ISSN 2436-0325

外国語教育研究ジャーナル

JOURNAL OF FOREIGN LANGUAGE EDUCATION AND RESEARCH

第3巻
2022年12月

Vol. 3
December 2022

FLER
立教大学 外国語教育研究センター
RIKKYO UNIVERSITY
Center for Foreign Language Education and Research

目次 / Table of Contents

【Research Article】

A Boundary of National Identity Discourse: Karl Florenz's Strategy for the Historical Writing of Japanese Literature

Daisuke Baba

Abstract

This study represents an attempt to illuminate the intellectual influence of Japanese national literary studies (*koku-gaku*) on the German Japanologist Karl Florenz (1865–1939) and his representative work, *A History of Japanese Literature* (*Geschichte der japanischen Litteratur*, 1906).[1] From 1889 to 1914, this lecturer of German literature at the Tokyo Imperial University established close relationships with his Japanese students and colleagues, some of whom cooperated on his studies of Japanese literature. Their knowledge about Japanese literature was based on the *koku-gaku* tradition that manifested its policy as a modern discipline. After describing the development of the *koku-gaku* at the Tokyo Imperial University during the modernizing Meiji era (1868–1912), the present study assesses both the manifestation and the function of national literature studies. In these historical contexts, this brief study points out Karl Florenz's activity and limits of his study of Japanese literature in order to clarify his strategy of the historical writing of national literature as a German oriental philologist, especially in contrast to William G. Aston's (1841–1911) *A History of Japanese Literature* (1899). The following textual analysis of *Geschichte der japanischen Literatur* demonstrates how Florenz transformed his knowledge of Japanese literature from the *koku-gaku* into the German historical writing of national literature. The conclusion of this study discusses Florenz's philosophy of the national frames of German and Japanese literature and its boundary function in the context of national identity.

Keywords: *German, Japanese, nation, civilization, culture*

1. Introduction

Japan in the late 19th century marks its modernization as a nation state while rapidly adopting civilization and cultures of European and American powers. In this process, one of the most urgent problems that the Meiji-government had to deal with was the general education of the Japanese people. Japanese intellectuals were convinced of the need for establishing a national institution for education. During this period, the first university in Japan, the University of Tokyo (*Tokyo Daigaku*), was founded as the highest graded educational organization (Yamamoto 2014: 63-86).[2] This university represented a modern example of higher education based on European and American university models. Political leaders primarily required academic education for Japan's material enrichment. In slight contrast to this movement, which was incredibly single-minded in its focus upon material

1 I appreciate the assistance of Professor Dr. Michael Heitkemper-Yates (Rikkyo University) for kindly proofreading this paper. In this paper, the Japanese surname will be written first in all instances. The contents of sections 2 and 3 are partly included in a previous publication: Daisuke Baba. (2020). *A Hybrid Origin of Modern Historical Writing of Japanese Literature. Karl Florenz's "A History of Japanese Literature" and the German–Japanese Contact of Academic Cultures* (*Kindai Nihon Bungaku-shi Kijutsu no Haiburiddo na ichi Kigen. Kāru Flōrentsu "Nihon Bungaku-shi" ni okeru Nichi-doku no Gakujutsu Bunka Sesshoku*). Tokyo: Sangensha.

2 *Tokyo Daigaku* was renamed "Imperial University" (*Teikoku Daigaku*) after the Decree for the Imperial University (*Teikoku Daigaku-rei*) by the First Cabinet of Itō Hirofumi (1841-1909) in 1886. When, in 1897, another Imperial university was also founded in Kyōto, the name of "Imperial University" in Tokyo was changed to "Tokyo Imperial University" (*Tokyo Teikoku Daigaku*). This study will refer to *Tokyo Daigaku* in this period only as "Tokyo Imperial University."

advancement, Fukuzawa Yukichi (1835–1901) published *An Outline of Civilization Theory* (*Bun'mei-ron no Gairyaku*, 1875). Fukuzawa's work discusses how Japan should go with European and American civilization in order to obtain its own international independence, thereby explaining civilization from the perspective of not only material aspects but also with a recognition of the moral and intellectual aspects of the nation (Fukuzawa 1959: 19, 85). In the early phase of Japanese modernization, the key word "civilization" (*bun'mei*) already played an essential role in setting up the foundation of Japan's educational system and national identity.

At the Tokyo Imperial University, some intellectuals of Japanese language and literature studies were employed as officials and teachers. In the early Meiji era, from 1870s to 1880s, they had already developed the clear awareness that their *koku-gaku* tradition from the 17th and 18th centuries could not be maintained during the modernizing process without the financial and institutional support of the new government. One of these old-guard intellectuals, Konakamura Kiyonori (1822–1895), declared on the basis of Fukuzawa's theoretical treatise that the main purpose of Japanese studies as a modern discipline should be to support the moral and intellectual development of the nation state (Fujita 2007: 210). The official intellectuals regarded these aspects as the "national body" (*koku-tai*)—the unique national harmony between the Japanese political state under the emperor and the spiritual attitude of the Japanese people (Wachutka 2016: 65-82). In order to clarify these spiritual aspects of nationality, early modern researchers of Japanese language and literature had attempted philological approaches to ancient (and partly medieval) poetry and prose (Hisamatsu 1969: 139, 145). The accumulation of those findings up to the Meiji era allowed the next generation to describe a history of Japanese literature from the ancient to the early modern period. Indeed, Konakamura suggested during a meeting in 1891 that the writing of a well-summarized history of Japanese literature based on these philological findings aimed at presenting a panorama of the subjects of the modern *koku-gaku* (1898: 22).

In 1889, Karl Florenz started his career as a lecturer of German language and literature at the Tokyo Imperial University, while privately studying Japanese literature. From 1883 to 1886, he had specialized in Sanskrit philology at the University of Leipzig to become better acquainted with Japanese students. Inoue Tetsujirō (1856–1944), a member of the scholarly elite who spoke German well, not only regularly taught him Japanese language, literature, and history, before Florenz arrived in Japan in 1888 (Inoue 1943: 218-20) but also recommended him to the government as a German teacher for the new university (Kamimura 2001: 423). Florenz's study of Japanese literature from 1889 to 1914 was supported by his Japanese colleagues and students. Another colleague, Haga Yaichi (1867–1927), who specialized in Japanese literature from 1892 to 1895 under Konakamura's instruction, was interested in German theories and methods of language and literature study in order to justify the *koku-gaku* tradition as a modern discipline in the abruptly westernizing Japanese university system. In conceptualizing his history of Japanese literature, Haga made a close friendship with Florenz and privately lectured him on Japanese literature. Meanwhile, Florenz privately requested Fujishiro Teisuke (1868–1927), a student in the newly formed department of German literature, to assist in his own study of Japanese literature. After the graduation, Fujishiro attended lectures of a *koku-gaku* scholar on *Manyōshū*, the oldest collection of Japanese poetry from the 8th century, in order to regularly explain to Florenz the detailed contents of the lectures in German (Fujishiro 1927: 138-46). Furthermore, Fujishiro translated a large number of Japanese literary works into German for Florenz (Tobari 1970: 299-300). In his seminar for the German department, moreover, Florenz gave all his students a task to write a thesis about Japanese literature in German (Tokyo Daigaku 1986: 768) in order to gain more precise knowledge of Japanese literature. The

knowledge of the students about Japanese literature undoubtedly consisted of lectures of *koku-gaku* scholars at the university. From this situation, it is highly evident that Florenz's study of Japanese literature and one of his major works, *Geschichte der japanischen Litteratur*, were strongly influenced by the *koku-gaku* tradition through his collogues and students.

Previous studies of Karl Florenz have rarely focused on the process of academic exchange between Germany and Japan itself, but have typically revolved around the assessment of Florenz and his works in the present discourse concerning modern academic exchange.[3] In contrast to the previous studies, this study proposes that a hybrid process of German–Japanese academic contact occurred during this period. The case for this proposal shall proceed by means of the following questions: 1) what is the background of Florenz's description of his history of Japanese literature; 2) for what reason did he transform the *koku-gaku* tradition through German methods of historical writing; and as a result, 3) what kind of character did his history of Japanese literature obtain. These historical perspectives on the academic modernization of language, literature, and education enable this present study to contribute to a better understanding of contemporary Japanese language and education.

2. Historical writings of Japanese literature: Haga, Florenz, and Aston

Toward the end of the 19th century, Florenz began to draft his history of Japanese literature. The historical writing of Japanese literature itself had started with Mikami Sanji's (1865–1939) and Takatsu Kuwasaburō's (1864–1921) *A History of Japanese Literature* (*Nihon Bungaku-shi*, 1890) and *A Reading Book for Our National Literature* (*Koku-bungaku Toku-hon*, 1890), edited by Tachibana Senzaburō (1867–1901) and Haga Yaichi (Takatsu 1969: 10). Mikami's and Takatsu's representation had already referred to English writing methods on the basis of civilization theory (Mikami and Takatsu 1890). The civilization theory related to the historical writing of national literature represented the concept that the national mind develops its literature and culture nearly parallel to the material advancement. On this theoretical basis, a number of histories of Japanese literature were published during the 1890s. In this context, it should be questioned how Florenz, who was certainly not able to describe the history of Japanese literature by himself, adopted former representations. In the foreword of his literary history, Florenz remarks that among the literary approaches, only a few works were worthy of his reference:

> As priceworthy exceptions I would like to firstly give Professor Yaichi Haga's imaginative *Ten Lectures on a History of Our National Literature* (*Koku-bungaku-shi Jikkō*), then Dr. S. Fujioka's *A Textbook of a History of Japanese Literature* (*Nihon Bungaku-shi Kyōkasho*) along with supplement and Wada's and Nagai's *A Short History of Our National Literature* (*Koku-bungaku Shōshi*). These books are also a moderate size and give more hinting outlines than detailed descriptions rounded off in itself. But these books are often valuable tools and guides for my orientation in the chaotic amount of the stuff to be considered. Of course, I have also drawn

3 Representative works include the following: Satō Masako's first extent biographical study (1995) argues that this German philologist played an intermediate role in the modern development of Japanese literature studies as a result of his importation of German theories and methods. More recently, Kamimura Naoki (2001) published a research study concentrating on Florenz as the founder of German studies in Japan. Furthermore, Michael Wachutka's first monograph, published in the same year, investigates Florenz's translation of *The Chronicles of Japan* (*Nihon Shoki*, approx. 720) and his religious interpretation of Shintō. Additionally, Tsuji Tomoki's doctoral thesis (2010) pays critical attention to Florenz's "euro-centric" attitude as an orientalist who depicts Japan as a subject of the Western scientific system.

> instruction about quite many subjects thankfully from W. G. Aston's 1899 published *A History of Japanese Literature*, the first history of Japanese literature, which was described by a European and is worthy of its high commendation (Florenz 1906: IV-V).

In this way, Haga's and Aston's histories of Japanese literature belong to the main reference books for Florenz's historical description.[4] This section attempts to answer the question of how and for what reasons the German philologist included Haga's knowledge and Aston's English translation of literary works into *Geschichte der japanischen Litteratur*.

Undoubtedly, Florenz and Haga were constantly cooperating with each other on their study of Japanese literature. Florenz, who lectured at the department of German studies about the history of German literature (Tobari 1965: 35), had surely gained exact knowledge about historical writing methods widespread in the German speaking areas up to the end of the 19th century. Haga, to whom Florenz not only taught German but also allowed to use the German books of his collection, provided in his 1895 published essay a concept of a historical writing of Japanese literature, thereby summarizing the names of prominent historians of German literature such as Georg Gottfried Gervinus (1805–1871) (Kaikawa [Haga] 1895: 188-98). Meanwhile, in 1891, Haga composed a review of Florenz's translation and comments of *The Chronicles of Japan* that consisted of Fujishiro's translated original manuscripts (Haga 1992: 6). In his first lecture on national literature in the summer of 1898, Haga made a short comment that Florenz had already started writing down his history of Japanese literature (Haga 1983: 187), certainly with Haga's aid. The relationship between these relatively young intellectuals led to a content framework for Florenz's historical description of Japanese literature based on Haga's knowledge.

During the summer vacation time of 1898, Haga Yaichi delivered at the Education Association of Tokyo Imperial University (*Teikoku Daigaku Kyōiku-kai*) *Ten Lectures on a History of our National Literature* to students of Japanese studies and teachers of Japanese language (*koku-go*). In the following years, a supplemental book version of these lectures was put together with a short foreword and contained these 10 lectures in 267 pages as well as an index for the names of authors and works and terminologies (Haga 1903). It is Haga's detailed knowledge in regard to periodization, content constitution, and objects of description in these lectures that Florenz seems to have articulated in his historical writing, certainly not by reading this work directly, but by getting Haga's explanations.

As for periodization, Haga (1983: 194-8) categorizes his history of national literature into five blocks of time according to the changes of the political system. The oldest period (approx. the 6th century to 794), which Haga called "above-old" (*jō-ko*), extends from the ancient time respectively before the reception of Chinese culture and Buddhism up to the movement of the capital from Nara to Heian-kyō (contemporary Kyōto). The second "middle-old" (*chū-ko*) period (794–1192) marks the aristocratic institution and culture in Heian-kyō until the establishment of the feudal military government in Kamakura. Afterward, during the third "near-old" (*kin-ko*) period (1192–1601), the military regime moved its capital from Kamakura to Kyōto by the Muromachi (or Ashikaga) shogunate. In the next "near-world" (*kin-sei*) period (1601–1868), or the early modern Edo period of Japanese history (Shirane, Suzuki, and Lurie 2016: viii), the country was under the rule of the Tokugawa shogunate in Edo (contemporary Tokyo). The Meiji Restoration in 1868 marks the end of the feudal society, when Japan was isolated from other countries at the beginning of the 5th period.

4 Florenz seems to have referred to Fujioka's *A Textbook for a History of Our National Literature* (1901) and Wada's and Nagai's *A Short History of Our National Literature* (1899) only for the extract of a poetry in *Kojiki* or *Record of Ancient Matters* (Baba 2020: 67-8).

Haga designates this as "present time" (*gen-dai*) in which the political system under the emperor ruled the Japanese people with a distinctly Western influence. These periodizations of political history, as we will see below, constitute the contents of Florenz's historical writing. Aston's history of Japanese literature, meanwhile, is divided into seven periods: 1) Archaic period (before A.D. 700); 2) Nara period (8th century); 3) Heian or classical period (800-1186); 4) Kamakura period (1186–1332); 5) Nanboku-chō and Muromachi periods (1332–1603); 6) Yedo period (1603–1867); 7) Tokio period (1868–1900).

Haga's lectures also show parallels with Florenz's literary history in descriptive constitution. In the introduction to the first lecture, Haga restricts his use of the term "literature" (*bungaku*) on the basis of the European general definition of all pieces of writing determined to be "artworks such as poetry and prose" (Haga 1983: 188–9). In the beginning of the explanations about each period, he summarizes its major historical process including political events and its influences on Japanese culture. Furthermore, he sketches out the main development of literature as a characteristic product of Japanese collective mind in each period. After the general description of literary characters and genres, Haga explains representative authors and works along with the secondary literature of Japanese literary studies for self-study. In the same way, Florenz and Aston introduce—instead of comments on the secondary materials—their translation of passages from various literary works, adding to their summaries of the context. Meanwhile, the notion of Japanese people (*koku-min*) as the subjects of the emperors from the ancient to the current period, especially within the Imperial state during the Meiji era, obviously plays a central role in Haga's descriptive constitution. According to his understanding, Japanese literature is a mirror of the collective mind of the Japanese people whose literary development matches their process of civilization (*bun'mei*), especially in comparison with the ancient influence of China and during the late 19th century influence of Western powers (Haga 1983: 188-9). This triad of the Japanese people, literature, and civilization functions characteristically as a component of Haga's lectures on national literature.

Haga's and Florenz's writings of Japanese literature seem to be also nearly matching each other in literary genres, authors, works, and the main objects of historical description. Haga features literature in the ancient Nara period as writings of the oldest thoughts through purely Japanese words (Haga 1983: 195) including *Records of Ancient Matters* (*Kojiki*), *The Chronicle of Japan* (*Nihon Shoki*), and *Collection of Ten Thousand Leaves* (*Manyōshū*). As the foundation of Heian literature in the next period, Haga depicts the reception of Chinese literature and Buddhism and the invention of the Japanese *kana* syllabary. He calls this writing system "national letters" (*koku-ji*) and sentences written using the *kana* syllabary "national prose" (*koku-bun*) (ibid.: 196). In addition, he characterizes the Heian period as "an effeminate period" (ibid.: 222) and its literature as "extremely elegant in perspective of both form and content" (ibid.: 252). Haga situates Murasaki Shikibu's *The Tale of Genji* (*Genji Monogatari*) and Sei Shōnagon's *The Pillow Book* (*Makura no Sōshi*) at the top of Japanese national literature (ibid.: 246). He also evaluates the prose works of both female authors as "the origins of our national literature" or "classical literature" (*mohan bungaku*), to which Japanese authors in the following periods referred for their own prose writing (ibid.: 196). In the next Kamakura and Muromachi periods, studying and writing were no longer active through repeated wars (ibid.: 267). In contrast to court women, who were representative of Heian literature, Haga highlights male warriors and Buddhist monks, the major authors during the Kamakura period, whose works reflect the depressive feelings of Buddhist pessimism ruling at the time (ibid.: 252). In addition, he points out the development of the new Sino–Japanese writing style mixed with the Japanese *kana* syllabary (*wakan-konkō-bun*) and the emergence of the following genres: the military

chronicle (*gunki*) in the Kamakura period, the textual part of the traditional *Nō* theater (*yōkyoku*), comic theater (*kyōgen*) and the Muromachi tale (*otogi zōshi*), and linked verse (*renga*) in the next Muromachi period. In the Tokugawa period, and afterward, Japanese literature was no longer dependent on the authority of the Imperial court, Buddhist temples, or military lords. Because of the government promotion of education and the spread of print capitalism, especially urban townspeople accepted and created literary works. As most representative poets and authors of the "folk literature" (*heimin bungaku*) (ibid.: 197), Haga recognizes Matsuo Bashō (1644–1694) in popular linked verse (*haikai*), Chikamatsu Monzaemon (1653–1725) in the writing of puppet-theater plays (*jōruri*), and Ihara Saikaku (1642–1693) and Kyokutei Bakin (Takizawa, 1767–1848) in novels and book-length prose (*yomi-hon*) (ibid.: 287–310). According to Haga's last lecture, Japanese people who have created their own literature since the reception of Chinese culture and civilization in the ancient period will finally "harmonize the civilizations of East and West" in the Meiji period to give birth to new forms of literature (ibid.: 317).

While Haga sees the development of Japanese literature in the civilizing process of the Japanese people and their culture, Florenz describes the literary development from the perspective much less of civilization than of culture. Florenz's literary history is divided into five periods: 1) the oldest time (*älteste Zeit*, until 794); 2) Heian period (*Heian-Periode*, 794–1186); 3) Kamakura and Muromachi period (*Kamakura und Muromachi-Periode*, 1186–1601); 4) Tokugawa period (*Tokugawa-Periode*, 1602–1868); and 5) Meiji era (*Meiji-Ära*, since 1868) (Florenz 1906: VII-X). In this history, "the mind of the Japanese people" (*der japanische Volksgeist*, ibid.: III) fundamentally functions as a subject developing itself repeatedly with prosperity or decay. According to this theoretical framework, attitudes of the collective mind typical of the period are reflected in representative works of Japanese literature. The expression of "the people" (*Volk*) in the German historical context of the late 19th century meant not a nation belonging to a state, rather a large group of people with the same language and culture (Grimm and Grimm 1984a: 425). Furthermore, Florenz depicts the appropriate way the Japanese people have repeatedly attempted to adapt forms of Chinese literature since their reception of Buddhism and Chinese culture in the 7th century in order to refine their own literature. According to Florenz, the development of Japanese literature reached its highest point in the women's court literature during the 11th century—what Florenz commends as a literary harmony of "form" adapted from Chinese literature and "content" reflected in the minds of the Japanese people (Florenz 1906: 208). Florenz argues that in the 12th century, the collective mind experienced a declining process as a result of repeated wars, when male warriors and monks became the representative authors during the Kamakura period (ibid.: 255). In the 15th century, the reception of forms of the Chinese drama led to the development of the *Nō* theater as a type of Japanese drama (ibid.: 387). Up to the late 19th century, Japan remained peaceful under the rule of the Tokugawa centralist government. At the time, intellectuals were engaged not only in the research for ancient Chinese literature, but some of them also labored nationalistically for a new conception of Japanese literature in order to counter the authority of Chinese studies (ibid.: 416). Florenz indicates that in the "renaissance" of the Tokugawa period, a literary "reform" occurred as a result of Matsuo Bashō's Haikai-poetry and caused a small and short flourishing time of literature (ibid.: 446). Nevertheless, in his short report on Meiji literature, Florenz comes to the conclusion that the Japanese mind is confronted with the difficult situation of the decline because of its one-sided reception of European literature and culture, despite its tradition of Japanese literature (ibid.: 612). For that reason, Florenz constitutes his Japanese literary history based on the German scheme of the cultural development of the collective mind, which barely appears in Haga's lectures of national literature.

Considering Florenz's knowledge of German historical writing and the instruction in Japanese literature provided by his colleagues and students, it is assumed that he did not necessarily refer to William G. Aston's history of Japanese literature in order to better write his own literary history. In December 1900, however, Haga sent his comments on Aston's literary history to Florenz (Haga 1992: 173). The key point to understand in Florenz's reference to Aston's book seems to be the self-obvious fact that Florenz's book was written for German intellectual readers such as his colleagues of oriental studies in Germany. Most of those German readers who understood English well were likely to peruse the first history of Japanese literature by a European. As Florenz mentions in his text, Aston's English work had earned a high reputation until the German literary work was published. Therefore, it is clearly suspected that Florenz was urged to produce a better history of Japanese literature than that of his English colleague.

To compare Florenz's with Aston's work from this point of view, it is evident that Florenz intentionally not only translated almost the same passages from the Japanese original texts as Aston had done but also added context to his descriptions of these passages.[5] As we have seen above, Florenz's study of Japanese literature was dependent on his student's translations from the original texts into German. Taking these situations into account, it seems plausible that Florenz requested that Fujishiro translate the passages of representative works, which had been introduced in Aston's literary history in order to demonstrate to German readers that *Geschichte der japanischen Litteratur* details more translations and explanations of representative works than Aston's *A History of Japanese Literature*.

Almost all the historical writings of Japanese literature in the Meiji era took the framework of nation, state, and civilization for granted. From this perspective, literature was considered a product of the national mind in both the advancement of its civilization and the development of its culture. In fact, the historical writings based on the English understanding of literature as all written artwork of poetry and prose, as an excellent result of civilization and culture of the nation, describe a parallel development of civilization and culture (Shinada 2012: 11-2). Fukuzawa Yukichi's work had introduced this civilization theory for the first time in regard to François Guizot's (1787–1874) *Histoire de la Civilization en Europe* (1828) and Henry Thomas Buckle's (1821–1862) *History of Civilization in England* (1857) (Matsuzawa 1995: 368-9). In contrast to Aston, who seems to have been well aware of the English writing method for European readers, Haga, the heir of Konakamura's concept of *kokugaku*, certainly made use of the civilization theory for his lectures on national literature in order to invent the consciousness where all of the people in Japan might be seen as belonging to the Imperial state. According to Haga's historical writing, Japanese literature represents a cultural heritage of the nation state, which seems to have been constant from the ancient times to the Meiji era. This attitude also claims that the tradition of Japanese literary studies is worthy of Meiji Japan and a necessary modern discipline for the Tokyo Imperial University.

---

5 This includes the following: *Manyōshū* (Aston 1899: 36-48; Florenz 1906: 94-126), *Collection of Ancient and Modern Poems* (*Kokinshū*) (Aston: 60-2; Florenz: 138-53), *Tosa Diary* (*Tosa Nikki*) (Aston: 68-76; Flroenz: 192-7), *The Pillow Book* (*Makura no Sōshi*) (Aston: 106-16; Florenz: 224-8), *The Record of the Rise and Fall of the Genji and Heike* (*Genpei Jōsui-ki*) (Aston: 137-9; Florenz: 304-7), *The Tales of the Heike* (*Heike Monogatari*) (Aston: 141-2; Florenz: 304-7), *The Ten-Foot Square Hut* (*Hōjō-ki*) (Aston: 146-56; Florenz: 325-9), *Chronicle of Great Peace* (*Taihei-ki*) (Aston: 175-8; Florenz: 311-5), *Essays in Idleness* (*Tsurezure-gusa*) (Aston: 189-96; Florenz: 331-8), and *Battle of Coxinga* (*Kokusen'ya Kassen*) (Aston: 280-7; Florenz: 596-8). Especially, it is a striking feature of Florenz's strategy against Aston's literary history that Florenz translates just the same passages of the 2nd chapter "The Broom-Tree" (*Hahaki-gi*) and the 5th "Young Murasaki" (*Waka Murasaki*) of *The Tale of Genji*, one of the longest novels of Japanese literature (Aston: 98-103; Florenz: 215-8).

3. From the advancement of civilization into the development of culture

In the context of former literary histories, Karl Florenz reformulated Haga's knowledge of Japanese literature through German methods of literary history. In other words, it is necessary to discuss how Florenz wrote the civilization theory into the scheme of cultural development that had been fundamental to the historical thinking of national literature in the German speaking areas of the 19th century. One point that must be addressed is his foreword in which Florenz explains the aim of his literary history and distinguishes "civilization" (*Zivilisation*) from "culture" (*Kultur*). Another point worthy of attention is his evaluation of *The Tale of Genji* and *The Pillow Book* by female authors in Heian period, both of which can be, according to the civilization theory, assessed as an excellent result of civilization and culture of the Japanese people or their most sophisticated prose through the Japanese original *kana* syllabary.

The foreword of *Geschichte der japanischen Litteratur* begins with the claim that since the Meiji Restoration in 1868, Japan has increasingly accepted European and American civilization and cultures.[6] In this passage, Florenz emphasizes that Japan is the only advanced country equal to European powers among the non-European countries. He indeed contrasts the words "civilization" and "culture" without definition, but from the passage, it is clearly understood that the "advancement" (*Fortschritte*) of Japan means less of its cultural development of the collective mind than its progress toward material development. In addition, after remarking that more and more Europeans and Americans have become interested in Japanese culture, Florenz insists that even before the European influence, Japan already possessed a "highly developed" (*hochentwickelt*) culture (Florenz 1906: III). This passage, therefore, represents Florenz's clear distinction between the advancement of civilization and the development of culture.

In this contrast, Florenz, referring to the recent influence from Japanese on European culture—hinting at the Japonism in the late 19th century—addresses the notion of literature as follows:

> Very much less than the visual arts, we have been knowing of another expression of the collective mind of the Japanese people, literature, although in this literature, the fineness of Japanese taste expresses itself not less sharply than in the visual arts. (Florenz 1906: III)

Florenz sets up the collective "mind of the people" (*Volksgeist*) that functions as a subject developing itself in his historical writing of Japanese literature from the ancient to the latest period. In this framework, the historical subject has created a variety of literary works reflecting circumstances in each period. It is remarkable that Florenz distinguishes between people of European and American powers as a "nation" (*Nation*) and the Japanese people as a "people" (*Volk*), although he does not define these terms specifically. Both *Nation* and *Volk* were synonymous as "language communities" in the 19th century. *Nation* in German impressed an image of the entire people in the German speaking areas after the emergence of civil society as a single community (Grimm and Grimm 1984a: 425). The term *Volk*, meanwhile, also meant a group or cultural community in ancient times (Grimm

6 "It has not passed half a century yet since the Japanese people (*Volk*) have inclined to come out of their own closure from the other nations (*Nationen*) in the world and appropriate the material advantages of our modern civilization (*Zivilisation*) first of all with interest in their self-preservation. Then, the not less blessed result that was not intended at first is that Japan increasingly transformed itself in harmony with the mental and moral ideal of Western culture (*Kultur*). With lively participation, we have observed from all the points of view the surprising advancement (*Fortschritte*) of the East Asian people (*Volk*). They are the only people until now who have succeeded in owning an equal position of the European and American nations (*Nationen*)" (Florenz 1906: III).

and Grimm 1984b: 454). Considering the general meanings and Florenz's use of these terms, it is reasonable to think that he calls the group of European and American people (who would be more closely related to his German readers) as *Nation* and the group of non-European people as *Volk*.

It is through these synonyms in German, nation and people, that Florenz reconstituted Haga's civilization framework of national literature as an understandable type of a literary history for his German readers, focusing on the cultural development of the collective mind of the Japanese people. German literary historians in the 19th century, on the one hand, did not clearly distinguished between *Volk* and *Nation*. These terms were used mostly as synonyms in the meaning of one language community of the German speaking people. On the other hand, Haga Yaichi, as we have seen in section 2, regarded the Japanese "nation" (*koku-min*) as the one language unit of the subject belonging to the Imperial state. Taking these differences between the German and the Japanese understanding of language community into account, Florenz probably starts the foreword of his literary history with the perspective of the civilization theory and sets up the Japanese people only in the German meaning of one language community in contrast to the European and American nations related to their modern state. Florenz seems to restrict the aspect of his historical writing to the development of Japanese literature regardless of civilization. The contrasting use of nation and people, indeed, is limited to the brief passage in the foreword. In the main text, Florenz describes the historical process mostly from a cultural point of view.

