな内容が不可欠なのか、優先度が高いものは何かというのはここでは言及されていない。また、より実践的なキャリア教育というのが具体的にどんな内容を指すのかも明らかでない。

最後に、「インターンシップ」については、「インターンシップを通じて、日本企業・組織における働き方や慣行等を経験し、先に掲げた「キャリア教育」の内容の一層の理解・定着を図る機会と位置付け、国内企業等における2週間程度以上のインターンシップを実施」するとしている。まず、「キャリア教育の一層の理解・定着を図る機会と位置付ける」ことが目的とされており、キャリア教育の内容はインターンシップの準備段階としての役

割もあることが理解できる。インターンシップの教育的効果が、日本語を使った職業体験を通じて、「日本語教育」「キャリア教育」を有機的に結び付ける総合的な教育実践の場として期待されていると考えられる。

しかしながら、前述したように留学生に特化した「キャリア教育」について明示的な内容がないため、現行のインターンシップの内容や到達目標はキャリア教育との有機的な結びつきによるというより、それぞれの企業が対応できうる範囲でプログラムを実施していると推察される。実際は様々であろう日本企業・組織における働き方や慣行の経験という漠然とした内容も、一律化したプログラム開発が難しいインターンシップの内容を示唆していると考えられる。

大学内でカリキュラム開発及び教育実践ができる「日本語教育」「キャリア教育」と異なり、インターンシップは企業側の実施協力なくして成立しない。加えて、言語・文化が異なる留学生が対象の場合、企業にとって国内学生（日本人学生及び一般入試で入学した外国籍学生）よりも受入れのハードルが高くなる可能性がある。

これらの内容を柱に留学生の国内就職促進のために創設された「留学生就職促進教育P」の認定状況を見ると、実施大学数は2024年度までの5年間で大学全体の3%に過ぎず、増加傾向は限定的である。この原因の一つとして、前述したインターンシップにおける内容の不明瞭さ、企業と連携して実施する困難さがあるのではないだろうか。

そこで、本稿では、留学生のインターンシップに焦点を当て、これまでのインターンシップに関わる先行研究を概観し、その役割や課題を明らかにする。さらに、留学生のインターンシップにおける具体的な到達目標について考察し、大学と企業がどのように連携して実施すべきかを検討する。

2. 大学生等のインターンシップの先行研究

インターンシップは欧米を中心に海外では古くから行われてきているため、まず海外のインターンシップ事例と日本のインターンシップを比較する。次に、留学生対象のものより研究が先行する大学生等の国内学生を対象としたインターンシップの研究について概観する。国内学生対象のインターンシップは、産業界の要請によって、国が方針を示すことでインターンシップ実施を促進してきた側面が強い。そのため、この留学生に特化しないインターンシップについては、まず、これまでの国の方針を概観する。

2.1 海外のインターンシップ事例

リクルートワークス研究所の調査報告（2015）によれば、アメリカで実施されているインターンシップの多くは正社員の採用・就職の手段となっており、6～12週間のものが多い。また、大学での学習と職場での3～6か月の実習を有機的に結びつけるCo-op（産学協同教育）と呼ばれる100年以上の歴史を有するプログラムもある。イギリスでは

2008年に起こった経済危機後の2009年以降、採用や就職への危機感から2年生の夏季休暇における4～8週間の有給インターンシップモデルが確立したと報告されている。オーストラリアでも夏季休暇中に8～10週間実施される有償プログラムが一般的とあり、いずれの国も1～2週間という短期間でインターンシップを実施する日本の状況とは異なることを指摘している。新卒一括採用により入社後の教育研修の実施が一般的に行われている日本では、インターンシップは就職活動におけるエントリーの一環として、採用の入口の役割を果たす場合が多く、直接的な採用手段とはなっていない。これに対し、欧米諸国ではインターンシップが採用とより密接に結び付いており、こうした位置づけの違いが実施期間の長さの差異に関連していると考えられる。そのため、海外のインターンシップ事例を、その目的や期間の点において、日本のインターンシップにそのまま適用することには慎重な検討が必要である。

2.2 インターンシップに関する国の方針

大学生等へのインターンシップに関して出された国の方針は、これまで、以下の3つの段階を経て、現在に至っている。

(1) 『インターンシップの推進に当たっての基本的な考え方』(1997年)

産業界のニーズに応える産学連携による人材育成の一形態として、1997年に『インターンシップの推進に当たっての基本的な考え方』(以下『基本的な考え方』)が国の方針として明示された。その意義として、大学等の教育機関にとっては「アカデミックな教育研究と社会での実地の体験を結びつけることが可能となり、大学等における教育内容・方法の改善・充実につながる。」とある。一方、学生にとっては「自己の職業適性や将来設計について考える機会となり、主体的な職業選択や高い職業意識の育成が図られる」とあり、これにより「就職後の職場への適応力や定着率の向上」、「企業等の現場において就業体験を積み、専門分野における高度な知識・技術に触れながら実務能力を高めることは、自主的に考え行動できる人材の育成」につながると述べている。

また、企業等受入れ側にとっての意義として、「インターンシップによって学生が得る成果は、就職後の企業等において実践的な能力として発揮されるものであり、インターンシップの普及は実社会への適応能力のより高い実践的な人材の育成につながる」「インターンシップの実施を通じて大学等と連携を図ることにより、大学等に新たな産業分野の動向を踏まえた産業界等のニーズを伝えることができ、大学等の教育にこれを反映させていくことにつながる」「大学等と企業等の接点が増えることにより、相互の情報の発進・受信の促進につながり、企業等の実態について学生の理解を促す一つの契機になる。これについては、特に中小企業やベンチャー企業等にとって意義が大きいものと思われる」とある。産業界と大学が連携することで大学の教育内容を将来的な人材育成に活かすことができ、知名度が高い大企業に向きがちな学生の関心の幅を広げるきっかけとなるため、特

に中小企業やベンチャー企業にとっての採用に有効であることを示している。さらに、留意事項として、「産学連携による人材育成のための教育活動であり、連携した効果的なプログラムの開発は重要である」との指摘がある。

この方針により、単位認定のインターンシップの実施大学がその後の約10年間で23.7%（1998年）から67.7%（2007年）へと3倍近く増えている。

(2) 「インターンシップの普及及び質的充実のための推進方策について」（2013）

その後、国際競争力強化の観点から、『日本再興戦略』（2013）で「我が国の将来を担う若者全てがその能力を存分に伸ばし、世界に勝てる若者を育てることの重要性に鑑み、インターンシップに参加する学生数についての目標設定や、キャリア教育から就職まで一貫して支援する体制の強化、インターンシップ活用の推進等」が提言された。それを受けて2013年に出された『「インターンシップの普及及び質的充実のための推進方策について」意見のとりまとめ』（以下『推進方策』）では、インターンシップの課題として受入れ企業数の不足、大企業志向の学生と受入れ先とのミスマッチ、受入れ企業拡大のための専門人材の不在、企業任せになっており大学側の教育的支援が不十分である点が指摘されている。

その対策として、大学側から企業等に対し、教育効果の高いインターンシップ・プログラム設計や運営方法を積極的に提案していく必要があること、外部人材のみに依存するのは適当ではなく、学部・研究科等の組織の枠を超えた学内の連携・協力体制が不可欠であることを指摘している。また、事前事後教育を体系化し、一連のプロセスとしての教育プログラムを開発していく必要があると述べている。さらに、学生に対してキャリア教育から就職まで一貫して支援する体制を強化するために、企業と大学・学生の間にとって質の高いインターンシップのプロジェクト設計が可能なコーディネーターの役割の重要性を指摘している。

留学生の観点から見ると、2013年の『推進方策』で特筆すべき点は、国際競争力強化を背景にしているため、対策として「留学生のインターンシップ」という項目が登場したことである。しかしながら、内容は「優秀な外国人留学生を確保するため、日本企業／日系企業による外国人留学生を対象としたインターンシップの実施を促進する必要がある。」という記述のみで、留学生のインターンシップが国内学生向けのインターンシップと具体的にどのように異なるのかには言及されていない。

尚、リクルートキャリア就職未来研究所（2017）の「平成8年度『先進的の大学改革推進委託事業』インターンシップ推進のための課題及び具体的効果・有用性に関する調査研究報告書」の学生に対する調査結果によれば、インターンシップの参加期間は「1日」が47%で最も多く、続いて「2日以上、1週間未満（2～5営業日）」44%と、短期間ながらも実施期間に幅があった。その内容は「企業等の見学」（54.6%）が最も高く、参加期間で最も多かった「1日」で実施できる内容として合理的であると考えられる。

(3) 「学生のキャリア形成支援に係る産学協働の取組み」(2022年)