As for Florenz's view of the literary history, it is remarkable that he once uses the term "cultural advance" (*kulturelle Fortschritte*) in order to explain Japan's reception of Chinese culture and its result. He emphasizes Buddhism and education coming from China in the 7th century as a main factor of cultural advance in Japan (Florenz 1906: 48). This expression includes the assumption of the "self-forming of the mind" (*geistige Bildung*), which argues that the Japanese people still in uncultivated conditions of life devoted themselves to learning the "highly developed" Chinese culture painstakingly in order to create their own literature (ibid.: 48-9). In other words, this perspective reveals that Florenz does not recognize Japan before the first reception of Chinese culture as a cultural at all. Without reasoning his viewpoint, Florenz additionally paraphrases in comparison with a foreign factor of Chinese culture an original factor of Japanese people as things "indigenously Japanese" (*einheimsch-japanisch*) (ibid.: 51). From this context, Florenz apparently assumes that the mixture of indigenously Japanese and Chinese factors through Japan's adoption of Chinese culture enabled the Japanese people to make dramatic progress and develop their own national literature. Over his literary history, furthermore, Florenz draws the reader's attention to Japanese people repeatedly receiving Chinese culture and improving their own culture such as the *Nō* drama in the Muromachi period and the style of novel in the Tokugawa period (ibid.: 287, 375, 517-8). For these reasons, Florenz's use of the expression "cultural advance" seems to be caused by his thinking of the first reception of Chinese culture as a crucial opportunity for Japanese people in uncultivated conditions to rapidly develop their own culture.

The description of culture in both aspects of advancement and development was unusual in the context of German historical writing in the 19th century. In German speaking areas, the historical process was described almost only from the perspective of the development of the national mind. As for the expressions of "advance" (*Fortschritt*) and "development" (*Entwicklung*), Leopold von Ranke (1795-1886) determined a comparative use in the context of Historicism. In the introduction of his *About the Epochs of Recent History* (*Über die Epochen der neueren Geschichte*, 1854), Ranke declares the historical idea that all of humankind has advanced constantly to be unproven. Instead of this idea (typical of the Enlightenment period), he insists on historical research to illuminate the force of the

national mind in a process of development, whereby his image of development represents a model of prosperity and decay (Ranke 1954: 5-7). This German standard historical view in the 19th century barely corresponds to Florenz's expression of cultural advance.

To sum up, Florenz seems to have used the notions of nation and people on the one side, and of advancement and self-development on the other side, in order to transform Haga's knowledge of Japanese literary history on the basis of civilization theory into a model of cultural development of the national mind through German historical writing methods. In the foreword, Florenz regards the European and American nations as communities of a modern confederated state or the resulting societies of high civilization. In contrast to this concept of national status, the Japanese people, according to the German general understanding of the literary history, means only a community of language and culture—regardless of modern national status. The notion of people matches better with the German theory of the development of the national mind than the notion of nation (as it was self-applied at the time). In the main text concerning the cultural development of the Japanese people, Florenz sets up the phases of both the advancement through the adoption of Chinese culture and the following self-formation. This notion of advancement, nevertheless, seems to be less familiar with the German historical model of prosperity and decay than with the civilization theory. The manipulation of cultural advancement and self-development is probably caused by the situation related to the historical writing of Japanese literature discussed in section 2. In this situation, Florenz had to describe Japan's relation to China to a certain extent from the perspective of civilization. As a result, he introduced his original use of the expressions of cultural advance and self-development or a hybrid point of view influenced by both Japanese and German discourses on national representation.

The theory of self-formation in literary development, especially in relation to a highly developed foreign culture, played a major role in the German speaking areas to define their cultural identification. Georg Gottfried Gervinus claims that within the European countries, only German people have appropriated classical literature and wisdom to such a high level as the ancient Greeks and Romans (Gervinus 1840: 11-2). The German philologist Wilhelm Scherer (1841–1886) insists in his assessment of the development of German literature that German people in the primitive conditions of their literature, repeatedly learned the literature of foreign countries to refine their own literature (Scherer 1883: 19, 21). In an earlier article about his concept of historical writing published in 1879, Scherer describes the cultural character of the German people as "the greatest possible increase in classical education" (*möglichste Steigerung der classischen Bildung*) (Scherer 1975: 397). Gervinus's and Scherer's works are recognized as the bestsellers of the history of German literature in the 19th century (Rosenberg 1989: 109), representing an academically approved self-image of German cultural features widespread in the civil society. Florenz, who studied in Leipzig and Berlin from 1883 to 1888, provided lectures on German literary history at the Tokyo Imperial University. Undoubtedly, he also shared the discursive type of German national self-understanding through the historical writings surrounding German literature—especially the German high receptivity of foreign cultures. This probably allowed his connection between German and Japanese cultural development to show German readers the self-forming process of the Japanese mind through Chinese literature and education.

In this view of cultural advancement and self-development, Florenz describes the Heian courtly literature of the 11th century, or the prose works by Murasaki Shikibu and Sei Shōnagon, not as a literary production of national civilization and culture, but as the most sophisticated harvest of Japan's reception of Chinese literature and the self-forming of its own literature. Florenz's following evaluation about both women's works that were, according to his comments, written nearly during

the same time round 1000 surely suggested to German readers an analogy with the development of German literature:

> To choose a round number, the year 1000 represents the highest point of literary production of the ancient Japan (Florenz 1906: 229).

Wilhelm Scherer's standard history of German literature sketches a regular development with prosperity around the years 600, 1200, and 1800, as well as with decay around the years 300, 900, and 1500 (Scherer 1883: 18). The second period of prosperity during the Medieval age is marked by the Chivalric romance or courtly poetry, particularly that of Gottfried von Strassburg and Wolfram von Eschenbach. Florenz, meanwhile, not only lists the Heian period as the "Medieval age" (*Mittelalter*) but also regards a characteristic point in the development of Japanese literature in comparison with European literatures as "early ripening" (*frühes Reifen*) (Florenz 1906: 229). His German readers who were familiar with Scherer's literary history were able to easily imagine that the courtly literature in the "Medieval age" in Japan reached the highest point of historical development 200 year earlier than in German speaking areas; not in poetry but in prose, and not written by men but by women. This type of analogous thinking functions as a discursive tie between German and Japanese literature in order for his intellectual readers to increase their interest in Japanese culture.

4. Conclusion

In this way of historical writing, Karl Florenz included Haga's knowledge of national literature from the *koku-gaku* tradition in the German context of the 19th century. Whereas Florenz, a specialist in ancient Japanese literature, had not been well aware of Heian courtly literature, his *Geschiche der japanischen Litteratur*, discussed above, evaluates both women's works, *The Tale of Genji* and *The Pillow Book*, as the classic prose of Japanese literature—differently from the previous histories of Japanese literature established by the civilization theory related to Japanese national identity discourse. In this reformulation of the literary knowledge, we can see the German Japanologist on a boundary between Japanese and German academic contexts in the late 19th century. On the Japanese side, intellectual leaders of the *koku-gaku* had set their academic policy on the basis of Fukuzawa Yukichi's civilization theory for the assimilation of the modern university system. On the German side, the growing interest in Japanese culture mainly within German people living in Japan (Schütte 2004: 62) and highly theorized historical methods for national literature enabled the German lecturer at the Tokyo Imperial University to describe his history of Japanese literature through the knowledge and assistance of his collogues and students.

The boundary on which Florenz stood while writing his literary history should be understood in the dynamic process between Japanese and German contexts. His historical writing was dependent on his situations as a German oriental philologist, an early German Japanologist, and a teacher within the department of German literary studies in a modernizing Japan. Under the influence from both academic contexts, Florenz's use of cultural advancement and self-development for historical writing led to his analogous thinking about the reception of highly developed foreign culture as a common similarity between German and Japanese national literature. With regard to this entire process, his understanding of Japanese literature distanced itself on the aspect of German historical writing to a certain extent from the Japanese discourse. At the same time, he had to follow the high evaluation of the Heian representative prose, which the previous literary historians had estimated very positively

on the basis of civilization theory and with which Florenz himself was probably not acquainted. The academic boundary about national identity included conflicts and comparability between German and Japanese literature studies in the modern period.

Bibliography

Aston, William George. (1899). *A History of Japanese Literature*. In *Collected Works of William George Aston. Vol. 5*, I-XI, 1-408. Edited by Tōru Haga. Bristol/Tokyo: Ganesha Publishing/Oxford University Press.

Baba, Daisuke. (2020). *Kindai Nihon Bungaku-shi Kijutsu no Haiburiddo na ichi Kigen. Kāru Flōrentsu "Nihon Bungaku-shi" ni okeru Nichi-doku no Gakujutsu Bunka Sesshoku* (*A Hybrid Origin of Modern Historical Writing of Japanese Literature. Karl Florenz's "A History of Japanese Literature" and the German–Japanese Contact of Academic Cultures*). Tokyo: Sangensha.

Florenz, Karl. (1906). *Geschichte der japanischen Litteratur*. Leipzig: C. F. Amelang.

Fujioka, Sakutarō. (1901). *Nihon Bungaku-shi Kyōkasho* (*A Textbook of a History of Japanese Literature*). Tokyo: Kaisei-kan.

Fujishiro, Sojin (Teisuke). (1927). F Sensei to Manyō no Soshaku (Chewing through Manyōshū with My Teacher F). In *Gahitsu Yoteki* (*Drops from My Quill*), 138-46. Tokyo: Kōbundō.

Fujita, Hiromasa. (2007). *Kindai Koku-gaku no Kenkyū* (*A Research on Japanese Literary Studies in the Modern Time*). Tokyo: Kōbundō.

Fukuzawa, Yukichi. (1959). *Bun'mei-ron no Gairyaku* (*An Outline of Civilization Theory*). In *Fukuzawa Yukichi Zenshū. Dai 4 kan* (*Complete Works of Fukuzawa Yukichi. Vol. 4*), 1-212. Tokyo: Iwanami Shoten.

Gervinus, Georg Gottfried. (1840). *Geschichte der poetischen National-Literatur der Deutschen. Bd. 1*. Leipzig: Wilhelm Engelmann.

Grimm, Jacob and Grimm, Wilhelm. (1984a). *Deutsches Wörterbuch. Bd. 13*. München: DTV.

Grimm, Jacob and Grimm, Wilhelm. (1984b). *Deutsches Wörterbuch. Bd. 26*. München: DTV.

Haga, Yaichi. (1903). *Koku-bungaku-shi Jikkō* (*Ten Lectures on a History of Our National Literature*). Tokyo: Fuzanbō.

Haga, Yaichi. (1983). *Koku-bungaku-shi Jikkō* (*Ten Lectures on a History of Our National Literature*). In *Haga Yaichi Senshū. Dai 2 kan* (*Selected Works of Haga Yaichi. Vol. 2*), 183-318. Tokyo: Kokugakuin Daigaku.

Haga, Yaichi. (1992). Shokan (Zen) (Letters (Complete)). In *Haga Yaichi Senshū. Dai 7 kan* (*Selected Works of Haga Yaichi. Vol. 7*), 1-144. Tokyo: Kokugakuin Daigaku.

Hisamatsu, Sen'ichi. (1969). *Nihon Bungaku Kenkyū-shi* (*A History of Japanese Literary Studies*). Tokyo: Shibundō.

Inoue, Tetsujirō. (1943). Furorentsu Hakase to Nihon-gaku (Dr. Florenz and Japanology). In *Kaisō-roku* (*Memoirs*), 215-31. Tokyo: Shunjūsha Kyōyō Sōsho.

Kaikawa [Haga, Yaichi]. (1895). Bungaku-shi Hensan Hōhō ni tsukite (About Compilation Methods for Historical Writing of Literature). In *Teikoku Bungaku. Dai 1 kan* (*Imperial Literature. Vol. 1*), 188-98. Unknown publisher.

Kamimura, Naoki. (2001). Doitsu Bungaku-ka Kyōshi toshite no K. Furōrentsu (K. Florenz as a Teacher of Department of German Literary Studies). In *Meij-ki Doitsu-go Gakusha no Kenkyū* (*Studies on Researchers of German Language and Literature in the Meiji Era*), 422-455. Tokyo: Taga Shuppan.

Konakamura, Kiyonori. (1898). *Yasumuro Zakkō* (*Konakamura's Various Considerations*). Tokyo: Yoshikawa Hanshichi.

Matsuzawa, Hiroaki. (1995). Kaisetsu (Commentary). In *Bun'mei-ron no Gairyaku* (*An Outline of Civilization Theory*), 363-91. Tokyo: Iwanami Bunko.

Mikami, Sanji and Takatsu, Kuwasaburō. (1890). *Nihon Bungaku-shi*. (*A History of Japanese Literature*.) Tokyo: Kinkōdō.

Nagai, Hidenori and Wada, Mankichi. (1899). *Koku-bungaku Shōshi* (*A Short History of Our National Literature*). Tokyo: Kyōiku Shobō.

Ranke, Leopord von. (1954). *Über die Epochen der neueren Geschichte*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Rosenberg, Rainer. (1989). *Literaturwissenschaftliche Germanistik. Zur Geschichte ihrer Probleme und Begriffe*. Berlin: Akademie-Verlag.

Satō, Masako. (1995). *Kāru Furōrentsu no Nihon Kenkyū* (*Karl Florenz's Japan-Study*). Tokyo: Shunjūsha.

Scherer, Wilhelm. (1883). *Geschichte der deutschen Litteratur*. Berlin: Weidmannsche Buchhandlung.

Scherer, Wilhelm. (1975). Die Epochen der deutschen Litteraturgeschichte. In *Über Literaturgeschichtsschreibung. Die historisierende Methode des 19. Jahrhunderts in Programm und Kritik*, 395-9. Edited by Edgar Marsch. Darmstadt: Wissenschaftliche Buchhandlung.

Schütte, Hans-Wilm. (2004). *Die Asienwissenschaften in Deutschland. Geschichte, Stand und Perspektiven*. Hamburg: Institut für Asienkunde.

Shinada, Yoshikazu. (2012). Haijo to Hōsetsu. Koku-gaku, Koku-bungaku, Haga Yaichi. (Exclusion and Inclusion. Koku-gaku, Koku-bungaku, and Haga Yaichi.) In *Koku-go to Koku-bungaku. 6 Gatsu gō*. (*National Language and Literature. The June Issue*.) Tokyo: Meiji Shoin.

Shirane, Haruo, Suzuki, Tomi, and Lurie, David (ed.). (2016). *The Cambridge History of Japanese Literature*. Cambridge: Cambridge University Press.

Takatsu, Yoshihiko. (1969). *Nihon Bungaku-shi no Hōhō*. (*Methods of Histories of Japanese Literature*.) Tokyo: Ōfūsha.

Tobari, Chikufū. (1965). Doitsu-go Zange (*My Repentance for German Learning*). In *Tobari Chikufū Tsuisō Ikōshū* (*Collections of Memoirs and Posthumous Manuscripts of Tobari Chikufū*), 10-200. Tokyo: Ikubundō.

Tobari, Chikufū. (1970). Fujishiro Teisuke ni Atau (For Fujishiro Teisuke). In *Meiji Bungaku Zenshū. Dai 40 kan* (*Complete Works of Meiji Literature. Vol. 40*), 358-61. Tokyo: Chikuma Shobō.

Tokyo Daigaku (ed.). (1986). *Tokyo Daigaku Hyakunen-shi. Bukyoku-shi ichi* (*A 100 Year History of University of Tokyo. A History of the Departments 1*). Tokyo: Tokyo Daigaku Shuppan-kai.

Tsuji, Tomoki. (2010). *Doitsu ni okeru Nihon-gaku no Manazashi. Kāru Furōrentsu no Nihon "Shōsan" ni Hisomu "Besshi" no Kōzu* (*A Look of German Japanology. Construction of "Contempt" Hidden in Karl Florenz's "Praise" for Japan*). Not published: Doctoral thesis to University of Tsukuba.

Wachutka, Michael. (2001). *Historical Reality or Metaphoric Expression? Culturally Formed Contrasts in Karl Florenz' and Iida Takesato's Interpretations of Japanese Mythology*. Hamburg: Lit.

Wachutka, Michael. (2016). Der kokutai-Diskurs im Staatsverständnis des 19. und frühen 20. Jahrhunderts: Kerngedanken und Entwicklung der Idee von Japans einzigartigem Nationalwesen. In *Staatsverständnis in Japan. Ideen und Wirklichkeiten des japanischen Staates in der Moderne*, 65-88. Baden-Baden: Nomos.

Yamamoto, Masami. (2014). *Nihon Kyōiku-shi. Kyōiku no "Ima" wo Rekishi kara Kangaeru* (*A History of Education in Japan. "Today" of Education in historical perspective*). Tokyo: Keiōgijuku

Daigaku Shuppan-kai.

【Research Article】

Psychometrics in L2 Groupwork: Development of the L2 Group Cohesion Scale

Deborah Maxfield

Abstract

Aim

Many communicative language teaching (CLT) classes require peer-to-peer cooperative learning and teamwork, and group cohesion has repeatedly been proven as important for motivation and task success in second-language (L2) contexts. Although psychometric scales have been developed to evaluate various aspects of the L2 learning experience, such as L2 anxiety and motivation, at present, no scale to evaluate L2 cohesion exists. Therefore, this study aimed to develop a new measurement tool, the L2 Group Cohesion Scale (L2GCS), by which L2 teachers can readily assess student experiences of working with others.

Procedure

An initial pool of 14 items investigating student experiences of group climate, L2 anxiety, and peer support was responded to by Japanese undergraduate students ($N = 98$). Items were tested using Pearson's correlations and *t*-tests to distinguish between weaker and stronger performing items, with exploratory factor analysis (EFA) used to uncover common factors. The L2GCS uses six items to assess two factors, Collaboration and L2 Anxiety Mitigation. Although this scale is the first of its kind, these factors appear to be consistent with established theory. The L2GCS demonstrated good-to-excellent reliability (Cronbach's $\alpha = .88$) and can be conducted and interpreted within a few minutes without in-depth statistical analysis.

Conclusions

Though further validation studies should be conducted using other student samples, the L2GCS questionnaire results appeared to constitute a valid and reliable measure of L2 cohesion that can be quickly and easily utilized by teachers to evaluate, isolate, and address issues with cohesion, L2 anxiety, and peer support.

Keywords: *L2 cohesion, L2 teamwork, L2 anxiety cohesion, CLT groupwork, cohesion questionnaire*

INTRODUCTION

Cohesion in L2 Classrooms

Group cohesion can be thought of as the unity of a group, the extent to which its members commit to and feel comfortable with the group (Dörnyei & Murphey, 2003). Cohesion has frequently been demonstrated to be important for motivation, which can be traced back to essential psychological drives via Self Determination Theory (SDT; Deci & Ryan, 1985). SDT divides the underlying rationale for behavior into two forms of motivation: intrinsic and extrinsic. Intrinsic motivation is based on the deep desire for competence and self-directed behavior; the need to feel successful and in control. Extrinsic motivation is driven by assumptions about the external consequences of behavior, such as gaining rewards or avoiding punishments.

Within a second-language (L2) learning context, one subtype of extrinsic L2 motivation is *identified regulation*, which drives students toward learning an L2 or undertaking social behaviors to achieve a valued goal (Noels et al., 2000), such as participating well with a team to succeed in a task. Writing specifically on how SDT can be applied within L2 motivation research, Noels (2013) describes

two main factors that can spark L2 motivation: *competence*, or belief in one's own ability to succeed on a task; and *relatedness*, the sense of connection with others. These twin principles relate to SDT as competence is based on self-directed intrinsic, behavior, while relatedness is connected to external, cohesion-based factors. Chang (2010) also found significant correlations between group cohesion and aspects of L2 motivation relevant to SDT, such as autonomy and self-efficacy. It was also found that cohesive groups can foster forms of external motivation, even in students who are not intrinsically motivated to study an L2 (Ushioda, 2003). Therefore, cohesion and perceived peer engagement are aspects of group dynamics that are particularly relevant for L2 motivation (Dörnyei, 1994; Tanaka, 2021).

Although the atmosphere or *climate* of the group will be collectively arrived at by the members composing it, teachers can aim to build cohesive classes to improve learning outcomes (MacWhinnie & Mitchell, 2017). Groupwork is an essential component of many CLT courses (Tanaka, 2021). The effects of groupwork have repeatedly been shown to boost motivation and improve learning outcomes in L2 classes (Pica et al., 1996). Tanaka (2021) found that L2 group work significantly affected motivation, in which greater group cohesion and engagement related to better learner experiences and improved motivation regardless of English proficiency level.

Previous ESL research has also investigated the relationship between cohesion and anxiety. Psychologically, anxiety can operate on cognitive and physiological levels as either a trait (a propensity to feel anxious in any situation) or a state, the chance of feeling anxious in particular settings (Maltby et al., 2010). One subtype termed *L2 anxiety* has been the subject of research for decades, which was recently defined by Teimouri et al. (2019) as anxiety occurring consistently and recurrently within language learning settings. L2 anxiety can reduce learners' willingness to communicate (WTC) in their L2 (MacIntyre et al., 1998), perceived competence (Ueki & Takeuchi, 2012), and retention (Poupore, 2013). However, working in cohesive groups has often been demonstrated to reduce students' L2 anxiety (Clement et al., 1994; MacWhinnie & Mitchell, 2017), as well as improve task performance (see Evans & Dion, 1991, for a meta-analysis).

Thus, from a psychological and ESL standpoint, working within a cohesive group increases self-esteem, reduces L2 anxiety, benefits task performance, and may improve memory, all of which would be beneficial to students within an L2 learning environment.

Prior Assessments of Cohesion

Meta-analyses have demonstrated that cohesion is moderately positive for group performance in various contexts (Evans & Dion, 1991; Gully et al., 1995), and measures of cohesion have evolved and adapted during a surge in research into this field (Greer, 2012).

Previous attempts to measure group cohesion have included the Group Environment Questionnaire (GEQ; Carron et al., 1985). Initially developed to evaluate cohesion in sports teams, the GEQ evaluates four factors related to social bonding and goal-based unity through 18 questions. Although the GEQ has been validated in other contexts, including educational and occupational settings, the use of both positively and negatively worded questions may have reduced internal consistency (α = .5 - .7), and some validation studies only found evidence for a two-factor model (Whitton & Fletcher, 2014). The Classroom Community Scale (CCS; Rovai, 2002) assesses student cohesion using 20 items on two subscales, Connectedness and Learning. The CCS was developed by selecting items based on content validity ratings by experts, with high internal consistency (overall Cronbach's α = .93). Both the GEQ and CCS are English-language measures designed for use by

native speakers.

At the time of writing, however, no scale exists for evaluating student experience of L2 teamwork. The ESL experience of teamwork might be very different from that within a native-language context owing to influences from different motivational systems and stressors, such as L2 anxiety. One useful way to evaluate L2 cohesion could be by using a short-form scale.

Short-form scales are shorter versions of full-length psychometric scales, which have been used in various psychological and educational contexts, including for test anxiety (Nasser et al., 1997) and socio-emotional experience in classrooms (Murray-Harvey, 2010). Short-form scales have proven useful for large-scale assessments (Heene et al., 2014), and can be valid in a variety of settings, provided that the scale's psychometric qualities - such as test-retest reliability and precision - are suitable for the settings in which they will be used (Ziegler et al., 2014).

While full-length questionnaires are required for clinical psychological diagnoses, the relative speed and ease of short-form scales make them applicable in a wider range of contexts than those offered by full-scale questionnaires. Short-form scales can be a useful way to explore links between pedagogic concepts; to explore relationships between L1 social experience and academic outcomes, Murray-Harvey (2010) used 12 items to evaluate academic performance, supportive and stressful relationships (α .74 - .89), finding strong connections between social and emotional experience and academic performance.

In contrast, Fraser et al. (1996) used an 80-item questionnaire to evaluate 10 aspects of classroom environment, including autonomy, student cohesiveness, and cooperation. Although investigating more aspects of a construct improves construct validity, increasing the number of questions tends to reduce reliability as measured by Cronbach's alpha, hence the subscales achieve varying levels of reliability, ranging from good (α = .89) to poor (α = .67). Furthermore, offering a scale with 80 questions would take considerable time for students to complete and for teachers to score. Drolet and Morrison (2001) manipulated the number of questions on a survey and found that respondents tended toward *"mindless response behavior*" (p. 200) as the number of similarly worded items increased, concluding that responding to more items takes longer and may increase response error. Hence, short-form questionnaires may provide more accurate, as well as faster, results.

As no questionnaire to evaluate student L2 cohesion experience had been found at the time of writing, it was determined that a pool of questions would be offered to a sample of L2 learners, and then exploratory factor analysis (EFA) would be used to uncover the structure of the questionnaire and find common factors. A similar method has been previously used in ESL research: Noels et al. (2000) used EFA to explore relationships between internal and external L2 learner motivation, uncovering seven subscales assessed by three to five items each (α .67 - .88; Noels et al., 2000). Mystkowska-Wiertelak and Pawlak (2016) also used EFA in their development of a questionnaire on L2 WTC, confidence, and motivation; the initial pool of 21 items was narrowed down to 13, and demonstrated good-to-excellent reliability (α = .88).

Short-form scales would seem to be a logical choice for L2 classrooms, in which offering a lengthy English-language questionnaire could affect time management of a lesson and increase the cognitive load on students. At present, no long- or short-form questionnaire exists to evaluate student experience of cohesion in L2 classes or teams, but as a shorter questionnaire would reduce cognitive load on students and be both faster and easier for teachers to use in the classroom, it was determined that developing a short-form scale would be a more effective and practical method to measure student L2 cohesion.

Previous Study

The data used for the development of this scale were originally collected in 2020 as part of a previous study (Maxfield, 2021). This questionnaire was designed to gather student's self-reported views on three interrelated constructs: team cohesion, anxiety in speaking English online, and anxiety with their team. The previous study investigated whether the use of teams allowed students to form cohesive groups and whether working in teams affected students' L2 or social anxiety.

All respondents were undergraduate students at a university in Tokyo and were enrolled in either Debate or Presentation classes, both of which were mandatory English-language courses for freshman students. Classes were held weekly during a 14-week semester, with around 20 students in each class. Previous psychological and EFL studies (such as those summarized above) had identified benefits of working within cohesive groups, including decreased social and L2 anxiety, improved task performance, and greater learning outcomes. Therefore, students were assigned to groups of four or five in the expectation that consistently working together would reduce L2 anxiety and increase cohesion, motivation, and peer L2 support. Teams worked together for four weekly lessons, spending considerable time working together on tasks, feedback, or discussions.

Quantitative data were gathered using a 28-item questionnaire regarding student experiences of online L2 use, perceptions of group cohesion and efficacy, and L2 anxiety. Both positively and negatively worded items were used to measure constructs, such as "I felt relaxed when speaking English with my teammates" and "I did not feel comfortable using English with teammates". Results indicated that a great majority of students had perceived their team experiences as helpful and enjoyable; 91.9% of respondents agreed that "I enjoyed working with my teams", and 93.9% that "working with a team helped me in this class" (Maxfield, 2021). Furthermore, groups with a positive social climate reported improved task achievement and reduced L2 anxiety in comparison with less cohesive teams. Correlations of around $r = .7$ can be regarded as "strong" (Dancey & Reidy, 2007); therefore, the relationship between "working in a team helped me to speak English" and "I felt relaxed with my teammates" ($r = .7$) suggests a strong link between positive social climates and improved L2 performance when students were able to form cohesive groups.

However, 11 of the 28 items used in the prior study referred to using an L2 online and hence would be irrelevant for face-to-face classes. As no previous scale has been developed for assessing group cohesion within an L2 environment, the first priority should be to develop a scale that is useful within the majority of learning environments. The Japanese Ministry of Education, Culture, Sports, Science and Technology , in line with the declining numbers and severity of COVID-19 cases, and perhaps concerned about student experience, recommended in October 2020 that universities resume face-to-face classes where possible (Government policy to name schools, November 2020), with the 2021 academic year seeing many Japanese educational institutions return to in-person classes. As developing an L2 team cohesion scale that can be used in the majority of learning environments should be a priority, it was hypothesized that using only 14 questions relating to team cohesion and anxiety might create a more streamlined and widely applicable scale for evaluating team cohesion in an L2 context.

As there remains a need for valid and reliable questionnaires to measure student cohesion (Lockee, 2021), this study aimed to develop and analyze the psychometric properties of a new questionnaire to evaluate cohesion in an L2 environment, which may be the first of its kind.

Study Overview

To create a valid and effective scale, items were retained or rejected based on skew, inter-item correlations, *t*-tests, EFA, and Cronbach's alpha if deleted. In terms of criterion validity, *t*-tests were undertaken to establish whether items could meaningfully discriminate between high and low scorers. Factor analysis was used to uncover the unknown number of factors, and various iterations were investigated to discover the best fit - items that load strongly onto one factor can be supposed to possess good construct validity. Reliability of the overall L2 Group Cohesion Scale L2GCS and the subscale Collaboration were evaluated using Cronbach's α, and *t*-tests were undertaken to establish whether these could significantly distinguish between high- and low-scoring groups.

METHOD

Design

This study utilized principal component analysis to explore relationships between questionnaire items and uncover related underlying psychological factors, termed as loaded onto. Data were collected using electronic questionnaires previously approved by the ethics review committee of the university.

Participants

All participants (N = 98) were undergraduate students enrolled in a university in Japan. All students were on one of two mandatory English courses: Presentation (N = 43, 43.9%) or Debate (N = 55, 56.1%). The classes were grouped by proficiency level, with students in Level 2 (N = 19, 19.4%) possessing greater English proficiency than those in Level 3 (N = 79, 80.6%). As part of providing consent, all respondents were asked to only complete the questionnaire if they were over 18 years old; although no demographic information was collected as part of this questionnaire, as all participants were in their first year of university, it is likely that respondents were largely aged between 18 and 20.

Materials

The data used in development of the short-form scale were originally collected in 2020 as part of a previous study (Maxfield, 2021) on student's self-reported views on three interrelated constructs: cohesion, anxiety in speaking English online, and anxiety with their team (Appendix 1). The initial study used 28 items that were translated into the students' L1, Japanese. Questions relating to experiences of online learning or open questions were removed from analysis as these topics lie outside the scope of the current paper; hence this study only uses data from 14 items.

Procedure

Ethical approval was gained before data collection began. The questionnaire was offered electronically using a Google Form. Participants read a paragraph written in both English and Japanese at the top of the questionnaire informing them of the research aims and the use of data,

which stated that they should only complete the following questions if they gave consent and were over 18 years old.

To respond to the questionnaire, participants first ticked boxes to indicate their class (Debate or Presentation) and proficiency level (Level 2, Level 3, or Prefer not to say). Students then responded to questions by clicking a box on a 6-point Likert scale that corresponded to their view. Positively valanced questions such as "I enjoyed working with my teams" were scored from 1 (*strongly disagree*) to 6 (*strongly agree*), meaning that a higher score indicated a more positive experience. The questionnaire also used negatively worded items such as "It was difficult to talk with my team"; to avoid student confusion and maintain consistency, these used the same response scale of 1 (*strongly disagree*) to 6 (*strongly agree*) as the positively worded questions, but scores were then reverse coded in SPSS. This process ensured that higher scores related to a better experience for all questions with possible total scores ranging from 17 to 102. Completing the questionnaire was estimated to less than 10 minutes.

Data Analysis

All data were entered into SPSS, reverse-scored where needed, and checked for missing or impossible scores (for instance, a response recorded as 10 on a 1 to 6 Likert scale). Descriptive statistics (Table 1) for each item were generated and checked to investigate distribution and outliers. The initial pool of items was narrowed down by assessing skew, criterion validity (through mean correlation), and *t*-tests, with weaker items removed from the analysis at each stage. Pearson's correlations evaluated relationships between items, and independent samples *t*-tests were conducted to determine whether items could meaningfully distinguish between high and low scores to test whether items showed good criterion validity.

Principal components analysis (PCA) is a statistical technique used to uncover factors shared by a group of questions. One subtype of PCA is EFA, which can examine relationships between variables without a predetermined hypothetical model (Parsian & Dunning, 2009). As no existing questionnaire on L2 cohesion could be found at the time of publication, EFA seems the most suitable method for delving into this new field. Moreover, as EFA aims to uncover the smallest number of factors needed to explain the greatest portion of variance in a dataset (Dancey & Reidy, 2007), it also lends itself well to development of short-form scales that rely on fewer factors and items than longer questionnaires. EFA was run several times to uncover the most accurate model for the data.

RESULTS

Table 1

Descriptive statistics for the initial 14 questionnaire items: mean, standard deviation, skew, and kurtosis, mean Pearson's correlation (M r), initial EFA factor loadings, and Cronbach's α if deleted (N = 98)

	Item statistics					Factor loadings			
	M	*SD*	Skew Z-score	Kurt. Z-score	*M r*	1	2	3	α if deleted
I felt relaxed when speaking English with my teammates	4.42	1.08	-0.67	-1.60	.42	**0.85**	0.05	0.07	0.54

Working with a team helped me to speak English	4.82	0.94	-1.42	-1.57	.48	**0.76**	-0.02	0.12	0.55
Talking with my teammates helped me to feel less anxious in class	4.96	0.98	-4.08	3.08	.33	**0.75**	0.06	-0.17	0.57
Working with a team helped me in this class	5.10	0.92	-2.88	-1.08	.51	**0.73**	-0.11	0.14	0.55
There was good teamwork in my teams	4.90	0.98	-1.92	-1.10	.45	**0.69**	-0.08	0.08	0.56
I enjoyed working with my teams	4.99	0.96	-2.92	-0.19	.56	**0.51**	-0.21	**0.46**	0.55
R-Sometimes my teams didn't work well together	2.69	1.35	1.88	-1.51	.39	0.00	**0.74**	-0.05	0.65
R - It was difficult to talk with my team	2.46	1.37	3.46	-0.23	.41	0.06	**0.71**	-0.32	0.66
R - My teammates rarely helped me	2.20	1.62	5.25	0.71	.33	-0.11	**0.67**	0.29	0.61
R - I did not feel comfortable talking with teammates	3.01	1.20	1.29	-0.82	.35	0.18	**0.65**	-0.26	0.61
R - I did not like working with the same people in several lessons	2.75	1.40	2.29	-1.51	.33	-0.24	**0.59**	0.26	0.63
I felt more relaxed when speaking English with my teammates than with other students in class	4.41	1.21	-1.25	-1.69	.34	0.13	0.01	**0.69**	0.56
I felt relaxed with my teammates	4.90	.96	-2.06	-0.77	.50	0.28	0.00	**0.68**	0.55
It was easy to make friends with my teams	4.30	1.25	-1.75	-0.52	.35	0.12	-0.08	**0.67**	0.57

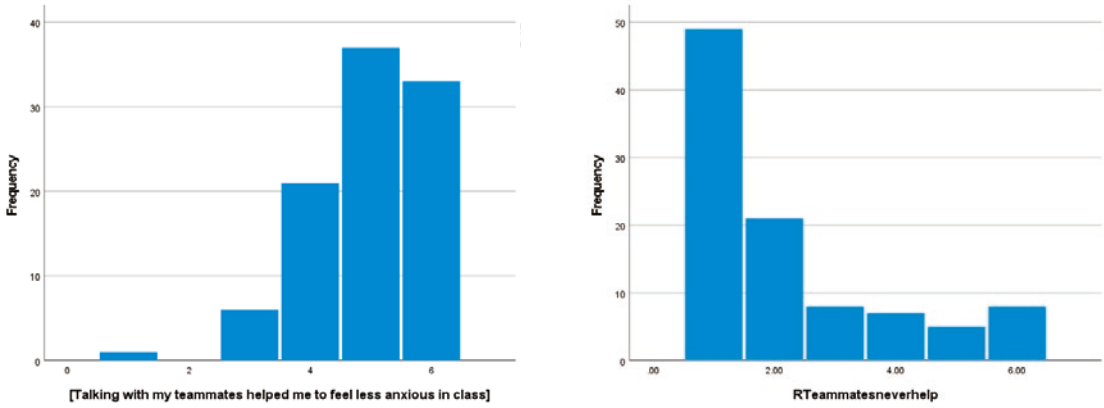
R – reverse coded; these negatively worded questions were reverse coded in SPSS

The data were checked for outliers to prevent these distorting mean values which mayd affect later analysis. Scatterplots did not indicate any impossible or irregular scores and there were no evident outliers. Checks of the minimum and maximum values (1 - 6) confirmed that no impossible scores had been entered and that no scores were missing.

Parametric assumptions were checked to determine the best correlational analysis. For the majority of items, histograms represented fairly normal distributions, and skew and kurtosis were within acceptable limits for medium-sized samples (≤3.25; as defined by Kim, 2013). However, negative skew was observed for "Talking with my teammates helped me to feel less anxious in class" and positive skew for "Teammates rarely helped me" (Figure 1 and 2); therefore, these items were further investigated as questions with extreme skew may display low differentiation (i.e., all respondents answer the same way). Quantile-quantile (QQ) plots of standardized residuals revealed a normal distribution for "Talking with my teammates helped me to feel less anxious in class"; however, points on the QQ plot for "Teammates rarely helped me" did not lie closely along the

Figures 1 and 2

Histograms investigating skew for Talking with my teammates helped me feel less anxious in class and Teammates rarely helped me



normal distribution line and showed several deviations. Levene's test (p = .001) and the Shapiro-Wilk tests (W = .74, p = .001) for this item were highly significant, indicating extreme non-normality. Based on these results, it was therefore determined to remove "Teammates rarely helped me" from further analysis due to high skew, but to retain "Talking with my teammates helped me to feel less anxious in class".

Parametric assumptions being met for the 13 remaining items, it was determined that Pearson's correlations would be appropriate to explore relationships. This is a key step in questionnaire development, as questions should be somewhat related to each other in order to measure the main construct; therefore, questions that are not related to many aspects of the construct being measured (i.e., those which produce very few significant correlations) should be removed. A correlation matrix was generated to check for singularity ($r \leq .1$, none found) and multicollinearity ($r \geq .8$), which would indicate that there were no practical differences between items; as correlations were below r = .70, there were practical differences between items. A correlation around r = .3 suggests a rather weak relationship (Dancey & Reidy, 2007); in terms of EFA, this could indicate items that poorly relate to each other or measure multiple factors; therefore, items with nonsignificant ($p \geq .01$) or weak correlations were also checked.

As negatively worded items (such as "I did not feel comfortable talking with teammates") tended to correlate only with other negative items, it was determined that inter-item correlations for negative and positive items should be considered separately. The average inter-item correlations for positive items (N = 9) were checked, and the mean correlation (M r = .44) was used as the criterion value for retention: any questions that had average correlations well below this number were therefore the weaker-performing items and were be removed. Three items with average correlations of r = .33 - .35 were removed at this stage. In terms of negative items (N = 5), the mean correlation was slightly lower at r =.37; hence, two items with mean correlations below this criterion value were removed (r =.33 - .35).

Exploratory Factor Analysis

As the eight remaining items displayed generally fair to strong Pearson's correlations (r =.4 - .7; Dancey & Reidy, 2007), the next stage was to uncover the factors to which these questions were related via PCA. As research on student experiences of L2 cohesion has not been undertaken in the

past, it was unclear how many factors may exist, therefore the EFA subtype of PCA was used to discover underlying factors. The dataset was determined to be suitable for PCA, as the initial Kaiser-Meyer-Olkin (KMO) statistic .823 suggested 'great' sampling adequacy based on Kaiser's thresholds (Parsian & Dunning, 2009) and Bartlett's test of sphericity was significant at $p > .001$.

The first round of EFA (Table 1) had been conducted on all of the initial 14-items to establish Cronbach's alpha for the full questionnaire undertaken in 2020, which indicated merely adequate reliability, $\alpha = .63$. The initial EFA had a KMO of .772, defined as 'good' (Parsian & Dunning, 2009) and Bartlett's test of sphericity was significant at $p > .001$. This solution explained 67.2% of the total variance and indicated four factors, although several items loaded onto more than one factor at above 0.4, indicating that the question does not reliably measure only one aspect of the construct under investigation. The scree plot appeared to show a two- or three-factor model, depending on how the bend of inflection was interpreted, as simply using all eigenvalues above 1.00 without reference to the scree plot does not guarantee the best solution (Cattell, 1966; Costello & Osborne, 2005). It was hoped that after removal of lower performing items based on skew and low correlations, the shorter eight-item PCA would reveal a better model of the different factors.

First, a Varimax method of PCA was undertaken, which analyzes variance under the assumption that the factors are not related. Two factors were extracted with eigenvalues above 1, which cumulatively explained 67% of the variance in the dataset.

The scree plot also showed a point of inflexion commensurate with a two-factor solution. However, use of oblique rotation, such as Oblimin, renders a more accurate solution than orthogonal if the factors are related (Costello & Osborne, 2005, p .3).

As cohesion, team performance, and L2 anxiety are related constructs (Maxfield, 2021), it seemed likely that any factors extracted could be related; therefore, Oblimin rotation was also undertaken. This solution also explained 67% of the variance in the dataset through two factors, although two questions loaded at around .4 onto both factors. Reliability analysis was undertaken for the eight-item scale, which showed higher Cronbach's alpha if two items were removed; however, removing items tends to alter all item loadings onto factors when the EFA is re-run without these items, and therefore removing items purely based on *Cronbach's alpha if deleted* does not necessarily build the best questionnaire. Various iterations of Oblimin rotation were also compared with or without these and other items to test four- to six-item scales, and by forcing a three-factor extraction to evaluate which model was the best.

After several rounds of EFA testing and comparisons with Cronbach's alpha for each version, the best fit for this data set was determined (Table 2). The final solution used six positively worded items that all loaded clearly onto one of two factors: this had a KMO of .813, explained 75.92% of the variance in the dataset, and Cronbach's $\alpha = .88$ indicated good-to-excellent reliability (Dancey & Reidy, 2007). This solution was deemed the most suitable as it matched the scree plot, revealed stronger factor loadings than on any other orthogonal or oblique analyses, and grouped items logically. As scale properties did not improve after testing with further item removals, , this solution was henceforth termed the L2 Group Cohesion Scale (L2GCS).

Table 2

Factor analysis of the six-item L2GCS: factor loadings for Collaboration and L2 Anxiety Mitigation, correlations with overall L2GCS, correlation with subscale, Cronbach's alpha if deleted, and t-test statistic

Item	Factor Loadings		Item-L2GCS	Item-Subscale	α if	
	Collab.	L2 A. M.	*r*	*r*	deleted	*t*
Working in a team helped me in this class	**.94**	.014	.82*	.86*	.85	10.93**
There was always good teamwork in my teams	**.83**	.09	.78*	.75*	.86	10.94**
I felt relaxed with my teammates	**.76**	.05	.71*	.75*	.84	13.70**
I enjoyed working with my teams	**.70**	.27	.83*	.85*	.84	10.55**
Working with my team helped me to speak English	**.69**	.18	.79*	.80*	.85	10.30**
I felt relaxed when speaking English with my team	.04	**.97**	.66*	-	.85	9.10**
Subscale Items	5	1	**L2GCS Items**	**Total**	6	
Subscale α	.89	-	**Total α**		.88	
Variance Explained	62.73%	12.24%	**Total Variance**		75.92%	

*$p < .01$ ** $p < .001$

Factor 1 was able to explain 62.73% of the variance in the dataset. Six items had loadings for this factor between .70 and .94, suggestive of strong fit with the factor. These items related to cooperation, group climate, and peer assistance; therefore, this factor was labeled *Collaboration*.

Factor 2 explained 12.24% of the variance through a single item "I felt relaxed when speaking English with my teammates", with a very high loading of .97 on this factor. As this item relates to diminished anxiety while using an L2 with a team, it was termed *L2 Anxiety Mitigation*.

Although it is impossible to test Cronbach's α for a single-item measure such as L2 Anxiety Mitigation, Cronbach's alphas for the overall L2GCS and for the Collaboration subscale were .88 and .89 respectively, demonstrating good to excellent reliability (Cooper, 2020). This suggests the L2GCS and Collaboration subscale each demonstrated high internal consistency. To ensure whether reliability could be improved by removing any items, *Cronbach's alpha if item deleted* was checked for the whole and subscale, but it was found that removal of any items would reduce rather than improve reliability.

To establish whether the scale could reliably distinguish between high and low scores, participants were sorted into three groups (Group 1 = low, Group 2 = medium, Group 3 = high). Independent samples *t*-tests were undertaken by comparing their total score against all six questionnaire items to check whether Group 3 had scored significantly higher than Group 1. The mean scores for Group 3 were higher than Group 1 for each of the six items, and all *t*-tests were significant at p = .001, suggesting that these items could significantly discriminate between high- and low-scoring groups. As all items displayed good levels of item discrimination, no further questions were removed.

Finally, the L2GCS was tested for construct validity. Where possible, new scales should be compared against existing measures to evaluate overlap between them, which can determine whether they possess convergent validity if new measures correlate well with existing scales. However, as no previous measure of L2 cohesion could be found at the time of publication, it was impossible to evaluate convergent validity for the L2GCS at this time.

Table 3
Pearson's correlations between L2GCS, subscales, and discriminant validity item

	Collab.	L2 A. M.	L2GCS
Collab.	-		
L2 A. M.	.50**	-	
L2GCS	.90**	.66**	-
Discriminant	.04	.15	.07

** $p < .001$

Another form of construct validity, termed discriminant validity, could be tested however. This method uses bivariate correlations to compare a new scale against an unrelated construct. Very low or statistically insignificant correlations would indicate that this scale does not measure irrelevant constructs. Discriminant validity was checked by comparisons of the L2GCS and subscales with an assumedly unrelated construct ("Speaking English online is easier than speaking English face-to-face"). None of the Pearson's correlations reached significance at $p = .05$ or lower with the discriminant item (Table 3), suggesting that neither the L2GCS nor its subscales measure irrelevant constructs.

Taken together, the results indicate that the L2GCS has good-to-excellent reliability and discriminant validity. Although the L2GCS is an original measure, results from EFA and correlations indicated a strong internal structure of the L2GCS, which may indicate strong construct validity.

DISCUSSION

The L2GCS (Appendix 2) demonstrates good-to-excellent internal consistency (Cronbach's α .88) and displays discriminant validity. Although the L2GCS is an original measure, results from EFA and Pearson's correlations (Table 3) indicated a strong internal structure of the L2GCS, which maps well onto existing cohesion research in the field.

The L2GCS consists of six self-report items that measure two subscales, Collaboration and L2 Anxiety Mitigation, using a 6-item Likert response scale. Collaboration (α = .89) relates to cohesion, similar to Connectedness within the CCS (Rovai, 2002), and covers social interaction within the group toward task success, as "one requires both social and intellectual interactions to accomplish learning goals" (Rovai, 2002, p. 199). L2 Anxiety Mitigation uses a single-item measure to assess students' affective experience of using an L2 with their team. The moderate correlation between the Collaboration and L2 Anxiety Mitigation subscales indicates that teams with a collaborative atmosphere tend to reduce L2 anxiety, which echoes prior findings (Clement et al., 1994; Poupore, 2013).

High- and low-scoring groups were investigated using the L2GCS. For this sample ($N = 98$), Debate class students ($N = 55$) tended to score slightly higher on the L2GCS (M 29.67, SD 4.63) than Presentation class students ($N = 43$, M 28.21, SD 4.59). Their higher average L2GCS score may be attributable to Debate students working collectively to research and develop arguments against a rival team; as their group debate skills were the subject of formal assessment during the course, it was explicitly stated that effective teamwork would be essential for successful group performance and higher final scores. However, Presentation teams fulfilled a more social than score-based role in discussions, peer support, and constructive peer-to-peer feedback. While Presentation teams may have provided social support to students, teamwork was less critical for their final grade than in

Debate classes. As reported by Gully et al.'s (1995) meta-analysis, the level of task interdependence may mediate the relationship between cohesion and task success; therefore, it seems logical that students working on interdependent tasks, such as those in Debate classes, would report higher overall cohesion on the L2GCS.

However, there was no great effect of L2 proficiency on L2GCS scores, as the mean score of Level 2 students (*N* = 19, *M* 29.11, *SD* 4.67) was only 0.4% higher than that of Level 3 students (*N* = 79, *M* 29.01, SD 4.67). This finding echoes Tanaka's (2021) conclusions that L2 proficiency does not significantly affect cohesion.

Ziegler et al. (2014) recommend that when developing a short-form scale, it is essential to address the construct being measured, the purpose of the scale, and the target population, which will be clarified here. The main purpose of developing the L2GCS was to measure cohesion within teams of L2 speakers who cooperated on shared tasks. While the questionnaire has only been tested and developed with Japanese undergraduate students (*N* = 98), it is likely that by adapting the wording of some items from 'teams' to 'class', the L2GCS could prove a useful tool for evaluating cohesion in larger groups.

Furthermore, translation of the items into the relevant L1 could allow the scale to be used internationally. It is also possible that the L2GCS may be useful with younger learners, although further testing with an appropriate sample would be required before it can be claimed that the L2GCS is valid for use with children. Though further validation studies of the L2GCS are needed before it can be reliably used with other populations, the internal consistency (α .88) and strong factor structure indicate that the L2GCS could prove a reliable instrument for evaluating cohesion in a fast, simple, and effective manner.

Limitations and Avenues for Further Study

Although the original questionnaire used both positively and negativelyworded items, the L2GCS uses only positively valanced questions. There could be debate on this point; employing both types of wording means researchers can check that respondents had not merely selected the same option for all questions without considering them carefully (for instance, a respondent selecting 'agree' to both "Working with a team helped me in this class" and "my teammates rarely helped me"). However, prior research has indicated that negatively worded questions may unduly affect factor analysis (Loomis & Wright, 2018), therefore the L2GCS uses only positively worded questions. However, a potential avenue for further research could be to compare the L2GCS with another questionnaire containing negatively worded questions, then compare the scales in terms of criterion validity and reliability.

A stricter cut-off than mean correlation could have been used to assess criterion validity. One disadvantage with this method is that it would have considerably narrowed the pool of items and thereby would have resulted in more limited, and therefore perhaps weaker, options for the final questionnaire.

The overall L2GCS and Collaboration subscale demonstrated good internal consistency (α .88 - .89), but reliability cannot be tested for the other subscale, L2 Anxiety Mitigation, as this is a single-item measure. The lack of reliability for single-item measures may trigger alarm; however, these are not always inappropriate, particularly within questionnaires that are deliberately designed as short-form scales. In support of single-item measures, Sarstedt and Wilczynski (2009) argued that single-item measures can perform acceptably on simple, singular constructs. Postmes et al. (2013) also

stated that singleitem measures can be sufficient, provided that the construct being measured is sufficiently narrow or homogenous. As the reduction of L2 anxiety provided by teammates seems to be quite a narrow construct, it would appear that a single-item measure may suffice in this instance.

Though efforts were made to maximize content validity and reliability of this new measure, it remains to be further tested before it can be claimed to be valid for other populations. The original study used participants from only one institution, which raises questions on whether it can be reliably applied to other populations (Hurley & Brookes, 1988). This is particularly problematic in cohesion research, as though the measurement of cohesion has certainly evolved over decades of research, its essential underlying factors and structures remain unresolved (Greer, 2012), and no prior scale for evaluating L2 cohesion can be used for comparison. Therefore, a logical next step would be to perform replication studies to validate the L2GCS by offering it in other settings or contexts, then using factor analysis to establish whether similar constructs and reliability were obtained.

Despite these limitations, development of the L2GCS opens new potential avenues for L2 cohesion research. For instance, the temporal stability of L2 teamwork, such as whether cohesion changes over time, could be assessed by providing a group with the L2GCS at spaced intervals during a semester and evaluating how scores change.

As no reverse scoring or artificial weighing of answer options is required, the L2GCS can be utilized in classrooms without specialized training or equipment, which will hopefully increase the universality of contexts in which it can be used. This is the first questionnaire designed for measuring group cohesiveness in teams in an L2 environment, and it is hoped that the L2GCS can provide teachers with a reliable tool for evaluating L2 collaboration and anxiety-mitigation in their classes. Gaining awareness of interpersonal dynamics within teams not only provides teachers with greater insights into student-to-student interactions, but also increases their awareness of how and where to target efforts to bolster peer-to-peer support, or to encourage social bonding.

CONCLUSION

This study aimed to develop the L2GCS, a new tool to measure student perceptions of cohesion in an L2 context, and establish its validity and reliability. Improving cohesion improves task performance, reduces L2 anxiety and increases WTC. The L2GCS offers teachers a further tool for bolstering motivation in their classes and improving learning outcomes, and one that can be undertaken in about five minutes without requiring special equipment or exhaustive statistical analysis. By applying the L2GCS, teachers in L2 university environments can better target their time, energy and resources onto issues being faced by groups, hopefully leading to more comfortable, supportive, and productive L2 learning environments for students.

REFERENCES

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245-276.

Carron, A. V., Widmeyer, W. N., & Brawley, L. R. (1985). The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire. *Journal of Sport and Exercise Psychology*, *7*(3), 244-266.

Chang, L. Y.H. 2010. Group processes and EFL learners' motivation: A study of group dynamics in EFL classrooms. *TESOL Quarterly*, 44(1). 129 - 154.

Clément, R., Dörnyei, Z., & Noels, K. A. (1994). Motivation, self-confidence, and group cohesion in the foreign language classroom. *Language Learning*, *44*(3), 417-448.

Cooper, C. (2020). *Individual Differences and Personality*. Routledge.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*(1), 7.

Dancey, C. P., & Reidy, J. (2007). *Statistics without maths for psychology*. Pearson Education.

Deci, E. L., & Ryan, R. M. (1985). Conceptualizations of intrinsic motivation and self-determination. In *Intrinsic motivation and self-determination in human behavior* (pp. 11-40). Springer.

Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The Modern Language Journal*, *78*(3), 273-284.

Dörnyei, Z., & Murphey, T. (2003). *Group dynamics in the language classroom*. Cambridge University Press

Drolet, A. L., & Morrison, D. G. (2001). Do we really need multiple-item measures in service research?. *Journal of Service Research*, *3*(3), 196-204.

Evans, C. R., & Dion, K. L. (1991). Group cohesion and performance: A meta-analysis. *Small Group Research*, *22*(2), 175-186.

Fraser, B. J., McRobbie, C. J., & Fisher, D. (1996). Development, validation and use of personal and class forms of a new classroom environment questionnaire. *Proceedings Western Australian Institute for Educational Research Forum* (Vol. 31).

Greer, L. L. (2012). Group cohesion: Then and now. *Small Group Research*, *43*(6), 655-661.

Gully, S. M., Devine, D. J., & Whitney, D. J. (1995). A meta-analysis of cohesion and performance: Effects of level of analysis and task interdependence. *Small Group Research*, *26*(4), 497-520.

Heene, M., Bollmann, S., & Bühner, M. (2014). Much ado about nothing, or much to do about something? Effects of scale shortening on criterion validity and mean differences. *Journal of Individual Differences*, *35*(4), 245 - 249.

Hurley, J. R., & Brooks, L. A. (1988). Primacy of affiliativeness in ratings of group climate. *Psychological Reports*, *62*(1), 123-133.

Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry and Endodontics*, *38*(1), 52-54.Lockee, B.B. (2021) Online education in the post-COVID era. *Nature Electronics* 4, 5 - 6.

Loomis, C., & Wright, C. (2018). How many factors does the sense of community index assess?. *Journal of Community Psychology*, *46*(3), 383-396

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, *82*(4), 545-562.

MacWhinnie, S. G., & Mitchell, C. (2017). English classroom reforms in Japan: A study of Japanese university EFL student anxiety and motivation. *Asian-Pacific Journal of Second and Foreign Language Education*, *2*(1), 1-13.

Maltby, J., Day, L., & Macaskill, A. (2010). *Personality, individual differences and intelligence*. Pearson Education.

Maxfield, D. (2021). Impact of group cohesion on anxiety and online task performance: A correlational exploratory analysis. *Journal of Foreign Language Education and Research*, *2*, 22-36.

Murray-Harvey, R. (2010). Relationship influences on students' academic achievement, psychological health and well-being at school. *Educational and Child Psychology*, *27*(1), 104.

Mystkowska-Wiertelak, A., & Pawlak, M. (2016). Designing a tool for measuring the interrelationships between L2 WTC, confidence, beliefs, motivation, and context. In *Classroom-oriented research* (pp. 19-37). Springer.

Nasser, F., Takahashi, T., & Benson, J. (1997). The structure of test anxiety in Israeli-Arab high school students: An application of confirmatory factor analysis with miniscales. *Anxiety, Stress, and Coping*, *10*(2), 129-151.

Noels, K. A. (2013). Learning Japanese; Learning English: Promoting motivation through autonomy, competence and relatedness. In M.T. Apple, D. Da Silva, T. Fellner (Eds.), *Language learning motivation in Japan*. (pp. 15-34). Multilingual Matters.

Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, *50*(1), 57-85.

Okubo, A. 2020, November 4. Japan universities baffled by gov't policy to name schools with fewer non-online classes. *The Mainichi*. Accessed online June 14, 2022 at https://mainichi.jp/english/articles/20201103/p2a/00m/0na/015000c

Parsian, N., & Dunning, T. (2009). Developing and validating a questionnaire to measure spirituality: A psychometric process. *Global Journal of Health Science*, *1*(1), 2-11.

Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners?. *TESOL Quarterly*, *30*(1), 59-84.

Postmes, T., Haslam, S. A., & Jans, L. (2013). A single-item measure of social identification: Reliability, validity, and utility. *British Journal of Social Psychology*, *52*(4), 597-617.

Poupore, G. (2013). The influence of L2 motivation and L2 anxiety on adult learners' socio-affective conditions and language production during communicative tasks. *The Asian EFL Journal Quarterly*, *15*(3), 93-128.

Rovai, A. P. (2002). Development of an instrument to measure classroom community. *The Internet and Higher Education*, *5*(3), 197-211

Sarstedt, M., & Wilczynski, P. (2009). More for less? A comparison of single-item and multi-item measures. *Die Betriebswirtschaft*, *69*(2), 211.

Tanaka, M. (2021). Individual perceptions of group work environment, motivation, and achievement. *International Review of Applied Linguistics in Language Teaching*.

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *41*(2), 363-387.

Whitton, S. M., & Fletcher, R. B. (2014). The Group Environment Questionnaire: A multilevel confirmatory factor analysis. *Small Group Research*, *45*(1), 68-88.

Ueki, M., & Takeuchi, O. (2012). Validating the L2 motivational self system in a Japanese EFL context: The interplay of L2 motivation, L2 anxiety, self-efficacy, and the perceived amount of information. *Language Education & Technology*, *49*, 1-22.

Ushioda, E. (2003). Motivation as a socially mediated process. Little, D., Ridley, J., & Ushioda, E. (Eds.), *Learner autonomy in the foreign language classroom: Teacher, learner, curriculum and assessment* (pp. 90-102). Dublin: Authentik.