インターンシップの実施自体は増えたものの、前述のように所要時間や内容が様々混在する中で、「インターンシップ」という名称が使われているのが実態だった。これを受けて、令和4年(2022年)には、「学生のキャリア形成支援に係る産学協働の取組み」が次の四つの類型に整理され、そのうちタイプ3及びタイプ4がインターンシップであると明記された。

(学生のキャリア形成支援に係る産学協働の取組の四つの類型1)

タイプ1 オープン・カンパニー

タイプ2 キャリア教育

タイプ3 汎用型能力・専門活用型インターンシップ

タイプ4 高度専門型インターンシップ(試行)

タイプ3・タイプ4の大学等のインターンシップの定義は、「学生がその仕事に就く能力が自らに備わっているかどうか(自らがその仕事で通用するかどうか)を見極めることを目的に、自らの専攻を含む関心分野や将来のキャリアに関連した就業体験(企業の実務を経験すること)を行う活動(但し、学生の学修段階に応じて具体的内容は異なる)」とある。このインターンシップの定義から、現状の自身の能力やスキルを把握することでキャリア探索を進めるために、専門や興味・関心に基づいて就業体験を行う活動であると理解できる。それを踏まえれば、インターンシップの内容は専門や興味・関心に応じて様々であり、準備段階としてのキャリア教育では、自らインターンシップ先の選択ができるようになることを目的としてキャリア探索を進めることが必要であると考えられる。

2.3 大学生のインターンシップの効果に関わる先行研究

前節で取り上げた国の方針ではインターンシップの意義や目的は具体的に示されているが、では、その効果についてはどうだろうか。尚、本節で取り上げる先行研究は特に留学生に対する言及はなく、少なくとも、留学生に特化した部分が該当する内容でなく、留学生も包括した大学生のインターンシップ全般に該当する内容と考えられる。

インターンシップの教育的効果について、松高(2021)は先行研究を分析した結果、「職業・就労意識の涵養」については効果が確認でき、「職業生活への移行支援」については明確ではなく、「学習意欲・深化」については確認できないという結論を示している。課題として、教育目的・教育課程への位置づけが弱く、専門教育との結びつきが薄く、職業意識・就労意識といった意識変容に重点が置かれていること、学生、企業、大学間で認識の乖離が大きく、目的の共有も図られておらず、教育プログラムとして成立していない可能性も指摘している。2013年の『推進策』で指摘されている課題と重なりがあり、あまり改善が進んでいないことが示唆されている。

初見他(2021a)は大学生のインターンシップの効果を、自分の将来・キャリアが明確になることを示す「キャリアの焦点化」と、自分の将来・キャリアについて可能性が広

がる「キャリアの展望化」の観点から調査した。その結果、文系の大学生・大学院生のキャリアの焦点化を促進するが、理系の大学生・大学院生では逆にキャリアの焦点化を抑制する可能性があること、また、文理を問わず、キャリアの展望化を促進する効果を持たないことを報告している。

初見他(2021b)は「インターンシップに参加して自分の中に生まれた変化について」と題して書かれた自由記述をテキスト分析し、インターンシップの効果を測る尺度開発を行っている。これまでキャリア探索は「自己探索」(self-exploration、自己について考え、評価を行う)と、「環境探索」(environment exploration、職業に関する情報を得る)の二つの側面があるとされた(Stumpf et al.1983)が、初見他は、この2つの側面だけでなく、効果因子として①キャリアの焦点化、②キャリアの展望化、③人的ネットワークの認知、④就労意欲、⑤自己理解の5つが見られることを明らかにした。「自己理解」は「自己探索」に該当し、「キャリアの焦点化・展望化」は「環境探索」にあたり、「③人的ネットワークの認知、④就労意欲」を他のキャリア教育には見られないインターンシップ特有の効果としている。

3. 留学生のインターンシップの先行研究

本章では留学生に特化したインターンシップの先行研究について概観する。前述の通り、国の方針として明示されている留学生に特化したインターンシップの具体的な方針というものは見られない。『留学生就職促進教育プログラム認定制度実施要項』によれば、言語や文化の側面から実践的な職業研修は必要度が高いということは理解できるが、明記しないことによって、国内学生向けの1997年の『基本的考え方』、2013年の『推進方策』に示されているキャリア探索の教育の一環であるとの方針と変わらないことを示唆しているのではないかと考えられる。