Ziegler, M., Kemper, C. J., & Kruyen, P. (2014). Short scales–Five misunderstandings and ways to overcome them. *Journal of Individual Differences*. 35 (4)

APPENDIX 1
14-item questionnaire from previous study

☐ Presentation ☐ Debate
☐ Level 2 ☐ Level 3 ☐ Prefer not to say

It was easy to make friends with my teams
Working with a team helped me in this class
I enjoyed working with my teams
There was good teamwork in my teams
I did not like working with the same people in several lessons
I felt relaxed with my teammates
Sometimes my teams did not work well together
My teammates rarely/never helped me in class
It was difficult to talk with my team
I did not feel comfortable talking with teammates
Talking with my teammates helped me to feel less anxious in class
I felt relaxed when speaking English with my teammates
I felt more relaxed when speaking English with my teammates than with other students in class
Working with a team helped me to speak English

APPENDIX 2
Recommended L2GCS format and instructions (English version)

Instructions to students: *Select the option (Strongly disagree, slightly disagree... strongly agree) that best matches your experience*

	Strongly disagree	Slightly disagree	Disagree	Agree	Slightly agree	Strongly agree
Working in a team helped me in this class	☐	☐	☐	☐	☐	☐
There was always good teamwork in my teams	☐	☐	☐	☐	☐	☐
I felt relaxed with my teammates	☐	☐	☐	☐	☐	☐
I enjoyed working with my teams	☐	☐	☐	☐	☐	☐
Working with my team helped me to speak English	☐	☐	☐	☐	☐	☐
I felt relaxed when speaking English with my team	☐	☐	☐	☐	☐	☐

【Research Article】

Incorporating Instructional Design Theory Into Presentation Slides to Improve Learner Engagement

Ian Hart

Abstract

With the increased popularity of technology-assisted learning in second language classrooms, English as a foreign language (EFL) teachers are looking for new and effective ways to provide instruction in various teaching environments. The focus of this study is in the area of instructional design theory (IDT) and blended learning (i.e., the use of technology in educational environments). This study aims to a) examine whether existing instructional design models can be used to design, develop, and implement more effective, personalized, and efficient instruction using presentation software (e.g., Microsoft PowerPoint, Google Slides) and b) investigate whether instructional design methods should replace traditional teaching methods such as teacher-fronted instruction or the use of coursebooks/textbooks. After the discussion of existing instructional design models, instructional presentation slides are designed using IDT principles, and then tested on second language learners. Qualitative and quantitative data were collected and analyzed through interviews and questionnaires. The findings from the study provide evidence that an instructional design model can be used to design technology-assisted materials that provide learners with more personalized and meaningful instruction, and based on the evidence presented in this study, it is recommended that more attention should be given to design principles when creating instructional materials for L2 learners.

Keywords: *Instructional Design Theory, Blended Learning, IDT, Motivation*

Introduction

Over the past few years, restrictions caused by the COVID-19 pandemic have forced teachers in all fields to consider new ways of conducting their lessons and providing instruction to learners. This has caused many teachers to turn to technologies and online platforms that give them the ability to present information in new and effective ways (Turan & Akdag-Cimen, 2020). In an earlier study (Hart, 2022), I looked at how online platforms could be used to provide learners with a "flipped classroom," where the presentation of course content takes place outside the classroom. For this to be achieved, PowerPoint slides were used to provide asynchronous instruction. While the focus of that study was on out-of-class learning and production, positive reactions to the PowerPoint slides were observed by the learners, other teachers who the slides were shared with, and course leaders. This led to an analysis of the effectiveness of well-designed slide presentations for English as a foreign language (EFL) instruction.

As a qualified product designer, I felt that my experience and knowledge helped me when designing teaching and learning materials for my second language learners. Design thinking is about the creator putting themselves in the shoes of the person who will be interacting with the design. With this belief echoing similar principles set by a student-centered approach to teaching, existing design theories that could be used to design and facilitate effective and student-centered instruction were investigated. While learners in a language classroom may come from similar educational backgrounds, they all have individual behaviors, values, interests, and goals. Therefore, this study focuses on instruction that is both learner-centered and customizable. One design approach that has

received increased attention in recent years is instructional design theory (IDT). Reigeluth (1999) describes IDT as a "theory that offers explicit guidance on how to better help learn and develop" (p. 5). Other definitions include IDT as being "prescriptive in nature" and offering "proven guidelines for creating optimal learning environments for intended learning content and the target audience" (Huang, 2013, p. 19). Like design theories in other fields, IDT puts emphasis on individuals with a learner-centered paradigm at its core. The role of an instructional designer is "translating principles of learning and instruction into specifications for instructional materials and activities" (Smith & Ragan, 1993, p. 12). By performing a needs analysis of the subjects (i.e., learners), more personalized and meaningful instruction can be designed and developed using instructional design models. While early instructional design theories took a behavioral approach to design, more up-to-date models take a constructivist approach, with constructivism receiving increased attention in several different disciplines, including language teaching/learning and instructional design (Bednar, Cunningham, Duffy, & Perry, 1991). In this study, existing instructional design models will be discussed and subsequently considered for the design of instructional materials in a second-language classroom.

In addition to the use of IDT, this study explores the use of instructional technology, such as presentation software, to help make the learner-centered paradigm more efficient and effective. With many English language courses being standardized with set learning materials, making instruction personalized to individual learners is challenging. However, Reigeluth (2014) explains that by using technology, the learners' individual needs and expectations can be accommodated, and more relevant and personalized learning experiences can be provided through the customization of instructional materials. This use of technology in an educational environment is known as "blended learning" (Vaughan, 2007), which is defined as being both a student-centered and flexible approach to learning. Focus is put on the design and use of Microsoft PowerPoint presentation software, due to its popularity, and with supporting research regarding the use of slide presentations to provide more interesting and motivational instruction (Szaboa & Hastings, 2000; Catherina, 2006; Wanner, 2015).

Finally, a research study was conducted. Instructional design models were used to design and evaluate instructional PowerPoint slides for two types of content and language integrated learning classes. Pre-study research was conducted through teacher interviews and questionnaires, and data were collected to provide a needs analysis in order to support the design and development of the specially designed slides. The slides were tested on second language university students at Rikkyo University, and post-study research was conducted regarding the effectiveness of the instruction to conclude whether IDT is effective in improving learner attitudes.

Literature Review

What is Instructional Design Theory?

Reigeluth and An (2021) describe IDT as a "deliberate and orderly, but flexible, process for planning, analyzing, designing, developing, implementing, and evaluating instruction in education" (p. 1). In education, IDT refers to a knowledge base that provides guidance on how to facilitate learning under different conditions (Reigeluth, 1999a). Reigeluth states that IDT is distinguished from learning theories. Rather than describing how learning occurs through descriptive explanations, IDT is prescriptive and design oriented (Reigeluth, 1999a). While it is not easy to apply the knowledge of learning theories to educational problems, IDT can identify methods for specific situations by offering detailed guidelines to design instruction. An early instructional design theory, called

"Theory One," was presented by Perkins (1992) that provided guidelines for what instruction should include to foster cognitive learning:

1. *Clear information:* Descriptions and examples of the goals, knowledge needed, and the performance expected.
2. *Thoughtful practice:* Opportunity for learners to engage actively and reflectively with whatever is to be learned – adding numbers, solving word problems, writing essays.
3. *Informative feedback:* Clear, thorough counsel to learners about their performance, helping them to proceed more efficiently.
4. *Strong intrinsic or extrinsic motivation:* Activities that are amply rewarded, either because they are very interesting and engaging in themselves or because they feed into other achievements that concern the learner.

(Perkins, 1992, p. 45, as cited by Reigeluth, 1999a)

To build on the above guidelines, Reigeluth (1999b) discusses the major characteristics that all instructional design theories have in common. The four major components of IDT are listed by Reigeluth as being:

1. *Instructional values:* To maximize the effectiveness of instruction, the values of the design theory should match those of the users. Therefore, IDT is only relevant to teachers who see the value in a learner-centered approach.
2. *Instructional outcomes:* These are the outcomes expected by the instructors, including expectations related to effectiveness, motivation, efficiency, and appeal. The focus of the outcome depends on the values of theories.
3. *Instructional conditions:* This includes factors that influence the selection of different instructional methods. Reigeluth (1999b) lists these as a) the nature of what is to be learned, b) the nature of the learner, c) the nature of the learning environment, and d) the nature of the instructional development constraints. These conditions "may influence which methods will work best to attain your desired outcomes"(p. 8).
4. *Instructional methods:* With instructional design theories being design oriented, they have methods that are situational and componential (i.e., can be done if different ways and made of different components). In addition, there are different ways in which these methods can be performed, depending on the way in which problems are presented or each scenario's characteristics. Therefore, some methods are "better than others (better for a given set of conditions and desired outcomes), but sometimes they're equally efficacious" (p. 10). When selecting methods, it is important to consider the values, outcomes, and conditions. It is also important to remember that these methods are probalistic, therefore, they "do not guarantee the desired instructional and learning outcomes" (p. 11). However, the goal is to attain the highest possible probability of the learners achieving their goals.

To sum up, instructional design theories offer methods that are situational, componential, and probabilistic. They help identify situations for which the method can be applied and identify the values of the goals they pursue to successfully attain them.

This has led to the creation of a variety of methods or models that can be used depending on the situation and requirements of the instructor. To find an appropriate model for this study, this paper

will consider two existing instructional design models.

Instructional Design Models

The most well-known ID model is the analysis, design, development, implementation and evaluation (ADDIE) model (Gustafson & Branch, 2002). Another more recent model is Reigeluth and An's (2021) holistic 4D model. The ADDIE model uses a systems approach in designing instruction, which begins with an analysis process that breaks down what should be taught into pieces, leading to the design of instruction for each of those pieces. The 4D model provides a more holistic approach that begins with a less clear vision of the instructional design, and then proceeds to work out progressively more details in additional cycles.

The ADDIE Model

The ADDIE model is the most widely used and simple approach to instructional design. It is generally agreed that it provides the most essential steps to the instructional design process (Molenda, 2003; Reiser & Dempsey, 2002). *Figure.1* shows an adaptation of the ADDIE model.

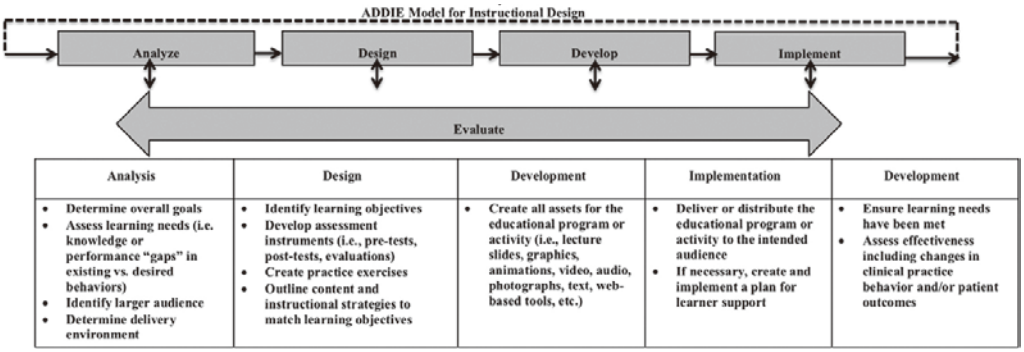


Figure.1: The ADDIE Model (Patel et al, 2018, p. 3)

The process begins with an analysis of the learner, instructional materials, and context to determine the overall goals. This includes identifying characteristics of the target learner, which may include their existing knowledge, values, motivations, and interests. After identifying the learning objectives, the design stage is used to create exercises, and outline content and instructional strategies to match the learning objectives. Instructional strategies include pre-instructional activities, content presentation, and pre-tests and/or evaluations with possible learner participation (Molenda, 2003). In addition, the types of media and delivery methods that will be used in the development stage are decided. The development stage includes the creation of all the assets for instruction. Regarding this study, this would include the development of presentation slides (e.g., Microsoft PowerPoint, Google Slides), containing media such as graphics, videos, photographs, audio, and animations. The implementation stage is where the developed materials is delivered or distributed to the learners. Learner support can be created, if necessary. Formative evaluation takes place throughout the entire process, allowing for changes and improvements to be made before implementation. In addition to the ongoing formative evaluation, summative evaluation takes place after the implementation stage to assess overall effectiveness.

The Holistic 4D Model

A more recent instructional design model is Reigeluth and An's (2021) holistic 4D model. They explain that the holistic approach "begins the design process by creating a fuzzy vision of the instructional system (top-level design) and proceeds to work out progressively more details for each part of it in two more cycles (mid-level and lower-level design) so that each part is designed with the other parts in mind" (p. 13). The benefit of this process is such that a) designers do not get mired in the details during the initial envisioning process, b) information obtained from the analysis is used immediately so it is still fresh, and c) all of the most important information is analyzed (Reigeluth & An, 2021). The cycles can be seen in *Figure.2*.

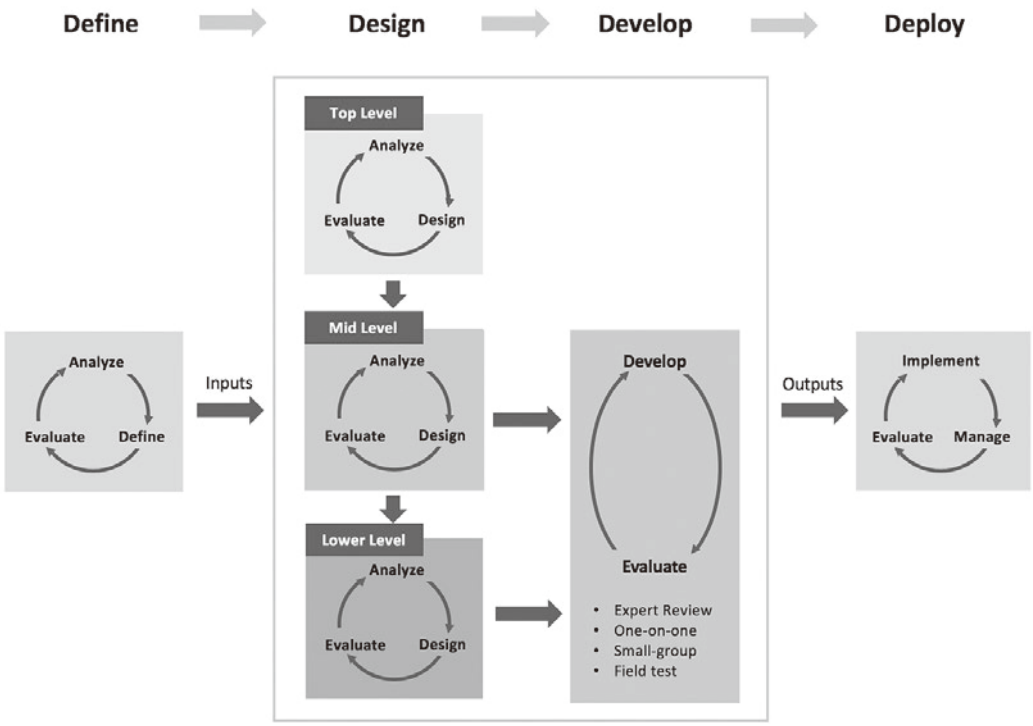


Figure.2: The Holistic 4D Model (Reigeluth & An, 2021, p. 22)

Within each design stage, analysis is required. At the top level (strategic), analysis is used to begin to determine what and how to teach. General content is identified that is later examined in the mid-level (operational) analysis and gives a "fuzzy vision" of what instruction should be like. The next level of clarity is provided in the mid-level, the general information gathered from the top level provides a meaningful context allowing for more detailed information to be gathered about what to teach and how to teach it. The lower-level design (tactical) is where a detailed blueprint for instruction is created for both "task focus" and "topic focus." As stated by Reigeluth and An (2021), "different kinds of learning require different kinds of mental processing for learning to occur, which in turn require different methods of instruction to foster the cognitive processes" (p. 14). The center box displayed in *Figure.2* represents the instructional design functions with "just-in-time analysis" and "ongoing evaluation" (p. 15). The 4 Ds in this design represent define, design, develop, and deploy, which are summarized here:

Define: The left box represents exercises that take place before the design process. They include "the

analysis of the need for instruction, carried out in a holistic and integrated manner that considers other forms of intervention" (p. 15). As the need for instruction is evaluated, project goals and objectives are defined.

Design: In the box, on the left side are three design levels. Within each level is an iterative process of analysis, design, and evaluation. Here, the analysis of learners, materials, and context is conducted, leading to decisions on both what and how to teach. The decision whether to conduct all kinds of analysis depends on the nature and scope of the project.

Develop: In the box, on the right side, the development process can be seen. The diagram also shows that careful evaluation takes place at each stage of development. However, evaluation at this stage differs to that of the design stage, as it involves testing the instruction on learners to make changes or improve it.

Deploy: The right side of the diagram represents the output, where activities are conducted following design and development. The system is used for regular, full-scale instruction (e.g., in schools, workplaces, and training workshops). The system is delivered by instructors, who manage and evaluate it. In addition, summative evaluation takes place at this stage.

Why use Instructional Design Theory?

EFL instructors often use a content perspective when approaching instruction, with a focus on *what to teach* their students. While instructional designers approach instruction from a problem-solving perspective that also includes what to teach, they also pay considerable attention on how to teach in a way that is effective, efficient, and motivational (Reigeluth & An, 2021). Many second language learners may lack the experience or knowledge to understand the content that is being presented to them; therefore, deciding the best way to teach it can be a complex problem that does not have a single solution. This means that instruction must "change from standardization to customization, from a focus on putting things into learners' heads to a focus on helping learners understand what their heads are into" (Reigeluth, 1999b, p. 19). Reigeluth (1999b) highlights the following three requirements:

- It requires a shift from passive to active learning and from teacher-directed to student-directed learning.
- It requires a shift from teacher initiative, control, and responsibility to shared initiative, control, and responsibility.
- It requires a shift from decontextualized learning to authentic, meaningful tasks.
- Most importantly, it requires a shift from holding time constrants and allowing achievement to vary, allowing each learner the time needed to reach the desired achievements.

(p. 19)

By defining the problem; determining what knowledge, skills, and attitudes are needed; focusing on what the learner need to learn and can achieve, and finally determining the methods that will best help the learner to master the content, instruction can be more effective, efficient, and motivating: a learner-focused paradigm.

Over the last decade, IDT has increased in popularity in the EFL/ESL industry, with comparisons being made to existing theories on second language acquisition. Examples can be found in the *Table.1*.

Second Language Acquisition	Instructional Design
Task-based language teaching (Nunan, 2004)	Backward design (Wiggins & McTighe, 2005)
Schema-based learning (Plaget via Wadsworth, 2004)	Generative learning (Osborne & Wittrock, 1983)
Student-centered pedagogy (Kumaravadivelu, 2003); Scaffolding (Vygotsky, 1980)	Learner-centered teaching (Weimer, 2013); Universal design for learning (Hall, Meyer, & Rose, 2012)
Sheltered instruction observation protocol (Echevarria, Vogt, & Short, 2013)	ARCs: Motivational theory (Keller, 2010)

Table.1: Second Language Acquisition and Instructional Design Relations

Instruction through Presentation Software

With technological tools becoming very common in the classroom, teachers have been looking at new ways to provide classroom instruction and keep learners engaged (Hart, 2022). This includes the use of a blended approach to learning (i.e., the intersection of human interactions with technology-assisted learning situations) as benefits include a) insights into students' learning styles, b) ongoing customization to meet learners' needs, c) use of technology to boost learning, d) increased learner engagement, and e) support for students who are absent from class (Fulton, 2012, as cited by Hart, 2022). Presentation software such as Microsoft PowerPoint and Google Slides are easy-to-use tools that have become the most popular method to provide such an approach. Szaboa and Hastings (2000) suggest that the use of PowerPoint can help learners enhance attention and reduce distraction, which is a belief supported by Catherina (2006) and Wanner (2015), whose research suggests that PowerPoint presentations are more engaging for learners than traditional lectures. This was the result of research conducted by Wanner, which tested the effectiveness of PowerPoint presentations on instructing university students. The research concluded that the presentations improved comprehension of the course content and provided content more efficiently (Wanner, 2015). Similar findings were recorded in Oommen's (2012) study that found that out of a class of 50 university students, 94% of them responded positively toward the use of PowerPoint, saying it was easy to follow, stimulated thinking, helped make better use of class time, and held their attention.

A key benefit of using presentation software for instruction is the ability to customize content to suit individual learners or learning environments. Information can be added or omitted in real time, providing just-in-time teaching. Feedback can be given visually and also synchronously, and the utilization of media such as videos, audio, and images can be made to provide instruction that is efficient, interactive, and visually pleasing. The presentations appeal to varying learning styles, such as the use of visual, auditory, kinesthetic, and creative instruction (Hart, 2022). Cashman and Shelly's (2002) research found that students learn most effectively when using their five senses, highlighting the effectiveness of using mixed media during instruction. The use of visuals to support instruction has been extensively researched, reaching conclusions such as retention being increased by as much as 80 percent (Burrow, 1986) and the belief that visual aids stimulate thinking, improve the learning environment, increase personal understanding, provide more relevant course content, and promote more consistent performance (Mohanty, 2001; Rather, 2004; Kunari, 2006). In addition, presentation software is continuously advancing, with new functions being added to provide new methods of

information presentation. One example is the use of timed animations to control when information is presented, allowing them to not only be a communication aid but also simulate the timed-instruction of a teacher (Levy, 1997). Now, presentation software can act as a surrogate teacher or manager of tasks (Hart, 2022). However, Chiquito, Meskill, and Renjilian-Burgy (1997) highlight the difficulties of making decisions between aesthetics and functionality. Instruction cannot only be about conveying information to the learner, but must also include a constructivist approach that helps learners build their own knowledge through purposeful learning.

With the learner-centered paradigm being an important part of instructional design theory, technological tools can be used to accommodate individual learner needs by providing personlized instruction through customization, and presentation software like PowerPoint provides teachers with the ability to do so.

Research Study

Aims of Study

The aim of this study is to a) determine whether existing instructional design models can be used to design, develop, and implement more effective, personalized, and efficient instruction using presentation software and b) draw conclusions regarding whether instructional design methods should replace traditional teaching methods such as teacher-fronted instruction (i.e., lecture style) or the use of coursebooks/textbooks.

Theoretical Framework

A blended approach to ADDIE instructional design model was chosen for the study, however, Reigeluth and An's (2021) holistic 4D model was also used during the design stage of the process. This allowed for a three-tier design system to be used before the initial development of the presentation slides. In addition, the 4D model allowed for evaluation to take place during the development stage, meaning the presentation slide design could be improved after being tested on the learners. Evaluation was carried out through observations by the instructor and feedback from the learners.

Other theoretical frameworks included the beliefs put forward by Levy and Stockwell (2006) and Fulton (2012) that include the need for the use of multiple types of media (e.g., videos, images, animations) and continuous customization to adapt to the needs of the learners.

Based on the success of my previous study (Hart, 2022), a "flipped approach" was considered. A flipped classroom allows learners to review the content of the course prior to the class session and complete exercises that would usually be conducted together in a face-to-face environment (Bishop & Vergleger, 2013). While computer-mediated communication (CMC) software could have been used to present the slide presentations to the learners prior to each class for asynchronous learning (Liu & Chen, 2007), it would have been impossible to monitor learner input in real-time. Therefore, while the slides were provided before and after each class, instruction was still provided using the slides during each lesson.

Test Subjects

The subjects of this study were 216 English language learners enrolled at Rikkyo University. The students were spread over 11 classes and 2 different subjects: academic debate class (93 students) and presentation class (123 students). The students' English language skills varied between classes, with learners being leveled based on standardized testing prior to the start of the courses. In addition, 10 EFL lecturers who taught the same classes completed an online survey with 5 of them taking part in a pre-study interview.

Research Methods

For this study, a needs analysis (i.e., a study of the language learning and teaching needs of students in a language program) was conducted using mixed methods research. Johnson, Onwuegbuzie & Turner (2007) define mixed methods research as "an intellectual and practical synthesis based on qualitative and quantitative research" that "recognizes the importance of traditional quantitative and qualitative research but also offers a power third paradigm choice that often will provide the most informative, complete, balanced and useful research results" (p. 129). They explain that "the research should strategically combine qualitative and quantitative methods, approaches, and concepts in a way that produces complementary strengths and nonoverlapping weaknesses" and "generates research questions and provides answers to those questions, as appropriate" (p. 127). A mixed method was chosen to provide more accurate data for both the design and evaluation phases of the instructional design and to answer the questions stated in the study's aims. In addition, Brown (2014) states that "any researcher that can do both quantitative and qualitative research in TESOL will have considerable advantages over those who can only do one or the other" (p. 6).

Research Design

Participants

A total number of 223 (n = 223) participants were involved in the study. This included 10 university teachers (n = 10) who had taught the same subjects (i.e., English debate and presentation) as the researcher, 93 first-grade university students (n = 93) who were enrolled in the researcher's English presentation classes, and 120 first-grade university students (n = 120) who were enrolled in the researcher's English debate classes. All the students were in their second semester and had taken mandatory English language classes in the previous semester.

Procedures

Pre-study research was conducting via an online questionnaire that was given to the university teachers (n = 10). Also, half of the teachers (n = 5) were interviewed. After the completing the study, a post-study questionnaire was given to the 213 university students (n = 213) who had participated in the study. In addition, one of the students (n = 1) was interviewed about his experience with the researcher's in-class instruction.

Pre-Study Research Questions

To *analyze* (Molenda, 2003) and *define* (Reigeluth & An, 2021) the need for different forms of instruction, a pre-study survey was conducted. A questionnaire was given to 10 university EFL teachers (n = 10) that belonged to the same department as the researcher. All 10 teachers had taught the university courses before in the previous fall semester; therefore, they had sufficient insight into the courses, the learning/teaching content, and possible student behaviors. The questionnaire included 23 question items (Appendix 1.1). Questions 1–8 focused on the teachers' experience with using presentation slides. Questions 9–17 focused on the teachers' perceptions of the use of slide presentations. Questions 18–22 asked about the impact of slide presentations on learners. At the end of the survey, the participants were given the opportunity to add additional comments (Item. 23).

In addition to the survey, five teachers (n = 5) from the survey study were interviewed. Questions were asked about their instructional methods and teaching styles, use of technology in language classrooms, and their thoughts on a learner-focused approach. The interviewees' responses were recorded and notes were taken by the interviewer.

Research Design, Development, and Evaluation

In the design phase, the learning objectives, lesson planning, media selection, and a "fuzzy vision" of the instructional system (top-level design) were established. Based on the data collected from the analysis phase (i.e., pre-research survey and interviews and the researcher's previous experience teaching the courses), PowerPoint presentations were designed and developed using a mixed media approach and existing course content from the assigned coursebooks. The focus of the study was on providing effective instruction and support, rather than the teaching of content such as target language and skills, as this was done through active learning and task-based learning, and the use of existing learning materials that were incorporated into the presentations. Instructional presentations were created for the first two lessons for initial implementation and evaluation. Evaluation included a) the reaction of the learners, b) the learners' resulting learning and any noticeable increase in knowledge from the new instruction, c) the learners' behavioral change, and d) the effects on performance during in-class exercises. As relevant visual media was required, videos were recorded by the researcher and other teachers from the same faculty.

After Lesson 1 (course introduction) and Lesson 2 (content-based lesson), an analysis of the learners, materials, and context was conducted. Observations by the researcher were made, and a group of learners were asked to give their opinions on the instructional presentations. This analysis and evaluation led to decisions on both what and how to design and develop the next set of instructional slides (mid-level design).

For the next set of instructional presentation slides, changes and improvements were made (e.g., designing more personalized presentations based on the learners' interests and lesson topics, reduced reading, the use of timed animations to present information more effectively and to reduce on-screen text, and the use of more attractive slide templates to make them more visually appealing). The mid-level design was continuously evaluated through observations and by monitoring the learners' behaviors and performance.

After an evaluation of the mid-level design, minor changes were made (e.g., allowing learners to add information to the PowerPoint slides, making them more personal and interactive) leading to the final lower-level design. This design was used to provide instruction for the remainder of the course.

All the instructional PowerPoint slides were uploaded on the university's CMC software called Blackboard. Therefore, the learners had access to the slides in advance and after the lessons for reflection and revision. The slides were also shared on the faculty's shared Google Drive, allowing other teachers to use them in their classes.

Post-Study Research Questions

After completing the courses, a survey was conducted with a focus on learner perceptions (Appendix 2.1 & 2.2). An online questionnaire was given to the learners (n = 213) during the final two weeks of the course. A simple 5-point Likert scale was used for most of the questions, ranging from *Strongly Agree* to *Strongly Disagree*. A 5-point Likert scale was chosen as it is easy to draw conclusions, reports, results, and graphs and make comparisons from the responses, and it provides a construct-centered approach to collecting relevant data (Messik, 1989). In addition, four or five points are desirable for young learners or learners with low motivation to complete the questionnaire because 5-point scales are easy to understand, and they require less effort to answer (Smith, Wakely, DeKruif, & Swartz, 2003). While some of the question items were the same between the two different courses, there were also additional questions that were relevant to the individual courses.

The online questionnaire for the presentation class contained 27 question items (Appendix 2.1). Questions 1–10 (Part 1) provided and evaluation of the researcher's instructional slides. Questions 11–18 (Part 2) provided an evaluation of the learners' own slides. Questions 19–21 (Part 3) focused on course content and 22–27 (Part 4) provided the opportunity for the learners to express any final thoughts through written comments.

The online questionnaire for the debate course contained 19 question items (Appendix 2.2). Questions 1–2 (Part 1) asked the learners about preferred teaching styles. Questions 3–12 (Part 2) asked the learners to evaluate the researcher's instructional slides. In Part 3, questions 13–15 asked about course content, and Part 4 (questions 16–19) allowed for more detailed explanations to be given through final comments.