国内学生向けのインターンシップと異なる留学生に特化したインターンシップの必要性や意義とは何かについて、先行研究から探る。

守屋(2014)は留学生に特化したインターンシップが必要な理由として、日本人向けのインターンシップに参加する場合、日本人として理解できることであれば前提として語られないため、インターンシップを通して異文化理解が深まらず、疑問がより拡大する恐れがあることを指摘している。そのため、異文化理解の差を埋めるためには日本企業による外国人留学生向けのインターンシップが必要であると述べている。

横須賀(2020)は留学生にとってのインターンシップは「企業現場の状況に埋め込まれた言葉や文化と結びつき、過去からつながる自分を新たな社会文脈の中に位置づけながら、未来の人生を歩む自分を探究する場である」とし、言葉と文化の理解や習得により、過去から将来につながる自己概念に関わる効果について述べている。

また、地域性や専門性を活かしたインターンシップの実践報告が複数ある。

藤巻（2022）は産官学連携プラットフォームの構築により、インターンシップを就職に結びつけた実践を報告している。インターンシップは留学生の就業体験のみならず、留学生の受け入れに不安を持つ企業にとっても協業体験できる機会であること、留学生がインターンシップに参加することで中小企業の精度の高い企業情報が把握でき、結果的に適切なマッチングにも生かされることを述べている。

泉谷他（2023）はインターンシップを含んだ外国人留学生と日本人学生、企業人がチームワークで課題に取り組む授業を行った結果、参加者の自身の変化について、異なる意見や価値観の違いによる活動の行き詰まりや、コミュニケーション上での葛藤を乗り越えた経験を通して、自己認識を深めたり、関係構築力を獲得する「周囲を巻き込む人間関係力」の高まりを示唆する記述が多かったことが報告されている。

一方、外国人向けインターンシップの課題として、栗原（2024）は、①インターンシップ先確保の難しさ、②求められる日本語力の高さ、③教育機関側がインターンシップ内容を把握することの難しさ、④インターンシップ内容を把握できないことによる評価の難しさを挙げている。②の日本語力の問題は留学生ならではの課題であり、入学時からの計画的な日本語教育カリキュラムが必要になってくる。③と④のインターンシップ内容の把握とそれに伴う評価の難しさは、2013年の『推進方策』でも指摘されている大学の関与の問題であると言える。これらを克服するための事例として課題解決型プログラムを実施することにより、教育機関側が内容を主導できる実践例を報告している。

結城（2025）は大学での留学生への就職支援のための教育プログラムの効果について、受講科目の有無と就職率の関係について調査している。その結果、ビジネス日本語・キャリア教育・インターンシップの全ての科目を受講した群が就職率・地域への就職率ともに高く、その次に一部科目受講の受講者群、全く受講しなかった群が一番就職率が低かったことを報告している。「ビジネス日本語」等コミュニケーションの支援だけでは十分ではなく、「キャリア教育」「中長期インターンシップ」が有機的に組み合わせられて効果が生れると述べている。ただし、具体的な効果については現段階ではグッドプラクティスの紹介にとどまっているため、効果検証の精度を上げる必要性を指摘している。

4. 留学生に特化したインターンシップの役割と課題

これまで概観した先行研究から、留学生のインターンシップの役割と実施する上での課題をまとめる。

4.1 留学生に特化したインターンシップの役割

守屋（2014）の指摘する留学生に特化したインターンシップだからこそ異文化理解が深まるという点、横須賀（2020）の企業の現場における文脈の中でこそ言葉や文化が習得でき、将来のキャリア探索の場となるという指摘を踏まえると、インターンシップの役割と

して以下の3つが考えられる。

①キャリア教育と関連したインターンシップの実践による、留学生のキャリアや日本での就労に関わる意識の涵養の促進

就職や生活など日本でキャリア形成を行うことに対して、実際の職場はどうか、自分はやっていけそうなのか、何が足りないのかといった現実的な自身への評価と判断を促進させることができる

②インターンシップ先での日本語コミュニケーションの実践

講義中心の大学生活においては、日本語は主に理解能力が求められる場面が多く、日常的に日本人学生のコミュニティに接している留学生を除けば、長時間にわたって日本語で課題を遂行する機会は限られている。この点でインターンシップは日本語使用を伴う実践的な学習の場と位置付けられる。

③職場での異文化理解のギャップを埋める

国内学生にとって暗黙の了解として共有されていることは、わざわざ取り上げることがないため、国内学生向けのインターンシップは、留学生にとって文化的理解が十分に進まない可能性がある。留学生に特化した内容にすることで、異文化理解のギャップに焦点を当てた設計が可能になる。

4.2 留学生に特化したインターンシップの課題

藤巻(2022)や栗原(2024)の指摘に加えて、2013年の『推進方策』で示す課題は、留学生のインターンシップにおいて、より顕著に現れる可能性があり、以下について検討する必要がある。

1) インターンシップ受入れ企業の拡大に向けた検討事項

- ①留学生のインターンシップ受入れは、企業にどんなメリットがあるのか
- ②インターンシップのプログラム開発を企業と協働でどのように行うか

2) 専門人材の養成

文部科学省(2017)は「大学等におけるインターンシップの推進に係る専門人材の育成・配置について」で専門人材に必要な要素として以下の4点を挙げている。①「大学等の教育目的に応じて、学生が身に付けるべき知識・能力の修得のためのプログラムの構築・改善」、②「学生の学習意欲の喚起や専門性の深化・向上、職業意識の育成」、③「教職協働による教育的効果の高いインターンシップの企画・立案・運営・評価」、④「企業等との協働による教育的効果の高いインターンシップの企画・立案・運営・評価」である。

留学生に特化したインターンシップの専門人材の役割として、上記に加えて、⑤「日本の職場における異文化理解を加味したプログラム開発」、⑥「インターンシップでの業務・課題遂行のための日本語コミュニケーションの養成」を含む、より広い観点から総合的にプログラムの開発、教育、実施運営を図る必要があると考えられる。

3) 効果検証をどう行うか

効果の検証について、少なくとも下記の4つの観点からの検証が必要だと考える。

①留学生の就職促進：卒業・修了後の国内就職に結びついているか

②留学生のキャリアや就労に関わる意識の涵養

インターンシップを通じてキャリア探索が進み、就労意識は高まったか

③日本の職場における異文化理解

企業ごとの個性差はあるものの、職場において求められる態度や姿勢が理解でき、自分なりに適切な振る舞いを判断し実行できたか。加えて、企業側にとっても、文化的に異なる留学生を迎えるための意識の涵養と具体的な対応がとれたか。

④職場で必要になる日本語コミュニケーション力

インターンシップ先での課題遂行が可能だったか。例えば、指示理解が不十分な場合、日本語で不明な点をきちんと確認する、自分だけで判断せずに相談するといった職場での業務遂行を可能にするコミュニケーション力を重視する。

5. 留学生のインターンシップにおける到達目標と実施プロセスの枠組み

先行研究から概観した留学生に特化したインターンシップの意義や課題に基づき、準備・実施・振り返りという各段階に応じて求められる具体的な到達目標と、それに伴う大学及び受入れ企業等による実施プロセスの図式化を試みたのが、図1の「インターンシップの段階別到達目標とそれに伴う大学・受け入れ企業等の取り組み」である。

留学生の到達目標については、自己探索、環境探索というキャリア探索項目に加えて、日本語、異文化理解という観点から検討した。尚、図1にある「▷」印は国内学生と共通する要素で、「▶」印は留学生に特化して求められる要素を示す。また、到達目標に応じた大学等と受入れ企業の取り組みについて個々の枠に、枠の重なり部分には連携して取り組む内容について示した。

まず、インターンシップの準備段階にあたっては、インターンシップ先を検討するために、留学生という特性を含めて自身の強みや関心を理解し、日本の就活や業界・企業研究について、自国や他国との関係も含めて理解する必要がある。また、日本の組織や雇用慣行、働き方への理解、業務遂行に求められる日本語力も重要である。大学等はそれらの内容をカリキュラム化した教育実践を行うとともに、インターンシップ実施に向けてコーディネーターの配置など実施体制を構築し、留学生ニーズの高い受入れ先の開拓を行う。受入れ企業等も主担者の取り決めや社内調整など受入れ体制を構築する。その後、連携してプログラム開発や参加者の選考、事前学習の実施を行う。