Issues Encountered

Most of the planned data collection was completed successfully; however, some issues caused by the COVID-19 pandemic affected data collection:

1. The first two weeks of the courses were conducted online due to restrictions implemented by the university. This made it more difficult to observe the learners in a regular classroom environment.
2. Many of the subjects were absent from the classes during the last two weeks of the study because of COVID-19-related situations. Out of the eight students who had agreed to be interviewed, only one was able to attend class.

Results

Pre-study Interviews: Teachers

During the teacher interviews (n = 5), notes were recorded regarding effective methods of providing instruction, different approaches to content presentation, and the use of technology in a

language classroom. It was interesting to hear that some of the teachers had reverted to using only the coursebooks after returning to face-to-face classes having taught online due to the recent pandemic. One teacher explained this was mainly due to convenience, where he did not have to carry a computer or set up a projector in his classroom. Another teacher explained how he pasted screenshots of the coursebooks content and instructions onto PowerPoint slides to help him remember the "flow of the lesson." In most cases, the learners were not considered, and choices were made for the convenience of the instructor or because of time restraints.

Most teachers used online learning platforms such as Google Classroom or Blackboard to administer tests or homework, with only one of the teachers using these platforms to provide instruction using a "flipped classroom" approach. His reasoning for this was that he believed students felt "less pressure" when working asynchronously, and it allowed them to "study at their own pace." This belief has been supported by other researchers whose research has shown that CMC provides learners with a safe environment to practice what they have learned and evaluate themselves (Fitze, 2006; Satar & Özdener, 2008).

In one interview, issues were highlighted regarding the students' lack of understanding during online classes. When instruction was given verbally, his students would sometimes go to breakout rooms and not notify the teacher of their lack of understanding. This meant that the students would wait in the breakout rooms in silence. He felt that PowerPoint slides would have helped provide clearer instruction, as he usually used a whiteboard in face-to-face classes. In another interview, similar points were made, as the teacher explained how he used the internet (e.g., YouTube, information websites, online articles) to provide examples, but admitted that time was sometimes wasted moving between sources.

In all the interviews, the teachers answered that they thought their students preferred classes with no textbooks or instructional materials. However, this belief was challenged in the learner questionnaire where 70.7% of students preferred classes with a mix of both "lecture style" and student-led "active learning."

Question	Answer	N	Question	Answer	N
1. Do you create slide presentations?	Always	4	13. I am able to deliver material easily by using slide presentations.	Strongly Agree	5
	Often	3		Agree	4
	Sometimes	3		Neutral	1
	Seldom	0		Disagree	0
	Never	0		Strongly Disagree	0
2. Do you use existing templates when creating slide presentations?	Always	1	14. I think students prefer the use of slide presentations over conventional methods.	Strongly Agree	1
	Often	2		Agree	3
	Sometimes	3		Neutral	5
	Seldom	3		Disagree	1
	Never	1		Strongly Disagree	0
3. Do you use pictures in your slide presentations?	Always	4	15. I think the use of media (e.g., pictures, videos, audio) helps students retain information.	Strongly Agree	4
	Often	2		Agree	6
	Sometimes	4		Neutral	0
	Seldom	0		Disagree	0
	Never	0		Strongly Disagree	0
4. Do you use videos in your slide presentations?	Always	0	16. I find it easy to create interesting or engaging slide presentations.	Strongly Agree	1
	Often	1		Agree	6
	Sometimes	3		Neutral	2
	Seldom	3		Disagree	1
	Never	3		Strongly Disagree	0

Question	Answer	N	Question	Answer	N
5. Do you use audio (e.g., music, sound clips) in your slide presentations?	Always	0	17. The facilities at the university support the use of slide presentations.	Strongly Agree	3
	Often	0		Agree	6
	Sometimes	3		Neutral	0
	Seldom	1		Disagree	0
	Never	6		Strongly Disagree	1
6. Do you use hyperlinks in your slide presentations?	Always	0	18. I put emphasis on the importance of slides when teaching presentation skills.	Strongly Agree	4
	Often	3		Agree	6
	Sometimes	3		Neutral	0
	Seldom	3		Disagree	0
	Never	1		Strongly Disagree	0
7. Do you usually use the same design for your slide presentations?	Always	2	19. I feel the use of slide presentations helps support the learners' speaking during their presentations.	Strongly Agree	4
	Often	4		Agree	5
	Sometimes	3		Neutral	1
	Seldom	1		Disagree	0
	Never	0		Strongly Disagree	0
8. Do you design your slide presentations based on the students' interests?	Always	0	20. I feel the use of slide presentations helps reduce learner anxiety during their presentations.	Strongly Agree	2
	Often	0		Agree	6
	Sometimes	5		Neutral	2
	Seldom	2		Disagree	0
	Never	3		Strongly Disagree	0
9. Slide presentations engage students in learning.	Strongly Agree	1	21. I feel good slides or visuals are important for a strong presentation.	Strongly Agree	5
	Agree	5		Agree	5
	Neutral	4		Neutral	0
	Disagree	0		Disagree	0
	Strongly Disagree	0		Strongly Disagree	0
10. Slide presentations help present learning materials more clearly.	Strongly Agree	7	22. The use of slides or visuals has an influence on how I grade my students.	Strongly Agree	2
	Agree	3		Agree	5
	Neutral	0		Neutral	1
	Disagree	0		Disagree	1
	Strongly Disagree	0		Strongly Disagree	1
11. Slide presentations provide inspiration to the learners when making their own presentations.	Strongly Agree	2	23. If you have any additional comments or observations regarding the use of slide presentations, please write them below.	N/A	N/A
	Agree	3			
	Neutral	5			
	Disagree	0			
	Strongly Disagree	0			
12. I prefer using slide presentations over conventional methods.	Strongly Agree	5			
	Agree	0			
	Neutral	4			
	Disagree	0			
	Strongly Disagree	1			

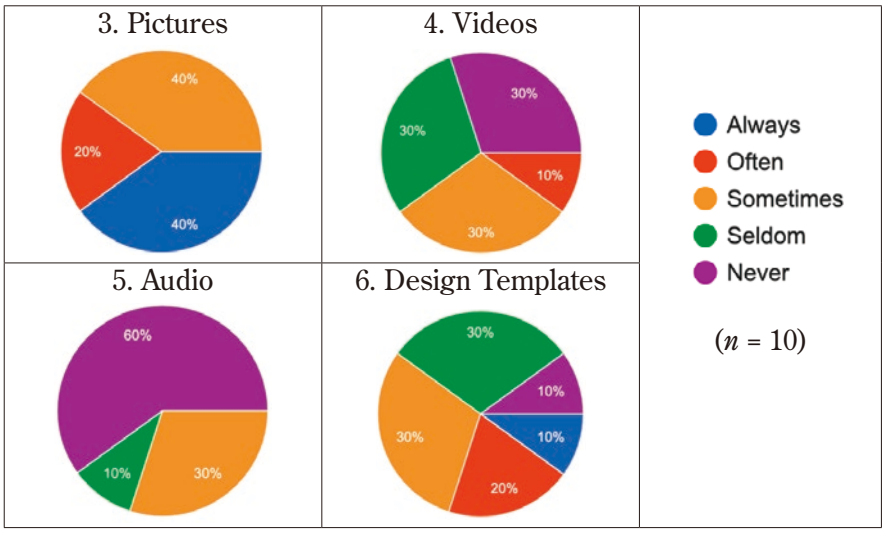
Table.2: Using Slide Presentations – Teacher Perceptions (n = 10)

Teacher Perceptions vs. Learner Perceptions

Almost all the teachers (n = 10) answered that they use slide presentations during their classes, with 40% always doing so. However, it was noted by a few of the participants that this was mainly due to their classes being moved back and forth from face-to-face to online. One of them stated that slides were "extremely important when teaching online," but "not so much in face-to-face classes" as he preferred using a whiteboard. Another participant said that he only used slides for "giving feedback" or "giving examples," with little use for them otherwise. One teacher saw slides as a distraction, while another felt that he lacked the skills to create effective PowerPoint presentations. Furthermore, 40% of the teachers preferred conventional methods of instruction (i.e., using a textbook and whiteboard),

which contrasts with a total of 86% of the students believing that the slides from the study were more useful than the textbook. Also, 50% of the teachers were neutral when asked if they thought students preferred slides, with only one teacher strongly agreeing that they do. When the students were asked if they preferred the use of slide presentations over conventional methods, 98.2% agreed that they did. In the debate course, a student wrote that "the flow of the debate" was "easier to understand than the textbook." During the student interview (n = 1), a learner was asked about his need for a textbook, and he replied by saying that he preferred the use of slides and felt that he had no need for a textbook in the class. However, while the slides were available online, the student asked for a digital PDF to be used for after-class review. This belief was not held by all students, with one student writing "I think using a textbook is important" and that they appreciated that page references had been added to the slides. The decision to do this was mainly because it was compulsory for all students to buy assigned textbooks from the university.

Data from the study supported Cashman and Shelly's (2002) belief that the use of mixed media promotes more effective learning as 90.2% of the students between the two courses agreed that the use of various media helped them understand the information more clearly. The students mentioned that the use of media helped them be more attentive and made the slides "interesting and fun to look at". This provides added support to the belief that PowerPoint slides help learners focus attention and reduce distraction (Szaboa & Hastings, 2000; Catherina, 2006; Wanner, 2015). Various students highlighted the use of animations and videos that made instruction "not boring" and "very clear than textbook," with one student saying that the slides "helped us understand what the teacher said" and made it "easy for us to understood important things." In the written comments from both questionnaires, when asked what the learners thought about the instructional slides, many of the comments put emphasis on the slide "design," with references to the "photos," "pictures," "videos," and "animations." Descriptive words such as "dynamic," "motivational," "fun" and "easy-understand" were used.



Figures.3-6: Teachers' use of mixed media

In contrast to the learners' perceptions, the use of mixed media by the teachers was not consistent (*Figures.3-6*): 30% of the teachers never used videos, 60% never used audio, and only 30% tended to use attractive design templates. The use of pictures was the only type of media that was consistently used by the teachers. It could be argued that the instructor themselves provide the

auditory and kinesthetic aspects of instruction; however, these aspects can be more varied and made more creative and appealing through the use of presentation slides.

A clear difference that was noticed between the teachers' slides and the ones used in this study is the focus on the learners' interests or the lesson topics: 50% of the teachers seldom or never consider the interests of their learners when designing slide presentations, with the other 50% only sometimes doing so. During this study's design process, changes were made after the top-tier design to utilize the learners' interests to make them more learner-focused and relatable. The learners' interests were discussed, shared, and listed in the Lesson 1 slides. Throughout the rest of the course, related media was used in the slides to provide instruction, demonstrations, or examples. In addition, to provide inspiration, the slides were often designed around the lesson topic or current events. By observing the students, and by evaluating instructional design in the initial phases, it became clear that the students responded better to slides that contained media that they could relate to. Examples included characters, famous people, places, or objects that they were familiar with. Visual aids also provided ideas that the students could draw from, reducing their cognitive load and allowing them to focus on the skills being taught. As suggested by Reigeluth (2014), the flexibilty and customabilty of technology provides the instructor with the ability to accommodate learners' individual needs and expectations and provide relevant and personalized learning experiences can be provided to the learners.

Question	Answer	N	Question	Answer	N
1. The teacher's slides were interesting.	Strongly Agree Agree Neutral Disagree Strongly Disagree	72 21 0 0 0	15. Using slides helped me feel more relaxed during my presentations.	Strongly Agree Agree Neutral Disagree Strongly Disagree	33 42 17 1 0
2. The teacher's slides were easy to follow and understand.	Strongly Agree Agree Neutral Disagree Strongly Disagree	75 12 6 0 0	16. Using slides supported my speaking when giving presentations.	Strongly Agree Agree Neutral Disagree Strongly Disagree	48 36 5 0 0
3. The teacher's slides made it easier to understand the teacher's instructions and lesson goals.	Strongly Agree Agree Neutral Disagree Strongly Disagree	75 18 0 0 0	17. I used ideas from my teacher's slides in my presentations.	Strongly Agree Agree Neutral Disagree Strongly Disagree	43 34 15 2 0
4. The teacher's slides held my attention throughout the class.	Strongly Agree Agree Neutral Disagree Strongly Disagree	57 33 3 0 0	18. I think using slides is important to give a strong presentation.	Strongly Agree Agree Neutral Disagree Strongly Disagree	60 21 9 0 0
5. The teacher's slides helped me remember information more easily.	Strongly Agree Agree Neutral Disagree Strongly Disagree	69 24 0 0 0	19. I prefer to use slides during a presentation instead of only speaking.	Strongly Agree Agree Neutral Disagree Strongly Disagree	52 27 12 2 0
6. The teacher's slides helped manage class time.	Strongly Agree Agree Neutral Disagree Strongly Disagree	60 27 6 0 0	20. The teacher's slides on Blackboard helped me prepare before each class.	Strongly Agree Agree Neutral Disagree Strongly Disagree	54 27 6 3 0

Question	Answer	N	Question	Answer	N
7. The teacher's slides were attractive and fun to look at.	Strongly Agree Agree Neutral Disagree Strongly Disagree	81 9 3 0 0	21. The teacher's slides on Blackboard helped me review and better understand the lesson after class.	Strongly Agree Agree Neutral Disagree Strongly Disagree	48 30 5 0 0
8. The use of media (e.g., pictures, videos, music, animations) helped present information more clearly.	Strongly Agree Agree Neutral Disagree Strongly Disagree	66 18 9 0 0	22. The teacher's slides were more useful than using a textbook.	Strongly Agree Agree Neutral Disagree Strongly Disagree	60 28 2 0 0
9. The teacher's slides motivated me to make my own slides better.	Strongly Agree Agree Neutral Disagree Strongly Disagree	54 30 9 0 0	23. What did you think about the teacher's slides?	N/A	N/A
10. I prefer lessons with slide presentations.	Strongly Agree Agree Neutral Disagree Strongly Disagree	63 26 1 0 0	24. What did you think of your slides?	N/A	N/A
11. I found it easy making interesting or attractive slides.	Strongly Agree Agree Neutral Disagree Strongly Disagree	24 39 24 2 24	25. Do your other teachers use presentation slides?	Yes, always. Sometimes. No, never.	45 42 6
12. I enjoyed making slides for my presentations.	Strongly Agree Agree Neutral Disagree Strongly Disagree	39 41 3 0 0	26. If yes, how are their slides similar/different?	N/A	N/A
13. Using slides made my presentations easier to understand.	Strongly Agree Agree Neutral Disagree Strongly Disagree	43 32 17 1 0	27. If no or sometimes, how do your other teachers present the course content?	N/A	N/A
14. Using slides helped me give my presentation more smoothly.	Strongly Agree Agree Neutral Disagree Strongly Disagree	49 36 6 2 0			

Table.3: Using Slide Presentations – Learners Responses (Presentation Students, n = 93)

When asked if their other teachers used presentation slides during their classes, 52.6% of the learners (n = 213) said "Yes, always," with the remaining students choosing "Sometimes." It is important to note that the other teachers whom they refered to were not only their language teachers but also teachers in their regular classes. However, comparisons can be made between different approaches to technology-based instructional design, and conclusions can be drawn from the participants' responses. Below are some examples:

a) *Ian's slides are more easy and interesting to look at.*

b) Other teachers just list words, but the slides in this class are animated and interesting.

c) The other teachers' slides don't have many illustrations, they are very mechanical.

d) Other teachers' slides are more descriptive and hard to read.
e) Other teachers' slides just give information so they are too simple. My debate teacher's slides have not only important information but also attractive motion and stuff.
f) To begin, the slides used introduced a debate battle. Next, it showed the schedule of debate. It was very convenient.
g) Other teachers' slides are not interesting and it is difficult to understand because they don't use pictures effectively.
h) Words are not so many as your slides.
i) Other teachers' slides are much simpler. Not as motivating as Ian's.
j) Other teachers' slides are only words and sentence. No pictures.
k) Similar, but it is more fun than others'.

The comparisons show that in many cases, the learners felt that the PowerPoint presentations designed for this study were more "interesting," "motivating," and "easier to follow," Also, the fun factor of the slides through the use of animations and pictures seemed to appeal to the learners. In statement *f*, the student highlights the convenience of the slides in providing a form of scaffolding, with videos showing the end goal, so that the students knew what they were aiming toward. Then, each phase of production could be broken down effectively, giving the learners a clearer understanding. While coursebooks/textbooks can provide similar instruction and scaffolding, the nature of this particular class (debate) means that students benefit more from witnessing the completed task in its natural form (i.e., seeing a full English debate in a video). In addition, the use of PowerPoint animations allows the teacher to present instruction and content when relevant. This minimizes the amount of information visible to the learner at any given time. When asked if they find it easy to create interesting and engaging slide presentations, 70% of the teachers felt that they did, with the others disagreeing or remaining neutral. Without testing their slides on the same students, it is impossible to observe differences between the slide designs; however, it is clear from the student responses that they notice differences in instruction and they are able to evaluate which method they find more effective.

Regarding learner performance, positive results were seen by both the researcher and learners. As the slide presentations were available before and after each class, nearly all of the students felt that the slides had helped them prepare in advance and better understand the lesson after the class. In the presentation class, more students agreed that they used the slides for preparation (90%) than review (83.8%), while in the debate class it was the reverse (57.5% for preparation, 70% for review). This may be explained by the course content, with the presentation class requiring more preparation outside of class time, requiring a wider varietyof language and skills to be learned, while the debate course is more task-based learning, with less focus on taught content and more focus on in-class planning and production. In these cases, the instructional slides acted as a "surrogate teacher," as seen in previous studies (Levy, 1997; Hart, 2022), with information being presented using a set order and effective timing, much like the verbal instruction of a teacher. By viewing the slides prior to the class, a kind of flipped classroom was created, with in-class instruction being more of a review of what was previously taught. As for the students who used the slides to review the class content, an interesting comment was given by one of the students who had said their other teachers do not use presentation slides: "I think PowerPoint is better because I know what what each lesson's goal is. Also, if using PowerPoint, I can use them to review everytime." The implication here is that the slides provide more structured presentation and help present learning goals more clearly. Moreover,

technology-based materials are more apealing to young learners, which improves motivation and engagement.

Question	Answer	N	Question	Answer	N
1. Which type of class do you prefer?	Lecture Style Active Learning Mixed	8 30 82	11. The teacher's slides motivated me to use the skills that were taught.	Strongly Agree Agree Neutral Disagree Strongly Disagree	45 42 33 0 0
2. Which teaching style do you other university teachers usually use?	Lecture Style Active Learning Mixed	69 9 39	12. I prefer lessons with slide presentations.	Strongly Agree Agree Neutral Disagree Strongly Disagree	66 48 3 0 0
3. The teacher's slides were interesting.	Strongly Agree Agree Neutral Disagree Strongly Disagree	87 27 3 3 0	13. The teacher's slides on Blackboard helped me prepare before each lesson.	Strongly Agree Agree Neutral Disagree Strongly Disagree	12 57 45 4 1
4. The teacher's slides were easy to follow.	Strongly Agree Agree Neutral Disagree Strongly Disagree	66 42 9 0 0	14. The teacher's slides on Blackboard helped me review and better understand the lesson after the class.	Strongly Agree Agree Neutral Disagree Strongly Disagree	15 69 30 6 0
5. The teacher's slides made it easier to understand the debate skills and lesson goals.	Strongly Agree Agree Neutral Disagree Strongly Disagree	60 51 9 0 0	15. The teacher's slides were more useful than using a textbook.	Strongly Agree Agree Neutral Disagree Strongly Disagree	48 39 24 6 0
6. The teacher's slides held my attention throughout the class.	Strongly Agree Agree Neutral Disagree Strongly Disagree	54 54 12 0 0	16. What did you think about the teacher's slides?	N/A	N/A
7. The teacher's slides helped me remember information more easily.	Strongly Agree Agree Neutral Disagree Strongly Disagree	60 36 21 3 0	17. Do your other teachers use slides in their classes?	Yes, always Sometimes No, never	69 51 0
8. The teacher's slides helped in managing class time efficiently.	Strongly Agree Agree Neutral Disagree Strongly Disagree	72 48 0 0 0	18. If yes, are there any differences with their slides? How do they use slides differently?	N/A	N/A
9. The teacher's slides were attractive and fun to look at.	Strongly Agree Agree Neutral Disagree Strongly Disagree	75 30 15 0 0	19. If no or sometimes, how do your other teachers present the course content?	N/A	N/A
10. The use of media (e.g., pictures, videos, music, animations) helped present information more clearly.	Strongly Agree Agree Neutral Disagree Strongly Disagree	72 36 9 3 0			

Table.4: Using Slide Presentations – Learners Responses (Debate Students, n = 120)

Another positive impact the instructional slides had on performance was that they provided inspiration to the students in the presentation class: 98.9% of the students admitted to using ideas from the teacher's slides when creating their own PowerPoint presentations. Clear differences were observed compared with the previous year, with slide presentations being more interactive, dynamic, and aesthetically appealing. Students used a vareity of media, and animations and transitions were used to present information and media more smoothly and effectively. As an observer, I felt that I was able to understand the taught information more easily, and the presentation of information was more memorable. Some of the students expressed that they found it difficult to make effective slides, whereas positive comments were given such as "using slides helped the listeners to understand my presentation when they couldn't listen to my voice." Even though the students did not use an instructional design model when designing their presentations, they appeared to be influenced by the teacher's presentations, resulting in clearer presentation and instruction.

In the debate class, major improvements were achieved with regard to learner progress. In the coursebook, debate skills are explained to the students in each lesson, with the first full debate taking place in lessons 8 and 9 (i.e., week 8 and 9). When evaluating the *top level* of the instructional design model, I noticed that the learners responded well to demonstration videos, and that the use of diagrams and animated examples made the skill-based goals more digestible for the students. This led to the combination of skills to show connectivity between them and displayed the structured progression of a debate more clearly. This also provided the students with examples of natural language use and fluent speaking. As a result, the students were able to conduct a full debate successfully by lesson 4. This created more time for practice, production, and testing. After the *mid-level* evaluation, slides were created to be more interactive, with students having access to the instructional material to add their own debate topics and images. I felt that the students appreciated having their ideas displayed in the learning materials, and with the lessons being more personalized, gains in learner participation, engagement, and motivation were observed.

During the length of the study, the presentation slides were shared with other instructors in my department who taught the same classes. Feedback was given by some of the teachers who had used the slides, with all of the responses being positive. One teacher explained that his students preferred the slides over using the coursebook, causing him to stop using the coursebook entirely. Benefits included increased participation, clearer instruction, more efficient use of time, the ability to provide instruction online, and more professional looking classes.

Implications, Limitations, and Conclusion

Implications

The study was relatively small in scale but clear implications can be made from the data gathered that support the use of IDT when designing technology-based instructional materials. Overall, the majority of the learners showed signs of preferring the instructional slides over conventional methods such as teacher-fronted verbal instruction or the use of the supplied coursebooks. The learners noticed clear differences in the presentation slides used in this study compared with slides used in their other classes, which proves that the learners are aware of the teaching methods of their instructors. Further evidence for this was provided by the end of course evaluation conducted by the university. When asked "Is there anything that you thought good about this course?," the answer "PowerPoint" was the most popular choice in all 11 classes (*Figure.7*).

Is there anything that you thought good about this course? [Multiple answers allowed]

Handouts (class resumes, etc.) / Handouts (Worksheets, including digital resources etc.) (1)
Write on the blackboard (Written communication in class, including use of digital whoteboards.) (2)
PowerPoint (12)
Video and other visual aids (This is not a video of the online class itself.) (0)
Syllabus (1)
None of the above apply / NA not applicable (0)

Figure.7: Example of course evaluation questionnaire answers

The learner-centered paradigm that is at the core of IDT allows for the learners to be more active participants, with many of the learners highlighting that the presentation slides were motivating and immersive, providing clearer and more efficent direction. Evidence for this was seen in the all subjects, with course goals being achieved sooner and at a higher standard. As stated by Reigeluth (2014), using such a personalized educated system "empowers learners" and supports their "self-directed learning" (p. 223).

The need for a holistic approach to evaluate instruction was also supported during the study. While it is important to identify instructional goals and conduct a needs analysis early on, continuous evaluation during the design and development helped create more effective outcome-based objectives. Further improvements to instruction could have been made to the instructional materials if more in-depth quantitative and qualitative data had been collected during the study (e.g., a number of students asked for even further reduction of text on screen at the end of the study). This implies that Reigeluth and An's (2021) holistic 4D model is better suited to the design of instructional presentation slides than the stardard ADDIE approach, as it provides more in-depth continuous evaluation throughout the design and development process.

Based on the responses made by the learners, and through observations made by other teachers and myself while using the newly design presentation slides, improved learner motivation was apparent. The use of an instructional design model helped cultivate students' motivation by focusing not only on their goals but also by utilizing their values and interests to provide more personlized and appealing instruction. This echos aspects of early motivational design models, such as Keller's model (1983) that identifies four major aspects of motivation in educational instruction: interest, relevance, expectancy, and satisfaction. Due to the university courses being mandatory, and having around 20 students per class, providing personalized learning materials for individuals students is not usually possible, other than the materials being level appropriate. However, instruction could be customized to provide learning experiences that accomodate learner interests, with optimal novelty, and a degree of learner control. Positive effects were present such as curiousity and increased engagement that appeared to enhance the learners' intrinsic motivation. Words such as "attractive," "interesting," "fun," and "motivating" were used by various students. With the use of mixed-media, tasks became more relevant by providing examples and inspiration, making them more meaningful. Without the use of control groups, the learners' willingness to participate in course-related tasks was difficult to measure; nonetheless, improvement in effort and production was evident in their presentations compared with the previous year. Data from the learner surveys also suggest improvement in learner confidence, with students stating that the classes were "easy to understand," making it easier for learners' to "understand the theme and goal" of each lesson. This feeling of confidence and success is known to facilitate learning (APA, 1993). Many of the students felt that the slide design influenced their own design choices when producing their own presentations. As for learner satisfaction, the

students were able to “understand deeply” about the skills that they were learning, and media and animated explanations provided opportunities to show them why the taught content was important and how they could use it effectively. This motivated the learners to pursue the same goals. By taking all these points into consideration, the study implies that IDT provided improved instruction outcomes with regard to effectiveness, motivation, efficiency, and appeal, as previously described by Reigeluth (2021).

Finally, with IDT being design oriented, the use of a technology (i.e., blended learning) allowed for a wider variety of instructional methods to be used in consideration of values, outcomes, and conditions. The transition between online and face-to-face classes was made smoother, as the instructional presentations could easily be customized to suit the conditions, and the study proved that the learners preferred the use of PowerPoint slides over coursebooks and the use of whiteboards. Additional instructional methods proved effective such as providing learners with a “flipped” classroom by providing animated slides prior to each lesson that presented information with timings that simulated a teacher’s instruction.

Limitations and Suggestions for Future Research

Due to unforeseen complications, only one student interview was conducted, and with insightful information coming from that interview, I believe the study would have benefited from more interviews from a wider variety of learners. I also believe that control groups would have provided a clearer understanding of changes in learner performance. As explained, a key aspect of IDT is the need for learner-centered instruction through a clear understanding of the learners’ values and personal goals. A wider variety of needs analysis and evaluation techniques could have been used during the design and development process, such as giving the learners the opportunity to provide continuous feedback during the courses (e.g., feedback forms, questionnaires, or interviews). Interviews with other teachers who used the presentation slides could also have been conducted to get a more accurate understanding of a wider range of learner needs.

Conclusion

This study provided support to the idea that an instructional design model can be used to design technology-assisted materials that provide learners with more personlized and meaningful instruction. By putting more focus on the individual needs, values, and personal interests of the learners, a shift can be made from standardized instructions to customized, where learner needs are more effectively met and learner motivation is improved. While plenty of existing research supports the use of active, student-centered learning, it is also important to provide student-centered instruction that is meaningful, efficient, and does not hold the time constraints that are found in traditional language classes. The difficulty is that if instruction should be personalized to individual learners, it is difficult to teach the same thing to a classroom of students, especially in standardized language courses. However, by using well-designed resources developed using instruction design models and instructional technology (e.g., presentation software), a more personalized experience can be achieved. Also, it is important that the instruction is focused not just on conveying information to the learner, but it also includes a constructivist approach as referred to by cognitive theorists as a process of helping learners build their own knowledge through the facilitation of purposeful learning (Ertmer & Newby, 2013).

Based on the evidence presented in this study, it is suggested that more attention should be given to design principles when creating instructional materials for L2 learners, and that further research needs to be carried out on the design and development of effective technology-assisted instructional presentations.

References

APA (1993). *Learner Centered Psychological Principles: Guidelines for School Redesign and Reform.* Washington, DC: American Psychological Association and the Mid-Continent Regional Educational Laboratory.

Bednar, A. K., Cunningham, D., Duffy, T. M., & Perry, J. D. (1991). Theory into practice: How do we link? . In G. J. Anglin, *Instructional Technology: Past, Present, and Future.* Englewood, CO: Libraries Unlimited.

Bishop, J. L., & Vergleger, M. A. (2013). The flipped classroom: A survey of the research. *ASEE National Conference Proceedings.* Atlanta, GA.

Brown, J. D. (2014). *Mixed Methods Research for TESOL.* Edinburgh: Edinburgh University Press Ltd.

Burrow, T. (1986). *Horizons in Human Geography.* London: Macmillan.

Caherina, F. (2006). Beyond presentation: Using PowerPoint as an effective instructional tool. *Gifted Child Today, 4*.