次に実施段階では、準備段階でキャリアの焦点化により選定したインターンシップ先での経験を通して、自身のキャリア観や興味・関心にどうマッチするか、強みをどう活かせる

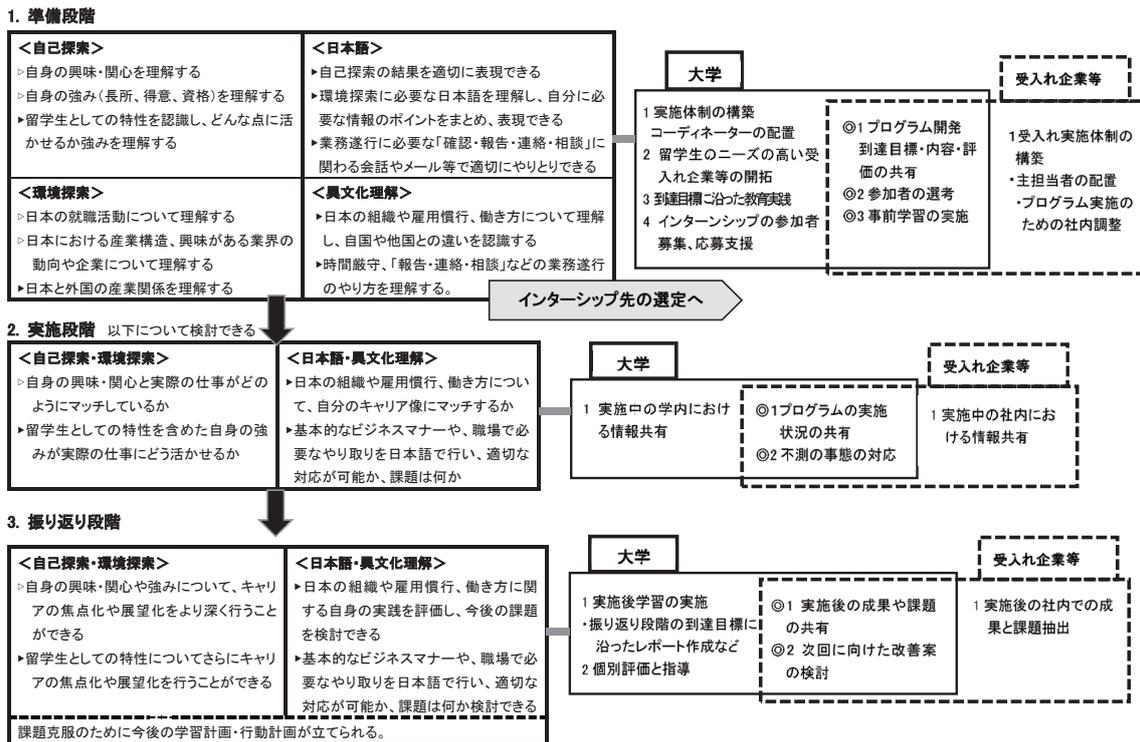


図1 インターンシップの段階別到達目標とそれに伴う大学・受入れ企業等の取り組み

るか検討を行う。また、日本語や異文化理解の面で、実践を通して自身の能力で対応が可能か、課題は何かを検討する機会となる。大学等と受入れ企業等はプログラムの実施状況を共有し、不測の事態に対応できる体制をとる必要があるため、それぞれの学内または社内での情報共有を密にする必要がある。

インターンシップ後は、自身を評価し、キャリア探索を促進し、課題を検討して今後の学習計画・行動計画につなげるための振り返りの段階である。大学等と受入れ企業等は連携して、プログラム内容や実施体制の面から、成果や課題、次回に向けた改善策を検討する。

6. 今後の課題

本稿では、留学生のインターンシップに求められる到達目標を、準備・実施・振り返りの三段階から整理し、これらの実践に向けて大学等と企業等がどのように連携して取り組むことが求められるのかについて、図1を用いて整理・提示した。しかし、本稿で示した内容は枠組みの提示にとどまっており、各到達目標に対する実際の習得状況や、インターンシップ参加が留学生に与える効果についての実証的検討が必要である。

今後は、大学と企業の連携のもと、プログラムの開発から実施後の評価に至るまで、その効果を実証的に分析し、具体的な課題を明らかにしていく必要がある。また、大学等の

課題として、留学生の採用意向も留学生側のニーズも高い企業について、景気動向等時代的背景や、留学生の属性の違いを踏まえて検討を行い、今後の企業開拓につなげていくことが挙げられる。

<引用文献>

- ・厚生労働省（1999年）『第9次雇用対策基本計画』
<https://www.mhlw.go.jp/topics/2002/07/dl/tp0711-1a1.pdf>（2025年9月20日閲覧）
- ・内閣府『日本再興戦略2016』
https://www5.cao.go.jp/keizai-shimon/kaigi/minutes/2016/0602/shiryo_04-1.pdf
（2025年9月20日閲覧）
- ・文部科学省「『留学生就職促進プログラム』選定大学の取組状況」
https://www.mext.go.jp/a_menu/koutou/ryugaku/1394574.htm（2025年9月20日閲覧）
- ・文部科学省「留学生就職促進教育プログラム認定制度実施要項」
<https://www.mext.go.jp/content/000368268.pdf>（2025年9月20日閲覧）
- ・リクルートワークス研究所（2015）「海外におけるインターンシップ最新事情1～14」
https://www.works-i.com/research/labour/item/160322_Internship_jp.pdf
（2026年1月14日閲覧）
- ・文部科学省・厚生労働省・経済産業省（平成9年/1997年）
「インターンシップ推進に当たっての基本的考え方」
https://www.mext.go.jp/component/a_menu/education/detail/_icsFiles/afielddfile/2014/04/18/1346604_01.pdf（2025年9月20日閲覧）
- ・内閣官房『日本再興戦略2013』
https://www.cas.go.jp/jp/seisaku/seicho/pdf/saikou_jpn.pdf（2025年9月20日閲覧）
- ・体系的なキャリア教育・職業教育の推進に向けたインターンシップの更なる充実に関する調査研究協力者会議（平成25年/2013年）
『インターンシップの普及及び質的充実のための推進方策について意見のとりまとめ』
https://www.mext.go.jp/b_menu/shingi/chousa/koutou/055/attach/1338000.htm
（2025年9月20日閲覧）
- ・リクルートキャリア就職みらい研究所（2017）
文部科学省「平成28年度先導的大学改革推進委託事業『インターンシップ推進のための課題及び具体的効果・有用性に関する調査研究』」
https://www.mext.go.jp/a_menu/koutou/itaku/1398414.htm（2025年9月20日閲覧）
- ・文部科学省・厚生労働省・経済産業省（令和4年/2022年）
「インターンシップを始めとする学生のキャリア形成支援に係る取組の推進に当たっての基本的考え方」
<https://www.mhlw.go.jp/content/11800000/000949684.pdf>（2025年9月20日閲覧）
- ・松高政（2021）「大学教育としてのインターンシップの現状と課題」
『日本労働研究雑誌』63（8）pp16-30 労働政策研究・研修機構
- ・初見康行・坂爪洋美・梅崎修（2021a）「多様なインターンシップ経験と効果の一考察」

- 『日本労働研究雑誌』63(8) pp41-57 労働政策研究・研修機構
- ・ 初見康行・梅崎修・坂爪洋美 (2021b) 「大学生のインターンシップ効果測定尺度の開発－テキスト分析とパネルデータによる実証研究－」『日本労務学会誌』Vol.21 No.3 pp18-42 日本労務学会
 - ・ Stumpf, S. A. A., Colarelli, S. M. & Hartman, K. (1983). Development of the Career Exploration Survey. *Journal of Vocational Behavior*, 22 (2), 191-226.
 - ・ 守屋貴司 (2014) 「日本における外国人留学生とインターンシップへの一考察－立命館大学のコーオプ演習・PBLの紹介」ウェブマガジン『留学交流』2014年8月号 Vol.41 日本学生支援機構
 - ・ 横須賀柳子 (2020) 「外国人留学生のキャリア形成－インターンシップ参加の実態調査と事例から」『日本語教育』175号 pp55-64 日本語教育学会
 - ・ 藤巻義博 (2022) 「新たな留学生向けインターンシップ：留学生に特化した高度なインターンシップの設計開発運営と産官学連携のプラットフォーム構築による地域企業への就職・採用の取組み」『静岡大学国際連携推進機構紀要』第4号 pp17-34 静岡大学国際連携推進機構
 - ・ 泉谷道子・小林修・伊月知子 (2023) 「外国人留学生・日本人学生・社会人協働PBL型インターンシップの効果（変革マインドセットの醸成に着目して）」『グローバル人材育成教育研究』11(1) pp2-11 グローバル人材育成教育学会
 - ・ 栗原由加 (2024) 「文系外国人留学生のためのインターンシップ実施方法の提案－教育現場と企業の事情に合わせた連携、実施方法の工夫－」『神戸学院大学グローバル・コミュニケーション学会紀要』9 pp83-91 神戸学院大学グローバル・コミュニケーション学会
 - ・ 結城恵 (2025) 「外国人留学生の日本就職・定着をめぐる政策に関する一考察」LOGOS (ロゴス)：群馬大学大学教育・学生支援機構論集第2号 pp1-14 群馬大学大学教育・学生支援機構
 - ・ 文部科学省 (2017) 「大学等におけるインターンシップの推進に係る専門人材の育成・配置について」https://www.mext.go.jp/b_menu/internship/1405622.htm
(2025年9月20日閲覧)

Challenges and Prospects of Internships for International Students as Highly Skilled Foreign professionals

Yuko WATABE

(Center for Global Education and Exchange, Toyo University)

Abstract

International students are expected to find employment and settle in Japan as highly skilled foreign professionals in the future. To promote this, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) has implemented the "Certification System for the International Student Employment Promotion Education Program." However, there are various challenges in conducting internships for international students, which is one of the main pillars of the program, and only 3% of universities had implemented the program by the fiscal year 2024.