Cashman, T. J., & Shelly, G. B. (2002). *Microsoft PowerPoint 2002 Introductoy Concepts and Techniques.* Boston, Massachusetts: Cengage Learning, Inc.

Chiquito, A. B., Meskill, C., & Renjilian-Burgy, J. (1997). Multiple, mixed, and malleable media. In M. D. Bush, & R. M. Terry, *Technology-Enhanced Language Learning* (pp. 47-76). Chicago, Illinois: NTC/Contemporary Publishing Group, Inc.

Echevarria, J., Vogt, M.-E., & Short, D. J. (2013). *Making Content Comprehensible for English Learners: The SIOP Model.* North Carolina: Pearson Education, Inc.

Ertmer, P. A., & Newby, T. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 26*(2), 43-71.

Fitze, M. (2006). Discourse and participation in ESl face-to-face and written electronic conferences. *Learning Language & Technology, 10*(1), 67-86.

Fulton, K. (2012). Upside down and inside out: Flip your classroom to improve student learning. *Learning & Leading with Technology, 39*(8), 12-17.

Gustafson, K., & Branch, R. (2002). *What Is Instructional Design? Trends and Issues in Instructional Design and Technology.* Saddle River, NJ: Merrill/Prentice Hall.

Hall, T. E., Meyer, A., & Rose, D. H. (2012). *Universal Design for Learning in theC Classroom: Practical Applications.* New York: Guilford Press.

Hart, I. P. (2022). Designing a Technology-Enhanced Flipped Classroom Model Using Instructional Slideshows and Computer-Mediated Asynchronous Discussions. *Journal of Multilingual Pedagogy and Practice, 2.*

Huang, W.-H. D. (2013). *Online Learning Engagement System (OLES) Design Framework for Postsecondary Online Learning Environments: A Synthesis on Affordances from Game-Based Learning, Social Media-Enabled Learning, and Open Learning.* Illinois: University of Illinois at Urbana-Champaign.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods

research. *Journal of Mixed Methods Research, 1*(2), 112-133.

Keller, J. (1983). Motivational Design of Instruction. In C. Reigeluth, *Instructional-Design Theories and Models: An Overview of Their Current Status.* Mahwah, NJ: Lawrence.

Kumaravadivelu, B. (2003). A postmethod perspective on English language teaching. *World Englishes*, 539-550.

Kunari, C. (2006). *Methods of Teaching Educational Technology.* New Delhi: Darya Gaj.

Levy, M. (1997). *Computer-Assisted Language Learning.* New York: Oxford University Press.

Levy, M., & Stockwell, G. (2006). *CALL Dimensions: Optons and Issues in Computer-Assisted Language Learning.* New York: Lawrence Erlbaum Associates.

Liu, G., & Chen, A. S. (2007). A taxonomy of Internet-based technologies integrated in language curricula. *British Journal of Educational Technology, 38*(5), 934–938.

Messik, S. (1989). Validity. In R. L. Linn, *Educational measurement* (pp. 13-103). New York: Macmillan.

Mohanty, J. (2001). *Educational Technology.* New Delhi: Rajouri Garden.

Molenda, M. (2003). In search of the elusive ADDIE model. *Performance Improvement, 45*(5), 34-36.

Nunan, D. (2004). *Task-Based Language Teaching.* Cambridge: Cambridge University Press.

Oommen, A. (2012). Teaching English as a global language in smart classrooms with PowerPoint oresentation. *English Language Teaching, 5*(12), 54-61.

Osborne, R. J., & Wittrock, M. C. (1983). Learning Science: A generative process. *Science Education, 67*(4), 489-508.

Perkins, D. N. (1992). *Smart Schools: Better Learning and Thinking for Every Child.* New York, NY: The Free Press.

Patel, S. R., Magolies, P. J., Covell, N. H., Lipscomb, C. & Dixon, L. B. (2018). Using instructional design, analyze, design, develop, implement, and evaluate, to develop e-Learning modules to disseminate supported employment for community behavioral health treatment programs in New York State. *Frontiers in Public Health: Curriculum, Instruction, and Pedagogy, 6 (113),* 1-9.

Raman, K., & Yamat, H. (2014). Barriers Teachers Face in Integrating ICT During English Lessons: A Case Study. *The Malaysian Online Journal of Educational Technology, 2*(3), 11-19.

Rather, A. R. (2004). *Essentials Instrctional Technology.* New Delhi: Darya Gaj.

Reigeluth, C. M. (1999a). *Instructional Design Theories and Models (Vol.2).* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Reigeluth, C. M. (1999b). What is instructional-design theory and how is it changing? In *Instructional Design Theories and Models: A New Paradigm of Instructioanl Theory (Vol.2)* (pp. 5-30). Hillsdale, NJ: Lawrence Erlbaum.

Reigeluth, C. M. (2014). The learner-centered paradigm: Roles for technology. *Educational Technology, 54*(3), 18-21.

Reigeluth, C. M., & An, Y. (2021). *Merging the Instructional Design Process with Learner-Centered Theory.* New York, NY: Routledge.

Reiser, R. A., & Dempsey, J. A. (2002). *Trends and Issues in Instructional Design and Technology.* Upper Saddle River, NJ: Merrill/Prentice Hall.

Satar, H. M., & Özdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal, 92*(4), 595-613.

Smith, P. L., & Ragan, T. J. (1993). *Instructional Design.* New York: Macmillan.

Smith, E. V., Wakely, M. B., DeKruif, R. E., & Swartz, C. W. (2003). Optimizing rating scales for self-efficacy (and other) research. *Educational and Psychological Measurement, 63*(3), 369-391.

Szaboa, A., & Hastings, N. (2000). Using IT in the undergraduate classroom: Should we replace the

blackboard wth PowerPoint? *Computers & Education, 32.*

Turan, Z., & Akdag-Cimen, B. (2020). Flipped classroom in English language teaching. *Computer Assisted Language Learning, 33*(6), 590-606.

Vaughan, N. (2007). Perspectives on blended learning in higher education. *International Journal of E-Learning, 6*(1), 81-94.

Vygotsky, L. S. (1980). *Mind in Society: The Development of Higher Psychological Processes.* Cambridge, MA: Harvard University Press.

Wadsworth, L. S. (2004). *Piaget's Theory of Cognitive and Affective Development: Foundations of Constructivism.* New York: Longman.

Wanner, T. (2015). Enhancing student engagement and active learning through just-in-time teaching and the use of PowerPoint. *International Journal of Teaching and Learning in Higher Education*, 154-161.

Weimer, M. (2013). *Learner-centred Teaching: Five Key Changes to Practice* (Vol. 2). Somerset, NJ: John Wiley & Sons.

Wiggins, G., & McTighe, J. (2005). Understanding by design. *Applied Linguistics Journal, 19*(1), 140-142.

Appendix 1 – Questionnaire Items for Teachers (Pre-Study)

1.1. Using Slides Presentations (Teacher Perceptions)

1. Do you create slide presentations?
2. Do you use existing templates when creating slide presentations?
3. Do you use pictures in your slide presentations?
4. Do you use videos in your slide presentations?
5. Do you use audio (e.g., music, sound clips) in your slide presentations?
6. Do you use hyperlinks in your slide presentations?
7. Do you usually use the same design for your slide presentations?
8. Do you design your slide presentations based on the students' interests?
9. Slide presentations engage students in learning.
10. Slide presentations help present learning materials more clearly.
11. Slide presentations provide inspiration to the learners when making their own presentations.
12. I prefer using slide presentations over conventional methods.
13. I am able to deliver material easily by using slide presentations.
14. I think students prefer the use of slide presentations over conventional methods.
15. I think the use of media (e.g., pictures, videos, audio) helps students retain information.
16. I find it easy to create interesting or engaging slide presentations.
17. The facilities at the university support the use of slide presentations.
18. I put emphasis on the importance of slides when teaching presentation skills.
19. I feel the use of slide presentations helps support the learners' speaking during their presentations.
20. I feel the use of slide presentations helps reduce learner anxiety during their presentations.
21. I feel good slides or visuals are important for a strong presentation.
22. The use of slides or visuals has an influence on how I grade my students.
23. If you have any additional comments or observations regarding the use of slide presentations, please write them below.

Appendix 2 – Questionnaire Items for Students (Post-Study)

2.1. Using Slides Presentations (Learner Perceptions: Presentation Classes)

1. The teacher's slides were interesting.
2. The teacher's slides were easy to follow and understand.
3. The teacher's slides made it easier to understand the teacher's instructions and lesson goals.
4. The teacher's slides held my attention throughout the class.
5. The teacher's slides helped me remember information more easily.
6. The teacher's slides helped manage class time.
7. The teacher's slides were attractive and fun to look at.
8. The use of media (e.g., pictures, videos, music, animations) helped present information more clearly.
9. The teacher's slides motivated me to make my own slides better.
10. I prefer lessons with slide presentations.
11. I found it easy making interesting or attractive slides.
12. I enjoyed making slides for my presentations.
13. Using slides made my presentations easier to understand.
14. Using slides helped me give my presentation more smoothly.
15. Using slides helped me feel more relaxed during my presentations.
16. Using slides supported my speaking when giving presentations.
17. I used ideas from my teacher's slides in my presentations.
18. I think using slides is important to give a strong presentation.
19. I prefer to use slides during a presentation instead of only speaking.
20. The teacher's slides on Blackboard helped me prepare before each class.
21. The teacher's slides on Blackboard helped me review and better understand the lesson after class.
22. The teacher's slides were more useful than using a textbook.
23. What did you think about the teacher's slides? Examples:
24. What did you think of your slides? Examples:
25. Do your other teachers use presentation slides?
26. If yes, how are their slides similar/different? Examples:
27. If no or sometimes, how do your other teachers present the course content?

2.2. Using Slides Presentations (Learner Perceptions: Debate Classes)

1. Which type of class do you prefer?
2. Which teaching style do you other university teachers usually use?
3. The teacher's slides were interesting.
4. The teacher's slides were easy to follow.
5. The teacher's slides made it easier to understand the debate skills and lesson goals.
6. The teacher's slides held my attention throughout the class.
7. The teacher's slides helped me remember information more easily.
8. The teacher's slides helped in managing class time efficiently.
9. The teacher's slides were attractive and fun to look at.

10. The use of media (e.g., pictures, videos, music, animations) helped present information more clearly.
11. The teacher's slides motivated me to use the skills that were taught.
12. I prefer lessons with slide presentations.
13. The teacher's slides on Blackboard helped me prepare before each lesson.
14. The teacher's slides on Blackboard helped me review and better understand the lesson after the class.
15. The teacher's slides were more useful than using a textbook.
16. What did you think about the teacher's slides? Examples:
17. Do your other teachers use slides in their classes?
18. If yes, are there any differences with their slides? How do they use slides differently? Examples:
19. If no or sometimes, how do your other teachers present the course content?

【Research Article】

Towards Rhetorical Genre Studies Some Conceptual Implications and Practical Considerations in Teaching Writing

Jeffrey Mok

Abstract

Rhetorical genre studies have been seen as a critical shift in the theory of understanding writing — a contextualized and situated nature of genre. Educators have been seeking to realize this socio-rhetorical theory in the classrooms and have found it daunting. This study seeks to briefly discuss the key concept of social action and implications to pedagogy: cultivation of mindset, metacognition, and inquiry-based approach. This is followed by a description of an attempt to implement a rhetorical genre studies classroom in Rikkyo University, reflecting on its implications the implications. The study reveals some challenges faced in a rhetorical genre studies classroom: situatedness of writing scene and assessment of metacognition.

Keywords: *rhetorical genre studies, reading and writing, theory, course design, classroom practice*

Introduction

Pedagogical considerations have always been a challenge for situated approaches to writing such as rhetorical genre studies (RGS). Writing acts are increasingly seen as "complex social participatory performance, in which the writer asserts meaning, goals, actions, affiliations, and identities within a constantly changing, contingently organized social world, relying on shared texts and knowledge" (Bazerman, 2013, p. 11). This understanding of writing as a highly contextualized and situated act within an activity system is shifting the teaching of reading and writing away from a narrow understanding about writing. As we approximate closer in our understanding of writing and its writing process, the challenge is to design a curriculum that best reflects these complexities. Traditional teaching of writing has become increasingly unsatisfactory today where genres are merging and transforming in our postmodern world (Miller & Kelly, 2017). Students are reading visual essays and interactive e-textbooks while writing a vlog, infographics, or twitterature (Aciman & Rensin, 2009). The conventional way of teaching writing does not seem to take into account these new genres, including emerging ones. The progression from product and process writing to genre analysis is a testament of this dissatisfaction. In recent decades, the introduction of genre analysis and emphasis on rhetoric have in many ways attempted to address this dissatisfaction. Today, RGS has emerged as one having the brightest prospect, having "a profound impact on the study and teaching of writing" (Bawarshi & Reiff, 2010, p 3).

The department responsible for the teaching of reading and writing courses at Rikkyo University was revisiting the writing curriculum and saw the need to revise the curriculum to respond to the changing landscape of the writing scene. As part of the committee, I decided to trial RGS in one semester and this paper is a description of that attempt. RGS may be a relatively new concept in the teaching writing at Rikkyo University and its choice is deliberate because of its apt representation of the natural thinking processes of a reader reading and a writer writing, under authentic settings. A critical and matured reader would immediately feel the effect of a text, its purpose, and the context

in which it is set. Similarly, a seasoned writer, about to pen her thoughts, would immediately think about the audience, her relationship with the reader, the conventions that befits the context of the writing genre, and the format to set the text in. RGS, as a writing approach, seeks to engender these thinking processes into the minds of students. In order to appreciate RGS, the paper begins with a discussion on its key concepts, followed by three pedagogical implications emerging from its key thrust of social action of writing acts. I will also attempt to show how these implications are played out in classroom teaching. The paper will end with challenges that the author feels that teachers face when using the RGS approach to teach writing.

RGS as a concept in writing

The notion of RGS began with the landmark reworking of the concept of genre by Miller (1984) into the central thrust that texts are forms of *social actions* — they deliver an effect on the reader with a social purpose. A genre is "not just a pattern of forms or even a method of achieving our own ends. We learn, more importantly, what ends we may have…" (Miller, 1984, p. 165). Indeed, building on Campbell and Jamieson's (1979) understanding of genre as language actions "driven by an internal dynamic" (p. 21), this "dynamic" merges the message and linguistic features in response to specific social situations. This "internal dynamic" is the intended social action of the writer seeking to affect the "ends." The writer interrogates the scene (context and specific situation) and bring forth the social communicative act of writing (or utterance) in a form (genre) that delivers the social effect. It is not dissimilar to Bakhtin's "utterances" (1986, p. 60), which positioned the understanding of text and speech acts in a highly situated context of human activity that is individualized with none rendered the same or similar in any way. Although Bakhtin's (1986) work on genre focused communicative acts as meaningful only when understood in their specific context, it was Miller (1984) who highlighted "that a rhetorical sound definition of genre must be centered not on the substance or the form of discourse but on the action it is used to accomplish" (p. 151). The rhetoric is thus accentuated in the genre to deliver the intended social effect adequately and appropriately. Learning to write is no longer just putting grammatically correct sentences in paragraphs and mimicking the genre form, but effecting a social action through the text.

Social action

Social actions are the "ends we may have…" (Miller, 1984, p. 165) in the communicative act. In other words, it is the outcome or effect felt by the reader. When this effect is felt by the reader, then the writer would have successfully achieved the writing act. The writer seeks this effect from his act of writing — the writer's desired effect on the reader. But this social action is the result of "the social motives, relations, values, and assumptions embodied within a genre that frame how, why, and when to act" (Bawarshi & Reiff, 2010, p. 77). It is a result of a much-considered act on the writing scene in its context and specific situation. These considerations are its purpose, context, topic, authorship, audience, structure, and format of the writing act. Arising from these many considerations, this social action becomes one of the "typified ways of acting within recurrent situations and as cultural artifacts that can tell us things about how a particular culture configures situations and ways of acting" (Bawarshi & Reiff, 2010, p. 76). The social action is both a conception from and an end of a multifaceted consideration. It is more than just achieving the purpose of writing a text.

To illustrate this, let us consider the writing scene of a student realizing that he is late for class

owing to a train delay — a common occurrence in Rikkyo. All acts of writing begin with a writing scene, a specific situation where the writer has a social action to achieve. In this case, the student wants to convey a message of excuse for being for late for class. This intended social action of the student sets in motion the consideration of the context of the writing scene. The student considers the *purpose* of the social action: — requesting for an excuse. Next, presumably, the student considers power distance of the *audience*, his teacher. Then, he may think about the *context*, the institutional requirements he needs to comply with in writing his request. The student may next think about the *topic*, what kind of content should be included in the message, what *structure* should be used in the message, and in what *format*. From the above illustration, we can see that a social action is derived from first interrogating the writing scene.

The student's primary social action, a request for an excuse, may carry alongside it other effects that he would like to achieve from his act of writing. In considering his *audience*, a professor, a person of a higher status and position, perhaps respect and goodwill should also be felt in the request. He also realizes when thinking about the *topic*, he should not only state the reason but also show proof to strengthen his *rhetoric* for the excuse. We can see that a social action that may have begun primarily as a request (*purpose*) for an excuse (*topic*), secondary social effects emerge alongside the persuasion (*purpose*) with proofs, and register (*audience*) of formal and polite language. There could be more, but the temporality of quickly putting this request together constrains the amount of time spent in the consideration. It is clear that a social action is not a single clean action but is accompanied by other equally desirable social effects shaped by the writing scene. This coalesced social effect is unique to the student and his writing scene. On a different day with a different writing scene, the message will be different. The interrogation particularizes the social effects as demanded of the writing scene.

RGS as a concept in teaching writing

RGS researchers have used this approach in teaching writing composition (Bawarshi, 2003; Devitt, 2004). It is contended that in order to learn how to perform a communicative act rhetorically in a situated manner, it is important to first understand how the rhetoric is conceived of in specific situation it is located in. Interrogating the genre texts prior to writing raises the awareness of not only how the genre is written but also the why, who, what and when. These questions when discussed, lift the learners from "an interior cognitive process located within individuals" to a "situated cognitive process" (Bawarshi, 2003, p. 10). The practice of writing is not a detached single enterprise, but a multi-dimensional and contextualized endeavor.

Learning through interrogating the social action (genre), including its context and situation, the rhetoric in the writing is understood in genre's situated message and linguistic features. The interrogation looks for regularities *and* irregularities of the social action that typifies the genre. It is not, however, the typified textual forms, but typified *actions*. As illustrated in the section on social action above, students are not learning the typified forms in order to replicate the forms in their writing. RGS is contrasted with genre analysis in that genre analysis expects the replication of these textual forms as expected of the genre. A classic example is Swales' Five Paragraph Essay as an academic essay structure, used for academic writing in most parts of North America (Johns, 2015). Genre analysis is akin to modeling the writing after analyzing the genre. This approach is to mirror what is out there (genre) and produce a similar textual form consistent to the genre type (Swales, 1990). To which, RGS is not.

RGS draws from these typified textual forms to see how they deliver the intended social action. While there are similarities in the genre analysis, the analysis takes it further toward the situatedness of the genre. The goal is to be aware that "genres predict…but do not determine…the nature of a text that will be produced in a situation" (Russell, 1997, p. 522). The focus is then on the awareness of how these typified forms perform the social action. This awareness informs and sensitizes the student to the textual forms that delivers the social action including its nuances. The interrogation of the genre's writing scene becomes the center piece of the RGS approach. The interrogation yields the typified actions and how they are derived and situated in its purpose, context, topic, authorship, audience, structure, and format. Thereafter, the writer having been sensitized, approaches his future writing scenes with the same interrogation *before* writing. He interrogates the writing scene that he is in, shapes his social action *through* this interrogation *with* the awareness of the learned typified social actions. The final coalesced social action will be realized in the writing act of the genre. The writer may end up with a different genre form or even transform the genre type, as dictated by the intended social action.

Implications to pedagogy

A pedagogical RGS will have a vastly different look from the traditional approach of product or process writing styles that most writing teachers are trained for or used to. I will discuss three implications to a RGS classroom. First, RGS is engendering a mindset of interrogation, the centerpiece of RGS: an interrogation of *both* writing scenes of genre types *and* that of the student's writing scene. The interrogation of the genre type looks for *typified social actions* embedded in the context and specific situation of the writing scene. This interrogation acquires the genre (social actions)— genre acquisition (Johns, 2015). Miller summarized this well in the following quote.

> "What we learn when we learn a genre is not just a pattern of forms or even a method of achieving our own ends. We learn, more importantly, what ends (*social action*) we may have…" [Italics mine] (Miller, 1984, p. 165).

This genre acquisition *builds* the knowledge of familiarity with different but typified social actions of different genre types. This genre familiarity building does not cease once the student has acquired this practice; the student by default uses interrogation each time they read to write. The goal is to engender the interrogation mindset in reading *every* single piece of text. In fact, this should become the default reading strategy to cultivate an informed sense of reading and not reading solely for comprehension. This is the first stage of interrogation to acquire the genre.

The second stage of interrogation is genre awareness. This interrogates the writing scene that the student is in *before* the act of writing. This foregrounds an awareness of the student's social action *within* the writing scene. As mentioned earlier, the student formulates their intended social action upon encountering their writing scene. But before they put ink to paper, they interrogate their writing scene to shape their intended social action.

> For the student, "genres serve as *keys* to understanding how to participate in the actions of a community" [Italics mine] (Miller, 1984, p. 165).

From the student's acquisition of typified social actions, they now use this knowledge and apply the

questioning of their intended social action and their writing scene to successfully participate in the community that they are in.

Therefore, this first implication of RGS in teaching writing is the *cultivation* of a mindset in the classroom. To cultivate habits of mind, students need to engage in cycles of cognition (Taczak & Robertson, 2016). The cycle of cognition in RGS is the repeated interrogation of the genre and the writing scenes. Through questions, students identify the purpose, context, topic, authorship, audience, structure, and format. Through these cycles of interrogation, students not only develop a robust understanding of genre but also and more importantly the habit of questioning the text.

The second implication is metacognition. It is established that metacognition (or reflective thought on connections between contexts) plays an important part in the transfer of skills and knowledge (Perkins & Salomon, 1988). Regular reflection can cultivate a "systematic activity keyed to transfer" (Yancey et al., 2014, p. 33). Transferability of writing skills has been of concern in the teaching of writing instruction for some time (Anson & Moore, 2016; Smit, 2004). Students have long been taught de-contextually. As a result, they encounter problems of transfer when they experience contextualized writing situations. RGS scholars have long argued that metacognition is cultivated when students engage in genre analysis (Bazerman, 2009; Beaufort, 2007; Taczak, 2015; Tinberg, 2015).

To engender a situated approach of interrogation, students should be provided with recurring opportunities for metacognition (Taczak & Robertson, 2016). When given the opportunity to question their own thinking as regards planning, monitoring, and evaluation, students engage in a situated mindset. "How will I do this differently? How will I alter my questioning to arrive at a more appropriate information? How does will I change my thinking when confronted with a different scenario?" These are metacognitive questions on an individual's way of thinking. By their very nature, metacognitive questions interrogate the individual's way of thinking in a situated manner. The implication of this in RGS is that students should be afforded the time and space to engage in metacognition to engender the mindset of interrogation. With these cycles of cognition and metacognition designed into the curriculum, not only will students develop the habit of mind (Meyer et al., 2010), but they will also encourage their own transfer to other writing contexts.

When should such metacognition of their thinking processes take place in the classroom? It can occur *after* they have completed their interrogation of the genre and their writing scene. Students' thinking about their thinking processes can be made visible through their written or verbalized reflections on these processes. Reflections can be done during the class after the genre interrogation or as a homework assignment. There can also be teacher and peer feedback on these reflections to learn from each other.

The third implication is the inductive and inquiry-based approach to learning that RGS affords and, in fact, is based on. RGS begins with asking questions of the genre and the writing scene. This characteristic of an inquiry-based learning approach contrasts teacher-led style of classroom. To begin learning with students asking questions is not only highly student-centered, but it also encourages a "bottom up" approach to discover the knowledge that needs to be acquired. Students, in asking questions, derive answers, make connections on their own, and set their own learning. Social actions are not fixed domains of knowledge but are interpreted and negotiated. Students interpret and negotiate their understanding of the social actions *from* the genres and similarly, interpret and negotiate their *own* writing scene to perform the act of writing to achieve their social action. After all, each individual act of writing is highly particularized in a contextualized and situated manner. What is the role of the teacher in a RGS class? The teacher facilitates the process of

interrogation, as well as probes and add more questions to deepen and broaden the students' journey of interrogation. In addition, an RGS class is highly amenable to the dialogic teaching process.
Having outlined three broad implications to teaching writing using the RGS approach, I will now describe how they were attempted in an RGS class in Rikkyo.

An RGS classroom

Rikkyo's reading and writing program is offered to first-year students as a mandatory course involving teaching reading and writing in an integrated fashion over two semesters. The RGS intervention was trialed in the second semester in two stages: genre acquisition and genre awareness. The first half of the semester began with genre acquisition as the starting ground for students to be familiar with the cycle of interrogating the genre. The goal of genre acquisition was for students to familiarize themselves with typified social acts. After the grounding of genre acquisition in the first half of the semester, the second half sought to build this mindset toward genre awareness of interrogating the writing assignment. It was hoped that the students would automatically first interrogate their writing assignment before embarking on the writing process. In other words, for genre awareness, students are expected to draw from their genre acquisition in the first half of the semester to write their first assignment. Selected genre texts were used to for genre acquisition. These were source-based academic writing of exploratory and argumentative nature. Appendix A shows the planned interrogation of the categories of the genre and corresponding tasks in a weekly schedule.

In the "genre acquisition" semester of the first six weeks of classes, students were put through a cycle of pre-reading and quizzes every week together with group discussion on the texts. Diagram 1 shows the learning activities that the students went through on a weekly basis. The target genres, where students were expected to familiarize themselves with the textual features, were read before class for comprehension. During class, students first underwent an individual quiz with questions on their comprehension of the texts. The texts/ topics revolved around description, explanation, and rationale of RGS. The intent was to make students understand the rationale for using RGS in the

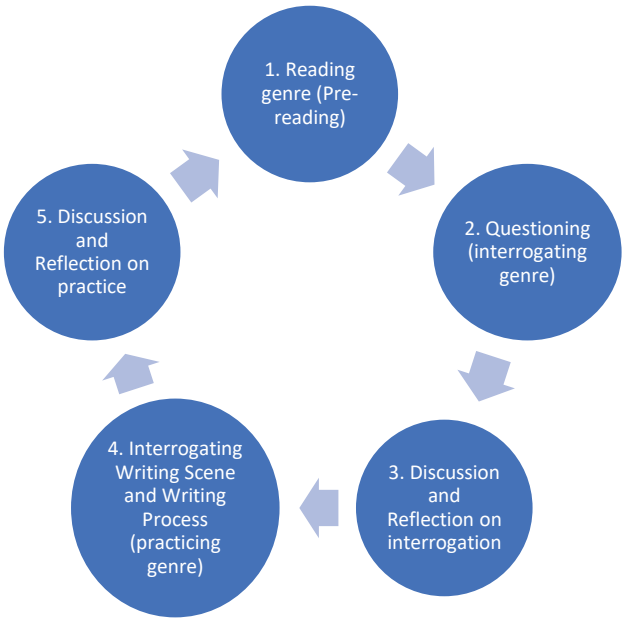


Diagram 1: Cycle of learning activities for genre acquisition and awareness

class. They also explained the different categories of questioning in the interrogation of the genre. This was the practice of reading skills as part of the required goal of this course. The second round of questions, however, were questions that interrogated the genre texts.

With each progressive week, the students were introduced to all the categories as shown in Table 1, interrogating the genre's purpose, context, topic, authorship, audience, structure, and format. Two additional categories were added to cater for language learning and social effect purposes. The category of "sentences" was included to have the students analyze the sentence structure and grammar while "rhetoric" was to draw the students' attention to the social effect of the writing at an earlier stage. This "rhetoric" category was revisited in the final set of questions in the later part of the semester to emphasize and focus on the rhetoric of writing.

Appendix B has examples of the quiz questions analyzing the genre that follows the different categories.

Table 1: Categories of genre analysis

Phases	Categories of questions
Scene and situation of genre	*Context*: Where does the genre appear? What are the circumstances under which the text is written?
	Topic: What is the topic? What issues and ideas does the genre address?
	Author: Who writes this type of text (genre)? What role do they perform as writers?
	Audience: Which type of text (genre) is this? What role do they perform when they read?
	Purposes: Why do authors write this genre and why does the audience read it?
Patterns of the genre's features (What recurrent features do the samples share?)	*Content*: What content is usually included? What is excluded?
	Rhetoric: What rhetorical appeals are used? What appeals to logos, pathos, and ethos?
	Structure: How are texts in the genre structured? What are their parts, and how are they organized? What do you notice is different in the patterns of the texts in the genre? Does the difference affect the purpose of the texts?
	Format: What is format of the texts of this genre? What is the layout or appearance? How long is the text in this genre?
	Sentences: What are the types of sentences? How long are they? What tenses are they in? Are they simple or complex, passive or active?
Rhetorical patterns	How does the rhetoric patterns affect the reader and the purpose for the writer? How appropriate is it to the situation of the writing context, the audience, and community of practice? How does this writing achieve the effect for the author and the reader?

The individual quiz allows the student to analyze the genre on their own with no feedback provided. The lack of feedback feedback allows later discussion on the same set of questions for contestation and social construction. Here at the discussion stage, students are expected to convince others of and defend their choice, affording a robust discussion and analysis of the question and in doing so, continue the interrogation of the genre. The discussion stage reveals the "correct" answer to the questions. Through this cycle of choice, discussion, and reflection, students interrogate the genre at length.