This paper focuses on internships for international students and aims to contribute to discussions on effective internship practices by reviewing previous studies and identifying key challenges associated with such internships. In addition, the study identifies specific learning outcomes and implementation processes required of international students and related organizations, examining three stages: the preparation stage, the implementation stage, and the reflection stage. Compared with domestic students, international students face additional factors in their career exploration process, such as intercultural understanding and Japanese language proficiency. Through this analysis, the study organizes the key issues that universities and companies involved in program development and management should consider in the future.

Key words

Highly skilled foreign professionals, international students,
internships, employment promotion education programs for international students

編集後記

東洋大学国際教育センター紀要第4号が刊行の運びとなりました。多くの方々にご尽力を賜り、感謝申し上げます。

本紀要は査読制をとっております。国際教育センターの教育研究に関連する国際教育、異文化理解、英語教育、日本語教育の研究領域において研究論文等を募集し、これらの研究領域の専門家によって投稿論文の審査を行っています。

第4号には研究論文5本と教育実践報告1本の投稿がありました。研究論文5本について、10名の査読者による審査の結果、研究論文1本と研究ノート1本が採択され、2本が研究報告として掲載されることとなりました。査読者よりいただいた多くのコメントは論文の改善に大きく役に立ったと確信しています。査読にご協力をいただいた先生方、誠にありがとうございました。

編集委員会の先生方には、ご多忙の中、査読候補者の選定、依頼にご尽力をいただき、また、国際部の担当職員の方々には事務局として紀要に係る様々なことで多くのサポートをいただきました。感謝申し上げます。

2026年3月

国際教育センター紀要編集委員会

委員長 董晶輝

Editor's Postscript:

The No. 4 of *the Bulletin of the Toyo University Center for Global Education and Exchange* has been ready for publishing. We are deeply grateful for the all who have given their efforts for this issue.

The manuscripts published in the bulletin are peer reviewed. We called for papers in the research fields of global education, cross-cultural understanding, English teaching, Japanese teaching which related to the education and research aims of the Center for Global Education and Exchange. The submitted papers are reviewed by the professionals of the above research fields.

There were five academic papers and one report of educational practice submitted to the 4th issue. The five academic papers were reviewed by the ten anonymous referees. As a result, one manuscript is accepted and the other one is accepted as a research note, the remaining two were accepted as a report of research. We believe that the valuable comments from the referees largely improved these manuscripts. We thank the referees for their cooperation.

Finally, I would like to thank the editorial committee members for their hard work to find the proper referees, and also the editorial staff members for their various supports.

March 2026

The editorial committee of *the Bulletin of the Center for Global Education and Exchange*

The editor in chief, Jinghui Dong

国際教育センター紀要編集委員

荒巻 俊也（センター長 国際学部教授）
董 晶輝（副センター長 経営学部教授）
スティーブン・グリーン（副センター長 法学部准教授）
花田 真吾（国際学部准教授）
内田 千春（ライフデザイン学部教授）
岩本 典子（理工学部教授）
平畑 奈美（文学部教授）

Editorial Committee Members:

Toshiya Aramaki (Director, Center for Global Education and Exchange;
Professor, Faculty of Global and Regional Studies)
Jinghui Dong (Deputy Director, Center for Global Education and Exchange;
Professor, Faculty of Business Administration)
H. Steven Green (Deputy Director, Center for Global Education and Exchange;
Associate Professor, Faculty of Law)
Shingo Hanada (Associate Professor, Faculty of Global and Regional Studies)
Chiharu Uchida (Professor, Faculty of Human Life Design)
Noriko Iwamoto (Professor, Faculty of Science and Engineering)
Nami Hirahata (Professor, Faculty of Letters)

国際教育センター紀要第4号

2026年3月19日発行

*Bulletin of the Center for Global Education and
Exchange, No.4*

March 19, 2026

編集・発行 東洋大学国際教育センター紀要編集委員会

Email: ml-iao-bulletin@toyo.jp

PDF制作 株式会社 洋文社