After all the questions are resolved, the students were given an opportunity to note down their doubts on the issues raised and request for further clarity. This stage allows students to reflect and think beyond what was the "right" answer as prescribed in the feedback. These challenges and questions were later discussed in the class. Here was a deliberate attempt to demonstrate to the students that their voice, with reason, were equally acceptable as "correct" answers other than what was prescribed by the teacher. The deliberation at this stage affords a deeper discussion on and

understanding of the genre.

After the discussion of the questions and answers and supplementary instructions, the groups moved on to the application stage where they collaborated on the interrogation on their writing assignment in groups. Students went through one round of asking the questions shown in Table 1. After the discussion, the students embarked on their individual writing assignments. The writing went through one draft with feedback provided by peers, guided with a criterion-based checklist , and the second draft was submitted as the final draft.

After the submission of the final draft, the groups of students were given a discussion activity to reflect on their writing process. Semi-structured questions were provided to guide students to evaluate and suggest improvements on their thinking processes. This activity was to engender a culture of metacognition. This cycle of writing was repeated for the second writing assignment in the latter half of the semester.

Classroom pedagogy for RGS

As discussed in the pedagogical implications of RGS, the learning philosophy of a RGS class is highly student-centered with students beginning the class by asking questions. In answering the quiz questions in groups, students began interrogating the genre initially as an individual and later as a group. The teacher's input came in the form of the automated answers in the quiz. Even then, students had opportunities to question the prescribed "correct" answers in the quiz as well as the questions. This inquiry style approach together with collaborative learning is reflective of the RGS's interrogative style as well as the constructivist notion of learning.

The repeated cycles of interrogation weekly on different samples of the genre not only allow the students to cumulatively acquire the genre type but also cultivate the habits of mind. It is hoped that these repetitive processes will engender this habit of mind of the students to engage in genre interrogation before writing. Similarly, opportunities for metacognition were afforded.

Having gone through what I would consider to be an experimental attempt at approaching RGS in my reading and writing class, there are several gaps that emerged. I will briefly discuss three of challenges: situatedness, metacognition, and assessment.

The challenge of situatedness

Each writing social act is embedded in a system of activity (Bazerman, 2017) in which we find ourselves. Whether it is in response to a written request or an initiation, we locate ourselves as an agent within the system, a system that affords its peculiar socio historical cultural context and network of agencies. This is the writing scene that the writers find themselves in. Moreover, when a teacher tasks the student with writing the system of activity that the student and teacher are in (let's call it the classroom genre) presents the situatedness of the writing context. How then does the student project themselves into performing the writing act of the target genre with this classroom genre situatedness "interfering" with the writing scene? The actual writing act, derived not out of a necessity or desire of the writing scene of the target genre, is supplanted by an injunction albeit contrived by the teacher. The teacher's requirement of a writing submission dominates the rhetorical genre analysis process where the student's consideration of his writing act is not located within the writing scene *but* located within the classroom genre system. In short, the student is confronted with and will be confused by *two* writing contexts: the target genre, and classroom writing scene

The nature of RGS in writing is a response, a social act, in a specific situation and context of the writing scene. In analyzing the genre within this framework, the student acquires the awareness of the writing genre by interrogating the writing scene. Through this process of interrogation, the student familiarizes themselves to the typified social acts of writing within these socio cultural contexts. In turn, with the familiarity of these typified socio cultural contexts, the student will arguably be able to respond adequately in similar social acts of writing upon encountering such socio cultural contexts. Inasmuch as this familiarity of these typified social acts are desired, the goal of RGS is not so much as acquiring this familiarity rather the ability to be aware of genres and begin engaging metacognitively as a writer in these writing scenes by first interrogating them.

If interrogation is the primary goal, then, when the writing scenes are of the classroom genre type of writing situation, the interrogation will be limited in its scope as students are regularly presented the same classroom genre throughout the course. The interrogation would repeatedly and invariably result in students identifying themselves as students and the audience as the professor to whom they are writing. Hence, for all the pieces of writing social acts, the contexts are rather fixed.

It is my suggestion therefore that for students to engender the interrogation of the context and situation in their future writing scenes, teachers can design role-playing writing scenarios for students to "act" out the scene. In so doing, students will interrogate their writing scene as they would, if they are encountering such similar situations in real life. While role playing is at best a simulation, it can offer a more realistic situatedness of the genre (social action) analysis when compared with the classroom genre type. Because "writing involves so many problem-solving judgments, it is best learned through a long sequence of varied problem-solving experiences in varied situations" (Bazerman, 2017, p. 34).

The challenge of metacognition

So, what exactly was being metacogitated? In RGS, students first familiarize themselves with the genre presented by interrogating other scenes of writing and thereafter interrogating their own writing scenes to write that genre. The metacognition is then, thinking about how they question the genre and how they practice the genre. Hence, it is not only thinking about the genre (Tinberg, 2015), the knowledge required for writing, but also the reflection on how they thought about the genre, including thinking about for thinking about their practice of writing and reflecting on their thoughts on their writing scene. It is instructive to note that the reflection here is "a mode of inquiry: a deliberate way of systematically recalling writing experiences to [frame or] reframe the current writing situation" (Taczak, 2015, p. 73). So, it is evaluating and improving their interrogation of the genre and not just recounting the process.

While my RGS class did afford opportunities for metacognition, there was no feedback and means to ascertain the types of metacognition that could have been discussed and improved. It was well and good that students engaged in discussion on what thinking to improve on, but how do I know this happened and what exactly was discussed? More likely my students were recalling their thought processes but did not seriously work on making changes to the way they think or question. The reflection activity assumed that metacognition occurred naturally or incidentally. While opportunities or activities for metacognition should be designed into an RGS classroom, some visible form should also be put in place for feedback and specificity.

Second, metacognition involves thinking and questioning one's planning, monitoring, and evaluation (Taczak & Robertson, 2016). How did I organize my interrogation of the writing scene?

Was the process adequately performed? What aspects should I prioritize in the interrogation? What was my criteria in evaluating the purpose, context, topic, authorship, audience, structure, and format? Did I give sufficient time to interrogate the writing scene? These questions are what my students should have examined in metacognition. Yet, though ideal, these may require more time and space for engendering a good practice of RGS interrogation. These are clearly difficult challenges to the practice of RGS.

The challenge of assessment

RGS views genre as a social action. The acquisition and awareness of the genre occurs by interrogating the social action and writing scene of the genre. The first challenge in assessment is assessing the social action. How do we assess social action? Social actions are complex forms of "typified ways of acting within recurrent situations and as cultural artifacts that can tell us things about how a particular culture configures situations and ways of acting" (Bawarshi & Reiff, 2010, p. 76). They are complex because of the very nature of "the social motives, relations, values, and assumptions embodied within a genre that frame how, why, and when to act" (p. 77). Clearly there are multifaceted and inter-connected values to consider when evaluating a social action in a piece of writing. The complexity of social action is a challenge to be observed and evaluated. If one were to deconstruct and concretize the social action into criteria for evaluation, it can be an unenviable task.

However, if social action are "typified ways of acting" embodying "the social motives, relations, values, and assumptions" of an actor, it is the actor who can provide the criteria for assessment. After all, the social act of writing that emerges out of the multifaceted and inter-connected values of purpose, context, and rhetoric of the writing scene seeks to realize *that* social act of the writer. Second, the social act that is intended by the writer has to be evaluated by the reader of the social act of writing. The reader, performing the role of the intended audience in the writing scene, is well positioned and valid to make judgments on the effect of intended social act. The reader, then, in his perception of the intended social act of the writer, makes the judgment.

Therefore, it would seem natural to include the intended social action of the writer as part of the assessment, as the intended social action would embody the confluence of the "the social motives, relations, values, and assumptions embodied within a genre that frame how, why, and when to act" (Bawarshi & Reiff, 2010, p. 77). This can be achieved with the writer, in this case, the student, to indicate the intended social action, which will form part of the assessment criteria of the writing assignment. The assessor, would in turn, play the role of the reader of the writing scene to evaluate the expressed intended social action. This would also "complete" the genre writing scene, unless of course if the scene entails action in a genre system. The inclusion of the student's criteria in assessment is reflective of the participatory act of students in having a say in the curriculum design (Eisner, 2001; Freire, 1993).

The second challenge to the assessment in an RGS class is the assessment of metacognition, the linchpin of RGS. RGS is about cultivating knowledge about writing (genre acquisition) through the practice of interrogating the writing scene (genre awareness) in order to perform the social action (genre). The first consideration is its necessary inclusion in the assessment of the writing act. Writing teachers have long evaluated pieces of writing as finished products including drafts. Metacognition, or thinking about thinking, is rarely observed, and assessed. If metacognition processes are prized in RGS, it would be meaningful to include it in the assessment.

But how do we assess this cultivation and practice of interrogation? Writing teachers have long

used written drafts to observe the progress of students' writing process and provide feedback. Writing portfolios that include prewriting and drafts of the same writing piece has been used to trace the development and editing of a piece of writing. However, cultivation and practice of interrogation are thinking processes and are not discrete knowledge of genre. How can we observe these processes of planning, monitoring, and evaluating of their thinking? Think-aloud protocols and reflective diaries are known to record metacognition of students, including surveys and interviews. Whatever forms it takes, it is suggested that in RGS classrooms, metacognition should be observed and assessed.

Conclusion

In discussing the concept of RGS in writing, recasting genre as a social action offers educators an authentic way to teach writing as it should be — interrogating the writing scene, both as a way to acquire the genre and a resource to "participate in the actions of a community" (Miller, 1984, p. 165). The pedagogical implications of RGS necessitate the engendering of habits of mind of constant interrogation of texts as a reader and interrogation of writing scenes as a writer. Such mindsets will surely put any student of reading and writing in a competent place to be a successful communicator of social acts. It is acknowledged that in seeking to cultivate habits of mind to engage in metacognition before and during a writing social act is not an overnight exercise but one that requires sustained and regular practice (Meyer et al., 2010). A study into the effects of this approach should suitably follow to illuminate this. The challenges to the RGS classroom emerge when the writing pedagogy enters uncharted territories of engendering of habits of mind of constant interrogation of texts. But when and if we are able to appropriate the challenges of situatedness, observe and assess metacognition, and cultivate mindset, we may reap untold treasures in our students' ability to write well in their ever-changing writing scenes and genre.

References

Aciman, A., & Rensin, E. (2009). *Twitterature: The World's Greatest Books Retold Through Twitter*. Penguin.

Anson, C. M., & Moore, J. L. (2016). *Critical Transitions: Writing and the Question of Transfer*. The WAC Clearinghouse: University Press of Colorado.

Bakhtin, M. (1986). The problem of speech genres. In C. Emerson, & M. Holquist (Eds.) *Speech genres and other late essays* (pp. 60-102). Austin: University of Texas Press.

Bawarshi, A. (2003). *Genre and the invention of the writer: Reconsidering the place of invention in composition.* Logan, UT: Utah State University Press.

Bawarshi, A., & Reiff, M. J. (2010). *Genre: An introduction to history, theory, research, and pedagogy.* West LaFayette, IN: Parlor Press.

Bazerman, C. (2009). Genre and cognitive development: Beyond writing to learn. In C. Bazerman, A. Bonini, & D. Figueiredo (Eds.) *Genre in a changing world* (pp. 283-298). Fort Collins, CO: The WAC Clearinghouse and Parlor Press.

Bazerman, C. (2013). *A theory of literate action*. Fort Collins, CO: The WAC Clearinghouse and Parlor Press.

Bazerman, C. (2017). The psychology of writing situated within social action: An empirical and theoretical program. In P. Portanova, J. M. Rifenburg, & D. Roen (Eds.) *Contemporary*

perspectives on cognition and writing. The WAC Clearinghouse; University Press of Colorado.

Beaufort, A. (2007). *College Writing and Beyond: A New Framework for University Writing Instruction*. Logan, UT: Utah State UP.

Campbell, K. K., & Jamieson, K. H. (1979). Form and genre in rhetorical criticism: An introduction. In K. K. Campbell, & K. H. Jamieson (Eds.) *Form and genre: Shaping rhetorical action* (pp. 9-32). Falls Church, VA: Speech Communication Association.

Devitt, A. J. (2004). *Writing genres*. Carbondale, IL: Southern Illinois University Press.

Eisner, E. W. (2001). What does it mean to say a school is doing well? *Phi Delta Kappen*, 82 (5), 367-372.

Freire, P. (1993). *Pedagogy of the oppressed*. New York: Continuum.

Johns, A.M. (2015). Moving on from genre analysis. *Journal of English for Academic Purposes*, 19, 113-124.

Meyer, J. H. F., Land, R., & Baillie, C. (2010). Editor's preface. In J. H. F. Meyer, R. Land, & C. Baillie (Eds.) *Threshold concepts and transformational learning* (pp. ixxlii). Champaign, IL: Sense.

Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70, 151-167.

Miller, C. R., & Kelly, A. R. (Eds.). (2017). *Emerging genres in new media environments*. Palgrave Macmillan/ Springer.

Perkins, D. N., & Salomon, G. (1988). "Teaching for Transfer." *Educational Leadership 46.(1)*, 22-32.

Russell, D. R. (1997). Rethinking genre in school and society: An activity theory analysis. *Written Communication, 14,* 504-554.

Smit, D. (2004). *The End of Composition Studies*. Carbondale: Southern Illinois UP.

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Taczak, K. (2015). Reflection is critical in the development of writers. In L. Adler-Kassner, & E. Wardle (Eds.) *Naming what we know: Threshold concepts of writing studies* (pp. 78-79). Logan, UT: Utah State University Press.

Taczak, K., & Robertson, L. (2016). Reiterative reflection in the twenty-first-century writing classroom: An integrated approach to teaching for transfer. In K. B. Yancey (Ed.) *A Rhetoric of Reflection* (pp. 42-63). Logan, UT: Utah State University Press.

Tinberg, H. (2015). Metacognition is not cognition. In L. Adler-Kassner, & E. Wardle (Eds.) *Naming what we know: Threshold concepts of writing studies* (pp. 75-76). Logan, UT: Utah State University Press.

Yancey, K. B., Robertson, L., & Taczak, K. (2014). *Writing across contexts: Transfer, composition, and sites of writing*. Logan, UT: Utah State University Press.

Appendix A
Outline of weekly schedule of RGS in a semester

	RGS progression	Tasks
1	purpose, context, topic, author, and audience.	Reading 1, Quiz, Discussion
2	purpose, context, topic, author, and audience.	Pre-Reading 2, Quiz, Discussion
3	purpose, context, topic, author, audience, and structure.	Pre-Reading 3, Quiz, Discussion
4	purpose, context, topic, author, audience, structure, and sentences	Pre-Reading 4, Quiz, Discussion
5	purpose, context, topic, author, audience, structure, and sentences	Pre-Reading 5, Quiz, Discussion, Writing
6	purpose, context, topic, author, audience, structure, format, rhetoric, and sentences.	Pre-Reading 6, Quiz, Discussion, Writing
7	purpose, context, topic, author, audience, structure, format, and rhetoric of Writing Assignment 1	Draft 1, Peer review
8	purpose, context, topic, author, audience, and structure.	Pre-Reading 7, Quiz, Discussion Draft 2 Submission, Metacognition
9	purpose, context, topic, author, audience, structure, rhetoric, and sentences.	Pre-Reading 8, Quiz, Discussion
10	purpose, context, topic, author, audience, structure, rhetoric, and sentences.	Pre-Reading 9, Quiz, Discussion
11	purpose, context, topic, author, audience, structure, and rhetorical patterns.	Pre-Reading 10, Quiz, Discussion
12	purpose, context, topic, author, audience, structure, and rhetorical patterns.	Pre-Reading 11, Quiz, Discussion
13	purpose, context, topic, author, audience, structure, and rhetorical patterns.	Pre-Reading 12, Quiz, Discussion
14	purpose, context, topic, author, audience, structure, format, and rhetorical patterns of Writing Assignment 2	Draft 1, Peer review Draft 2 Submission (1 week after class)

Appendix B

Question 8.

This question is about content. In the first sentence, this phrase, "...students are given the space to question, analyze, and understand why and how a text is written... (L6-L7), what is the writer using the content for?

Answer

a. The author is describing the audience of the RGS.
b. The author is describing what RGS is and the process of RGS.
c. The author is describing what RGS is and knowledge transfer.
d. The author is describing what RGS is and the disadvantage of RGS.

Question 6.

This question is about the structure of Pre-Reading 3A. Which of the following best describes how the different parts are organized?

Answer

a. Introduction of RGS - Summary of what is RGS - Benefits of RGS.
b. Introduction of RGS - Benefits of RGS - Future of RGS.
c. Introduction of RGS - Benefits of RGS - Summary of what is RGS and benefits.
d. Introduction of RGS - Summary of what is RGS and benefits - Future of RGS.

Question 6.

This question is about the sentences in the Pre-Reading. What types of sentences does this text generally use?

Answer

a. Present tense.
b. Past tense.
c. Continuous tense.
d. Future tense.

Question 7.

This question is about the language AND organization of the Pre-Reading 4. How does the author use language to organize the important ideas?

Answer

a. Uses cause and effect: hence, therefore, etc.
b. Uses problem and solution: state the problem and give the solution.
c. Uses numbering to sequence the ideas: 1st, 2nd, 3rd, etc.
d. Uses conjunctions to connect the ideas: Lower case. but, however, etc.

Question 5.

This question is about the Lower case use in the Pre-Reading 7. How did the author convince you convince you of the number of the rhetorical strategies to use at the end of the essay?

Answer

a. The author uses logos to convince me to write them well, which is better than using all three.

b. The author uses ethos to convince me to write them well, which is better than using all three.
c. The author uses pathos to convince me to write them well, which is better than using all three.
d. The author uses words like “write them well” and “poorly” to convince me to write them well, which is better than using all three.

Question 6.

This question is about the Lower case use in the Pre-Reading 7. Which of the following phrase did the author use as part of the rhetorical strategy to convince you on how many of the rhetorical strategies to use at the end of the essay?

Answer

a. “While it may be more convincing to use all the three strategies,..”
b. “It is more important to write them well in our essays to achieve the effect.”
c. “A poorly written essay with all three may not be as persuasive as one that is well written.”
d. “Rhetorical strategies such as logos, ethos, and pathos are necessary...”

Group Discussion Application Questions

Refer to the Sample essay titled, “Is Fashion Really Important?” (315 words)
Discuss in your group and answer this question.
Analyze the structure of this essay.

1. How many reasons did the writer use to support his claim, “that fashion is important” (L2)? State the number.
2. Did the writer write something negative about fashion? (Yes or No)

Refer to the Sample essay titled, “Gun Control Essay” (313 words)
Discuss in your group and answer these questions.

TAQ4: Analyze the claim of this essay.

1. How many supports did the writer write in his claim? State the number.
2. Is the claim clearly stated? (Yes or No)
3. Did the writer restate the claim of his essay again in the conclusion? (Yes or No)

【Research Article】

Factors Influencing the Establishment of Professional Development Network During Emergency Remote Teaching: An Activity Theory Analysis

Joshua Rappeneker, Satchie Haga, James Carpenter, and Daniel Beck

Abstract

Professional development networks (PDN) can improve learning, knowledge sharing, and school objectives. Most research conducted on PDNs are evaluative and examine the outcomes of existing PDNs or highly structured newly formed PDNs developed specifically for research purposes. Much is unknown about the sociocultural influences that affect the development of organically formed PDNs. Using qualitative methods and cultural–historical activity theory (CHAT) as a lens, this study examines the experience of four teachers engaged in the development of a newly formed professional development network focused on developing the digitals skills necessary to cope with the switch to emergency remote teaching during a global emergency—COVID-19. Results indicate that the usual formal/informal dichotomy between PDNs does not necessarily always hold true. A CHAT analysis of the formation of the committee demonstrates how an ad hoc PDN can be transformed into a formal top-down network. Furthermore, the analysis demonstrates the significant mediating role of communication tools in establishing and maintaining a PDN.

Keywords: *Emergency Remote Teaching, Professional Development Networks, Cultural Historical Activity Theory, ICT*

Background

With the advent of the novel Coronavirus (COVID-19) and the subsequent declaration of the global pandemic in March 2020 (Cucinotta & Vanelli, 2020), universities worldwide started to rapidly adapt their curricula to comply with health and safety guidelines (Marinoni, et al., 2020). In Japan, the majority of universities switched from face-to-face classes to hybrid or fully online learning (Mainichi, 2020). At Rikkyo University in Japan, an ad hoc committee was formed in the foreign language department in order to facilitate the development of the necessary skills in quickly transitioning to online instruction among faculty members. This digital communication committee (DCC) was tasked with developing and distributing materials to assist faculty in the use of information and communication technology (ICT). The emergency nature of the committee, along with the unprecedented nature of their task, meant that the specific goals and methods of the committee were initially undefined. Thus, the committee was forced to improvise in developing these materials whilst navigating the technological and cultural landscape of early emergency remote teaching in Japan—a country that consistently ranked the lowest amongst developed countries in terms of the use of online tools across industries prior to the pandemic (Inoue, 2021).

As the committee's work progressed, it began to focus on three key activities: the development of a website with original information and tutorials on the use of ICT in online instruction, as well as curated lists of links to other existing sources of information; the creation and sharing of videos explaining the use of specific ICT tools (e.g., Blackboard, Zoom); and the creation and running of a helpdesk to answer faculty questions about ICT in the context of online language teaching. In this way, the activities of the committee intersected with the activities of other academic units on campus, becoming influenced by the university's internal logic (Shangraw & Crow, 1998). The overarching

design of an organization that guides the decision-making process of individual actors is difficult to describe (Crow & Shangraw, 2016). In this study, we adopt the analytical vocabulary of cultural–historical activity theory (CHAT) because it provides analytical concepts for describing the sociocultural relationships, that is, the relationship between individual actors, the organization in which they act, and the larger cultural (or global) context, that dictate what teachers can and cannot do. In adopting this theoretical framework, we hope to clarify the relationship between the activities of the committee, others in the same department, and the rest of the campus. As we will show, the DCC acted as the nexus for a professional development network (PDN) unique to the historical moment: The rapid transition from face to face to online teaching in response to COVID-19. This PDN was both peer-to-peer or bottom-up—in that it started as an informal collection of like-minded instructors attempting to prepare for and share knowledge about emergency remote teaching (ERT)—and regulated or top-down—as it was integrated into the foreign language department's official committee structure. This combination of peer-to-peer and formally structured elements affected how the committee functioned within the university system.

Literature Review

Professional Development Networks

Broadly construed, PDNs are platforms, or combinations of platforms, that allow instructors to search for and share professional development knowledge (Trust, 2015). They make use of social networks, collaboration, and online tools such as wikis, forums, and social media groups (e.g., LINE groups), to find and share resources, links to useful websites, lesson plans, and other relevant knowledge (Trust, 2017). In the last decade, education scholars and practitioner researchers have investigated the impact of ICT on learning spaces. In an early influential paper, Siemens and Matheos (2010) theorized that such spaces—in both formal and informal educational contexts—will become more fluid: Educators will take advantage of the network aspects of ICT to create personal learning environments or open network learning environments. In these spaces, teachers and learners creatively integrate different online tools in highly personalized ways to accomplish teaching and learning tasks (e.g., Tu, et al., 2012). Whereas this research has largely focused on how students accomplish tasks in online spaces, the concept of the PDN provides a means of describing how teachers, functioning within a larger organization, accomplish tasks in a networked environment.

PDNs can be categorised as either formal or informal groups. Lantz-Andersson et al. (2018) examine both "formally-organised and informally-developed"[1] professional development groups and explore the contexts in which these groups formed. Lantz-Andersson differentiates between formal and informal groups in terms of the manner of the groups' initial creation. They define formal groups as "top-down professional development endeavours, initiated by schools…" (p. 304). In contrast, informal groups are defined as "bottom-up initiatives involving a group of practitioners who choose to come together to discuss, share information and work together". Macià and García (2016) focus their survey on informal networks and communities for teacher professional development. They define informal communities or networks as "collaborative processes of knowledge sharing" of informal learning mediated by technology, using Watkins and Marsick's (1992) definition of informal learning as "learning from experience that takes place outside formally structured, institutionally sponsored, classroom based activities".

Previous studies have demonstrated the benefits of PDNs on development of institutional goals

(Rieckhoff & Larsen, 2012), teacher professional skill development, confidence, and knowledge sharing (Alemdar & Rosen, 2011; Cutts, et al., 2017; Trust, 2017), and student learning and social justice (Lawrence & Dubetz, 2001; Thomas, 2007). However, PDNs are socioculturally complex spaces (Burns Thomas, 2004) with barriers that can negatively impact and shape the components of the community.

Lantz-Andersson et al. (2018, p. 310) found that time issues were a significant barrier to participation in informally developed online learning communities. Teacher participation in the communities was "often conducted after regular school hours" and could be seen as intruding on teachers' time outside of work. Macià and García (2016, p. 300) discovered that a "gradual lack of engagement" results in members using the network less to the point of "drop out". They also found that reluctance to participate was the result of fear of criticism, and insecurity in "sharing [their] own ideas" proved a barrier to participation.

In contrast to the large number of studies of formal professional development (e.g., Cutts et al., 2017; Lavicza et al., 2010) communities, there is a dearth of research on informal PDNs, and in many cases, the informal communities that are analysed were not "organically" formed, but were developed specifically to be researched (Macià & García, 2016). Moreover, despite the extant literature demonstrating the sociocultural complexities that shape PDNs and their components (Thomas, 2007; Trust, 2017; Warschauer, 2020), research that specifically examines sociocultural influences that affect PDN formation has unsurprisingly received little attention until now. This study aims to add to the literature by examining the sociocultural factors that influenced the emergence of an informal PDN created to address the challenges facing a language department at a university in response to a global emergency—including the changes that PDN underwent as it was formalised by the institution.

Emergency Remote Teaching (ERT) Context

ERT is defined by Hodges and associates (2020) as the rapid and temporary move to online learning in times of crisis or emergency. Hodges states that it is important for institutions to recognise the difference between ERT and well-planned online learning experiences. Rapid transitions to online learning present many challenges and obstacles (Crawford et al., 2020); thus, how instructors go about the work of planning and conducting their classes will be significantly different from stabler times. The same distinction applies to how teachers search for and share professional knowledge during times of stability as opposed to times of crisis (Webster-Wright, 2009), where the most important factor is time: The transition from face to face to online in ERT situations is by necessity short.

One significant difference intrinsic to ERT is the need for flexibility in defining pedagogical policy, and the ways that faculty knowledge is shared. Gacs and associates (2020) recommended that whilst universities should provide teachers with ample technological and other professional development training for online teaching, during emergencies, instructors "may need to rely on ad hoc or already established personal learning networks" (Gacs et al., 2020). Similarly, as Hodges et al. (2020) state, expectations of quality must also be tempered. During the first year of the COVID-19 pandemic, two thirds of higher education institutions reported replacing face-to-face teaching with distance learning (Marinoni et al., 2020). As noted above, Japanese universities were notoriously slow to adapt to online technology before the pandemic (e.g., Funamori, 2017), which makes the implementation of ERT policies particularly difficult (Inoue, 2021).

Theoretical Framework

What is Cultural Theoretical Activity Theory (CHAT)?

CHAT provides a framework for analysing decision making in communities through examination of relationships between goals, actors, and tools, and the cultural and historical context within which they act as members of some community. (Cole & Engeström, 1993). CHAT considers the *activity system* as the "primary unit of analysis". (Engeström & Sannino, 2021). Activity systems define networks of sociocultural contexts and the structures and tools that shape actors and their behaviour in attempting to achieve goals (Engeström, 2000). Activity systems can also include the interaction of smaller sub-systems (Engeström, 2001) (See Figure 2). Each individual actor in a community can be considered a system, as much as collections of actors in a like-minded group can. Activity systems consist of an object, one or more subjects, mediating artefacts (tools), a community, division of labour, rules, outcomes, and contradictions (See Figure 1) (Cole & Engeström, 1993). The individual elements are described further below. As noted above, CHAT provides useful tools to analyse complex sociocultural situations (Bligh & Flood, 2017) and thus was selected as an appropriate theoretical framework to examine the evolution of the PDNs in this study. Figure 1 below depicts the key features of the chat model and the dialectical relationship between them.

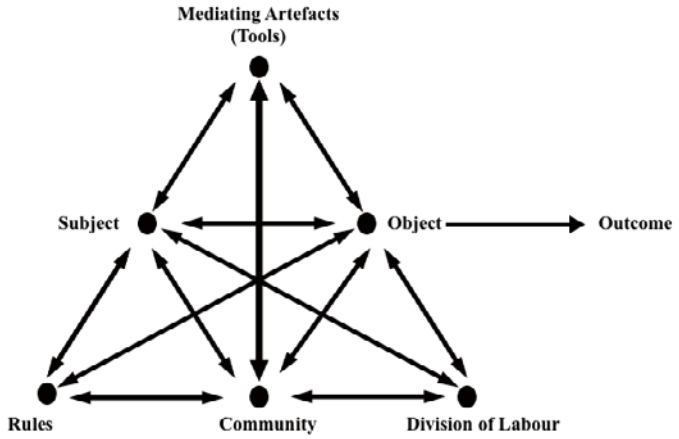


Figure 1. An activity system and its dialectical relationships

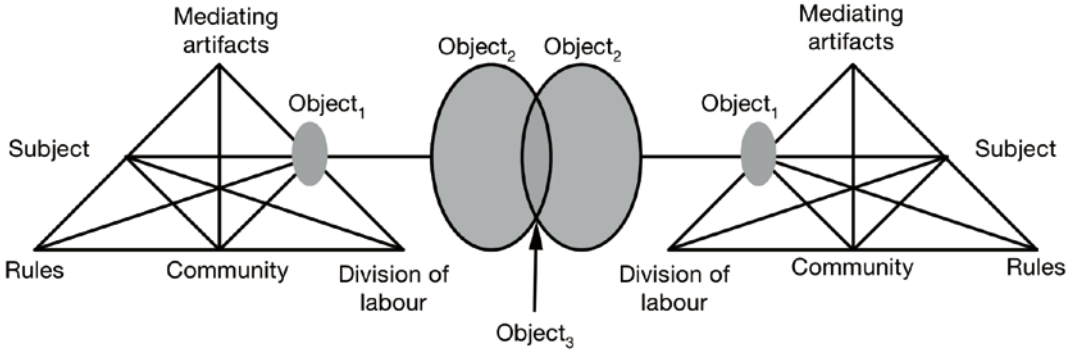


Figure 2. Two CHAT activity systems interacting inside a larger system (Engeström, 2001)

In order to reduce the scope of analysis, this study focuses on the DCC activity system in and of itself and will not be considering the greater activity system of the institution in the data analysis.

Object

According to Leont'ev (1978), actors are motivated by and towards an external object. This object is the goal or motivating force for an activity system. Objects are transformed into outcomes by the actions of subjects within a system (Engeström & Sannino, 2021). Each subject within a system may have a different motivating object, and therefore may make different decisions than other actors in the same system.

Subject

The subject of a system is the person or persons acting within the system towards some object. Subjects are affected by their actions, the actions of others, and their conflict or interaction with other elements of the system. Subjects' motivations towards the object shape their experience and behaviour within the system.

Mediating Artefacts (Tools)

Tools are resources (cognitive or material) that mediate the subject's actions within a system. Cognitive tools can be socially shared frameworks (e.g., mathematics, times tables) or personal understanding of the external world. Material tools can be physical (e.g., a ruler, a pencil) or more abstract (a website or social media page). The tools chosen by subjects shape their decisions and their potential actions within a system. Changing toolsets greatly modifies the context of an activity system.

Community

Community is the group of subjects working towards the object within a system. Whilst individual subjects within a system may have differing objects, they are understood to be mutually motivated to act within the activity system to share some of the system's cultural context.

Division of Labour

Each member of a system has a different role, which is negotiated by the community based on sociocultural rules, the member's abilities and knowledge, and the member's interests and motivations. Members' roles may change as subjects interact with the system and pursue the object.

Rules

Rules regulate the behaviour of the subjects in their pursuit of the object. Rules may be formal (e.g., work contracts, labour laws) or informal (e.g., sociocultural norms, community hierarchy). These rules create the structure within which culturally appropriate action may be undertaken in the system and indicate the place of subjects within the community.

Outcome

The outcomes are the consequences that a subject experiences owing to their pursuit of the goal. These outcomes need not be the realisation of the object of the activity. They may be the unintended consequences of the activity system, or even results that the subject perceives as negative.

Contradictions

Contradictions are the internal conflicts of an activity system. These conflicts may exist within a single activity system (e.g., between rules and division of labour), between two or more objects (e.g., tension between subjects with different goals), or two or more activity systems (Roth & Lee, 2007; Trust, 2017). Contradictions may lead to changes in the activity systems at a fundamental level.

Using CHAT as a framework allows researchers to understand how instructors formed a PDN and pursued the goals of finding and sharing knowledge with peers in the context of the early ERT period. CHAT in particular provides insights into the tensions and contradictions of interactions between objects, tools (significant in the case of ERT), and subjects in this fraught sociocultural context.

Methodology

Research Question

- What factors influenced the formation and functioning of the DCC?

Data Collection

Data collection was designed to, as much as possible, fully reflect the experience of each subject in the activity system. Data were initially collected through a 90-minute interview and discussion over Zoom with all four members of the committee. At the beginning of the discussion, the basic elements of CHAT were outlined and explained to the members of the group (object, subject, tools, community, division of labour, rules, outcome, and contradictions). The reasons for choosing CHAT analysis were discussed, and then the discussion proper was undertaken. Each element of CHAT was discussed in the context of the committee's work, and each member was asked to give their definitions of that particular element and describe their experience of this element. For example, the first question asked during the discussion was "What did you perceive the purpose of the committee to be?" This question was designed to elicit the object of the activity system from the point of view of each subject. Another question asked later during the discussion, in order to elicit subjects' perceptions of contradictions within the system, was "Did you experience any other kinds of… friction or tensions in any of …the rules or the community or the tools…?"

The discussion was conducted via Zoom considering Zoom's automatic transcription feature. This made analysis of the discussion afterwards much simpler. Any issues with the automatic transcription were clarified with the relevant members via email.

The first question of the survey relating to the framework is show in Table 1:

Table 1

Question 2 of the survey

Object=The goal of the activity.
"Creation of a committee that provided resources to help teachers get online." How far would you agree/disagree that this was our "object". If you were to change it, how would you describe the object we were working on?
Members of the committee were given a textbox to add their optional response.

Each subsequent element of CHAT was presented in a form with the answers that the researcher had extracted from the initial discussion. Outcomes were split into two questions, positive and negative. Question 8 (positive outcomes) was as follows:

Table 2

Question 8 of the survey

	Outcome=result of the activity. (any other positive outcomes)? Positive:
i)	Provided faculty support for technical issues during the first academic semester of the pandemic.
ii)	Experience of working between different cultural groups and committees within Japanese universities
iii)	Insight into the different users that would need faculty development related to digital skills
iv)	Recognition from peers who were helped during the time the committee was active and those that still use the videos today
v)	Recognition in the Zenkari newsletter about our activities
vi)	Conducting this study informed self-reflection and transformative learning
vii)	Empowered - gave the ability to make the website and videos
viii)	Motivating - respect for expertise separate from language teaching
ix)	Individual skills improved - e.g., how to create an institutional website
x)	Developed stronger professional relationships with those we helped

Again, members were asked to provide further examples or corrections.

Data Analysis

An iterative approach was used to analyse the data. This began with multiple readings of the interview transcript to immerse the researcher in the data and gain an overarching perspective of the responses (Hsieh & Shannon, 2005). Data immersion and discussions about the data enabled the researchers to engage in a reflective process reflecting on personal beliefs that could potentially bias the data analysis and interpretation (Creswell & Poth, 2016). Following this the transcript was entered into Atlas.ti and an inductive approach based on grounded theory (Corbin & Strauss, 1990) guided the coding where in vivo coding procedures were used to maintain the integrity of the participants' words and experience (Miles et al., 2014).

Once this initial coding was complete, a constant comparative method (Glaser, 1965) was performed on the codes to group similar ideas together and create categories cantered on the factors that influenced the development of the PDN identified by the committee members. These data were organized within the activity theory framework categories.

The coding and analysis of the transcript was conducted by one member of the research team who then conducted a member check with the other members to triangulate her findings and provide stronger reliability of the qualitative analysis. This second data collection was conducted via online survey form. Initial key themes that emerged from the data analysis were shared with the other team members to resolve differences in interpretation and examine any new ideas related to the categories that emerged from the data in relation to the central research question guiding the study.

Findings

Figure 3 demonstrates the emergent themes from the DCC activity system. The following sections details the elements and how they shaped the development and functioning of the DCC.

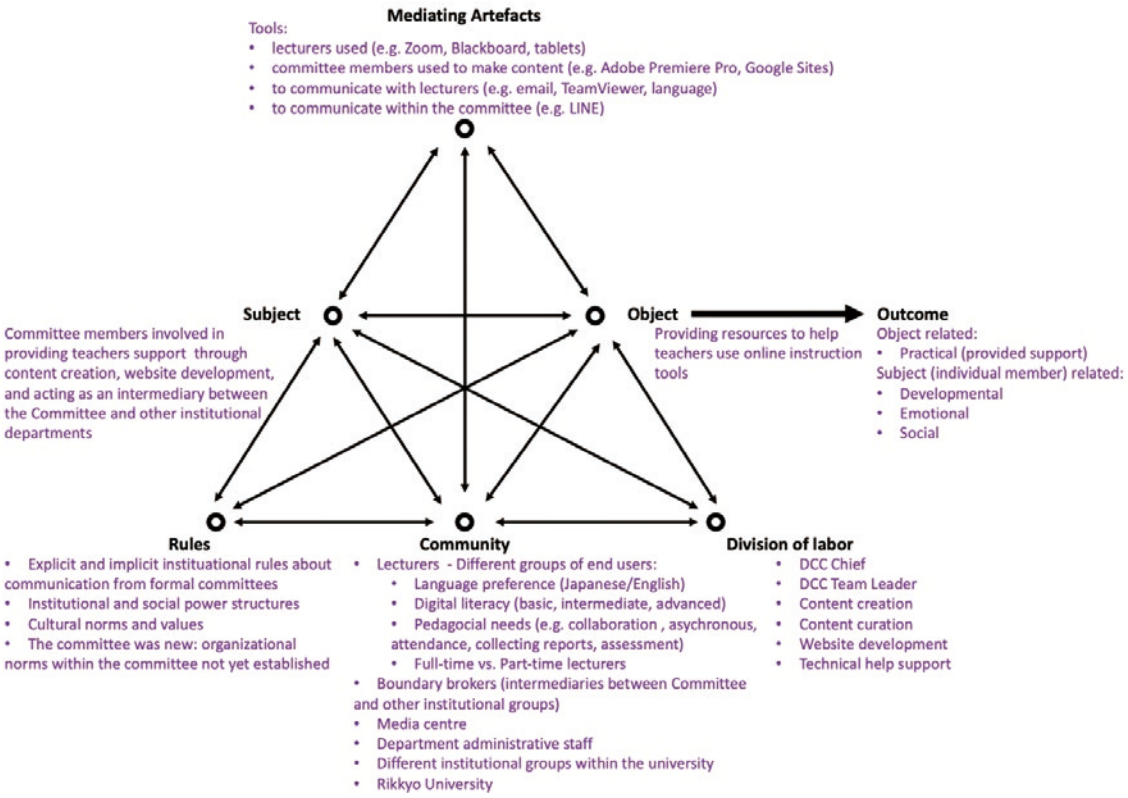


Figure 3. The Digital Communication Committee activity system elements

Object

Committee members' actions were shaped by the object of providing resources to help teachers be able to teach their coursework online during the pandemic, which for health and safety reasons, compelled the university to stop face-to-face classes and offer online instruction. However, given the new formation of the committee the object changed with time and as rules and boundaries became apparent (discussed in more detail later). For instance, before officially forming the committee, members were involved with support in informal development meetings with teachers that explored different tools, and pedagogical methods to support online instruction. However, upon formation as a formal committee the object was limited to provision of support for university endorsed tools (i.e., Blackboard, Zoom, Google for Education – Classroom, Drive, Docs, Slides, Sheets).

Subject

The subjects included the committee members and the committee chief. The committee members were involved in the sourcing and creation of content and provision of technical support to faculty that needed help getting online. The committee chief acted as an intermediary between the DCC and other committees and departments within the university and exercised final decision-making authority regarding what would be displayed on the website and through email. Each subject brought with them their own pedagogical values, social and institutional expectations, and digital expertise that also shaped the activity. For instance, some committee members had already been using Blackboard extensively and thus were able to provide many resources and step-by-step tutorials on how to use Blackboard in relation to specific courses at Rikkyo. Others had experience using the Google for Education suite and were able to provide resources on that. Members had been

involved in informal groups supporting faculty before the formation of an official committee. Requests and feedback from these meetings initially informed early content creation and included innovative pedagogical methods such as collaborative online tasks and discussion forums. However, institutional expectations informed the chief's decisions on what could be published on the final website, which then altered the content created by the members. As such the social and institutional positioning of the subjects mutually informed each other serving to redefining the object and the outcome.

Tools (Mediating Artefacts)

The mediating artefacts in the DCC can be grouped into four categories:

Tools lecturers needed for instruction

This included technology to support synchronous teaching (e.g., Zoom, Google Documents, Slides), technology that supported classroom management (e.g., Blackboard, Google Classroom), and the devices that the teachers and their students used (i.e., tablets, smartphones, laptop computers, web browsers). These tools mediated DCC activity as it provided a basis for the content for the DCC website. Moreover, there were different types of technical support offered to the teacher depending on the tools that they had. For instance, a teacher was not able to access any of the online tools provided by the university because of blockers installed in her browser. The team leader interacted remotely with her to resolve this issue.

Tools that committee members used to make content

These included technology for editing such as Adobe Premiere Pro for videos, Audacity for sound clean-up, screen capture tools, Google Sites, and websites to source professional license-free images and content (e.g., Blackboard how to videos). Video creation and editing took on average 1 hour per video. This affected the speed at which the content could be uploaded to the website. The team overcame this by also curating content from official sources (e.g., Blackboard, Google for Education) and then removing the links as official videos addressing specific needs of the university were added (e.g., how to merge courses for eLearning).

Tools and artefacts used to communicate with lecturers (users)

These tools included Zoom for meetings, TeamViewer to interact with their computers remotely, email for official notices of the committee work, and as a place for teachers to contact when they had issues. Apart from the chief, the committee members are all native English speakers with intermediate to high levels of Japanese proficiency. However, the lecturers that the committee supported included Japanese teachers as well as other foreign language teachers for example French, Italian, Chinese, Spanish, Korea. The website was primarily in English, and this may have mediated the support for lecturers who were not confident with their English abilities.

Tools for communication within the committee (e.g., LINE, email, Zoom)

Email and the Japanese social networking platform, LINE, were used for simple and immediate communication; however there were limitations in terms of the depth we could discuss topics. Consequently, Zoom was used for synchronous meetings with the team members. This allowed faster decisions when we needed to clarify collective objectives. However, it was hard to coordinate a

time at which everyone could meet simultaneously.

LINE mediated the communication within the committee. As it was attached to our smartphones, we received notifications which compelled us to respond to comments right away. This affected our work–life balance as we were more likely to respond to LINE communication off-hours compared with email, which we might wait on until the following business day to respond. Furthermore, the informal nature of LINE seemed to affect communication as there were jokes and comments in the LINE group that might not have been in an email thread.

Community

The community in which the DCC operated consisted of several groups of individuals that affected the DCC activity.

The largest group were the lecturers. These were the end users of the DCC, and their characteristics influenced the object and outcome of the DCC. These characteristics included the following:

- language preferences (Japanese/English/other)
- digital literacy (basic, intermediate, advanced)
- pedagogical needs
- full-time vs. part-time lecturers

These different characteristics affected communication and the support provided for example, language could affect the type of tools they had (e.g., English/Japanese/other language operating systems), digital literacy affected the type of technical support they needed, and pedagogical needs informed the type of content they wanted. Teachers wanted information on how to teach in the way they taught in their face-to-face classes. Those that used collaborative active learning methods wanted information on how to do the same online, while others who preferred teacher-fronted lecture-based methodologies sought technology and systems to allow them to do the same. Finally, lecturers could be divided into full- and part-time lecturers. Many of the support issues that we received were from part-time lecturers. Often part-time lecturers are working in several different universities with different systems. They may only have two classes with the university so many are not as familiar with the tools of a particular university compared with the full-time staff. Moreover, they sought to minimize the new technology that they had to learn. For this reason, some indicated that they preferred Google Classroom over Blackboard because they could easily create new classroom spaces for classes in the other universities they worked at but not Blackboard, which required institutional support.

The university media centre provided extensive technical support; however, it was all in Japanese. Many foreign language lecturers did know about the support they could get due to limited Japanese (e.g., they might not read the Japanese emails sent out by the media centre). The DCC committee connected the lecturers to the media centre by linking the DCC website to the media centre and the specific page where lecturers could go for additional support for online teaching.

Other departments and the Rikkyo university community also affected the activity of the DCC. Although within the foreign language department, there are different pedagogical methodologies, for the most part, lecturers sought communicative approaches due to the nature of our subject matter. However, across the university there was a wide range of subjects that did not use similar methods.

As such (as discussed in more detail in the rule section below) our "official" activities and communication were constrained by what was deemed appropriate for the overall community.

Division of labour

DCC members performed different roles within the committee (Table 3). The DCC chief was a tenured professor and acted as an intermediary between the committee and other institutional groups. As the chief, they interpreted the institutional needs, rules, and expectations for formal communication from the committee and thus exercised final authority of what content could be included on the website and how the committee was positioned in emails.

The DCC team leader performed all the same tasks as the committee members and additionally acted as an intermediary between the chief and the regular members. This entailed presenting the result of discussions collectively made within the committee to the chief. This role relieved the chief of being involved in all discussions and helped save the face of committee members during discussions as they would be able to express opinions and concerns freely.

All committee members except the chief were responsible for content creation (making new content), content curation (sourcing relevant, appropriate content), website development, and technical help support (responding to lecturers' requests by email, Zoom, or in person).

Table 3
Roles within the committee

Role	Description
DCC Chief	● Intermediary between the committee and other institutional groups. Final decision-making authority of content published on formal lines of communication.
DCC Team Leader	● In addition to all the duties of a committee member, the team leader acted as an intermediary between the committee members and the chief.
Committee member	● Responsible for content creation, curation, website development, and technical help support.

Rules

Explicit and implicit institution rules affected the activity. Once formed as an "official" committee, the content that we could prepare to provide information on the official communication channels (e.g., DCC website, email) was limited to technologies that were officially endorsed by the university (e.g., Zoom, Blackboard, Google for Education). Explicit and implicit expectations about how the technology was presented also informed content. For instance, as an official committee, members worked harder to make content better using Rikkyo colors and spent longer editing to make sound and images flawless. Further, to avoid conflict or confusion, materials that were viewed as potentially controversial, such as those containing pedagogical methods, were not included. For example, how to create a discussion board on Blackboard was included, but how to implement online discussion boards in teaching was excluded.

The committee was newly formed so organizational norms were not yet established, and the tenured chief was also relatively new to the university and still learning about institutional expectations. As such, the rules and organizational flow of the committee was in an early development stage; nonetheless activity was informed by institutional and social power structures. For instance, implicit expectations for staff to create the content and resolve issues with minimal disruption to the

chief who was busy with other duties. Power differentials between committee members on limited-term contracts and the tenured chief affected communication. Members were careful in their wording about conflicting views. Decisions that the chief made were final and accepted without question.

Outcome

Outcomes are grouped into two categories: object and subject related.

Object related

A positive object-related outcome is that the work of the activity provided practical support to the lecturers during the first stage of the pandemic. A negative outcome was that limitations on the content of the website to only include technical applications meant that once lecturers knew how to do those things, they did not need to use the resources and the value of the committee declined until it dissolved.

Subject related

Object-related outcomes informed the individual subject-related outcomes and can be grouped into three categories: developmental, emotional, and social.

Developmental

Committee members felt that their skills developed while engaging in the committee work. They learned from each other the different ways to teach and manage courses online. Moreover, the experience of interacting with the different institutional groups gave them insight into the different users and concerns related to faculty development of digital skills in the Japanese university context. Technical skills improved as members learned how to make the institutional website and edit videos and content. Finally conducting the research for this study informed self-reflection and transformative learning.

Emotional

Committee members reported both positive and negative emotional outcomes. Providing faculty support and obtaining appreciation from those who used the services was motivating and encouraging. Being trusted by the department to develop these resources was empowering and members were thrilled to be recognized for the skills they were passionate about that were separate from their language teaching duties. However, the limitations of the activity, derived from what could be shared officially, led to frustration about wanting to do more but being constrained. Additionally, members expressed disappointment when material with ideas on how to use the technology (e.g., ways to use discussion forums) was deleted from the site, as well as collective sadness that the committee was no longer recognized.

Social

Positive social outcomes included developing stronger professional relationships with the committee members and those we helped, recognition from other lecturers and the university Zenkari newsletter about our efforts (Zenkari 48, p. 5). Negative social outcomes were related to effects on our work–life balance and private lives. Troubleshooting concerns during the pandemic

while teachers were stressed created an urgency to respond to requests and LINE messages immediately. Many hours were spent creating the videos and content for the website during typically non working hours.

Contradictions

Members identified several contractions between the elements. The most prominent contradiction existed in the subject–object–rule–community relationship. As a newly formed committee, the subjects and different groups in the community had different views of the expected object of the activity. Before the committee was officially formed, members were providing informal support on a variety of tools (e.g., FlipGrid, meetings discussing ways to use Zoom breakout rooms); hence, initially members viewed the object with long term potential to continuously support the faculty development of digital skills. However, during the activity, institutional regulations placed on content viewed as controversial for different community groups clarified the perceived object to be more of an English help desk for officially endorsed technology. Moreover, there was a tension between the different pedagogical approaches between the committee members and the chief who acted as a boundary broker between the different institutional groups. Committee members wanted to send out a survey to the lecturers who could identify their critical needs, but this idea was rejected. Finally, the removal of resources that were created in response to questions from lecturers caused some confusion about the audience. This contradiction was later resolved by members through deployment of different communication channels: informal support versus formal communication.

Another tension experienced by members was in the subject–tool–object relationship. Members found that the type of tool mediated their communication. Initially communication was over email, which then changed to a LINE group. The first LINE group contained all members including the chief. However, there was some discomfort about having private conversations (including jokes and face-threatening acts such as questions and concerns) as we realized that our activity was “exposed” to the chief, which in turn constrained our dialogue. This discourse would typically not be conducted over email; however, the informal nature of LINE induced familial conversation despite status differentials. Later the team leader created a sub-LINE group for all members but the chief to reduce this tension.

Discussion

Through applying the CHAT framework, the analysis above describes the sociocultural relationship between the DCC members and the various university communities with which they interacted; as shown above, the decisions that the subject (the DCC) made were constrained by the fluid nature of its object. As noted above, the atypical nature of ERT has created a paucity of studies devoted to describing how PDNs evolve in crisis situations. Whereas research on ERT defines such situations in terms of flexibility, our findings suggest that how flexible an ERT situation actually is depends on the point from which it is viewed. Figure 3 above makes explicit the relationship between the activity system of the (DCC) PDN and the larger PDN with which it attempted to influence the object.

As Figure 3 indicates, as PDNs are made up of members who are themselves also members of the broader institution, the tensions between the DCC and other PDNs become more explicit. The exact nature of this tension is subtle and difficult to describe: The DCC operated from the perspective of a peer-to-peer or bottom-up PDN; the COVID-19 pandemic was poorly understood and, in the

absence of immediately forthcoming information, the committee members felt empowered to take matters into their own hands. In contrast, the other academic units on campus—in particular the administrative units associated with the language center—appeared to be operating according to the logic of a regulated or top-down PDN. In this way, our findings suggest that the relative fluidity of the object, that is, the objectives that the DCC intended to achieve, was evaluated as something far more rigid in the activity systems of other PDNs. With this in mind, we will venture two important features of our analysis.

On the Use of Tools

To begin with, Fucoloro (2012) suggests that two major themes in the choice of social network for developing a PDN are "community" (or camaraderie) and "convenience". As noted, the majority of communication within the committee was mediated via LINE, communication tended to be less formal than emails. The upside of this informal, asynchronous communication style was that it allowed for group decisions to be made rapidly by the committee, without the preamble and social niceties common in email chains. It also allowed previously discussed topics to be easily searched and viewed in their chronological context within the LINE app—something that is more difficult in typical email threads, which have a tendency to split into multiple conversations and become increasingly difficult to read (Sobotta, 2016).

Conversely, the use of LINE to mediate conversations in some ways negatively influenced the subjects' relationship with the community. The more informal nature of the medium, in conjunction with the installation of the software on the subjects' personal devices, meant that the line between work life and personal life became ill-defined. This, in the context of emergency remote work, which further reduced distinctions between private and the professional spheres (Sandoval-Reyes et al., 2021), subjects felt significant stress when interacting through this medium. As a result, as the committee evolved, the LINE group was used much less frequently, with some members turning off notifications for the group on their personal devices. This is consistent with Macià and García's (2016, p. 300) findings that virtual communities suffer from a "gradual lack of engagement…which results in reduced user participation." and lends further support to Lantz-Andersson and associates' (2018) findings that these communities interfere with their members' private lives, overwhelming members with an overabundance of information (Davis, 2015).

On Rules and Community

Relatedly, the unique constraints intrinsic to the tools that the DCC used figured into the larger dynamic of the community itself. Our findings differ from previous research (Lantz-Andersson et al., 2018; Macià & García, 2016) in that we do not classify the DCC as either an informal or formal network, but rather as a single PDN that transitioned from the first state to the second. CHAT is the ideal lens for analysing the process by which this occurred. Tensions experienced by the subjects can be explained by a mismatch of objects between the original team and the other institutional communities with which they were now interfacing. The DCC's activity system did not have the historical cultural context of the larger institution. When it interacted with these new groups, their historical discourse and decisions (and subsequent implicit cultural rules) conflicted with the DCC's. This conflict was resolved by submitting to the academic hierarchy. The DCC chief deferred to the cultural context of the institution, and thus, the DCC itself was changed. This interaction of activity

systems is not sufficiently explained by the form of CHAT discussed above; nonetheless (Cole & Engeström, 1993) the third generation of CHAT allows for the analysis of interacting systems.

The implicit rules of the activity system were mediated by the ERT context and the communities with which the DCC interacted. Initially, the emergency nature of the task and the ad hoc formation of the group meant that the implicit rules for knowledge gathering were loose. Typically, PDNs gather informal knowledge from members who in turn gained knowledge using the "interactive" model (Clarke & Hollingsworth, 2002; Macià & García, 2016). In this model, teachers' professional development is facilitated by "external sources of information lead[ing] to new experiences in the classroom which… can lead to new insights" (Macià & García, 2016, p. 292). In the context of ERT, the subjects were unable to rely on their interactively produced knowledge. As none of them had extensive experiences with remote teaching, knowledge provided for the PDN was necessarily sought after specifically in order to fulfill the pressing need.

Wenger (2009) argues that knowledge sharing requires "relationships of trust" developed through "mutual engagement". In online spaces, gaining trust without the benefits of face-to-face interactions is fraught with difficulties (Ridings et al., 2002; Young & Tseng, 2008). If this is the case, then perhaps this potential lack of trust, in combination with the lack of a more traditionally gathered informal knowledge base, led to the change in explicit rules about what the committee could share online on its website.

Implications

When preparing to develop a PDN, instructors need to be aware of the complexities of the activity system that they are engaging with. CHAT analysis may help them contemplate the sociocultural elements that will shape the endeavour. Instructors may wish to consider making a formal statement of the object of the PDN, to avoid any ambiguity or tensions thus derived. The choice of tools mediating communication between members of the PDN will require careful reflection. Ease and speed of use will need to be reconciled with a consideration of the divide between the members' personal and professional lives. The historical cultural rules of the institution in which the PDN is built will need to be carefully considered. Any tension between the object of the PDN and the culture of the institution will need to be resolved, ideally before the PDN is formed. Finally, the cultivation of community trust should be considered as trust is one of the "key enablers of knowledge sharing in online communities" (Booth, 2012). For example, as Booth (2012) reported, smaller closed communities engender trust among members.

Limitations and Future Research

Conducting this study using CHAT and the PDN activity system as the "primary unit of analysis" (Figure 1) provided a lens that helped reveal the sociocultural issues that emerge during an organically forming intercultural PDN with the context of a crisis. However, a clear finding of this study is that professional development activities occur across different institutional systems that influence and shape each other. As such, the analysis of one activity system from the perspective of only the teachers whose objects were mediated by other systems that they were not aware of provides a limited view of the true complexity of forming a successful PDN within an intercultural context. Future research that examines other interacting systems in more detail, including, for example, the clients they serve, institutional faculty, and administrative support staff through the lens

of third or fourth generation activity theory (Engeström, 2001, 2009), would provide a more holistic view of the mediated activity.

Conclusion

With the proliferation of ICT tools and proficiencies amongst instructors since the ERT era, many new teachers now have the ability to establish their own PDNs. However, the development and maintenance of these networks can be fraught with complex issues and tensions. Using CHAT analysis to study the sociocultural elements of the formation of these networks and their maintenance will allow teachers to more effectively create their own PDNs and enjoy the many researched benefits of sharing knowledge with their peers. With sufficient preparation and careful planning, these PDNs could be successfully integrated into their larger institutional context to provide knowledge findings and share it institution wide.

References

Alemdar, M., & Rosen, J. H. (2011, June). Introducing K-12 teachers to LEGO Mindstorm Robotics through a collaborative online professional development course. In 2011 *ASEE Annual Conference & Exposition* (pp. 22-959).

Bligh, B., & Flood, M. (2017). Activity theory in empirical higher education research: Choices, uses and values. *Tertiary Education and Management*, *23*(2), 125-152.

Booth, S. E. (2012). Cultivating knowledge sharing and trust in online communities for educators. *Journal of Educational Computing Research*, *47*(1), 1-31.

Burns Thomas, A. (2004). *Inquiry within and against: The experiences of four new teachers in an urban professional development network*. (University of Pennsylvania Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3138081)

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(8), 947-967. https://doi.org/10.1016/S0742-051X(02)00053-7

Close-up: COVID Universities, 60% of classes distance learning only, (2020, June 7). *Mainichi Shimbum*. https://mainichi.jp/articles/20200607/ddm/003/040/037000c

Cole, M., & Engeström, Y. (1993). A cultural-historical approach to distributed cognition. *Distributed Cognitions: Psychological and Educational Considerations*, 1-46.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3-21.

Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., Magni, P., & Lam, S. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses. *Journal of Applied Learning & Teaching*, *3*(1), 1-20.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Crow, M. M., & Shangraw, R. F. Jr (2016). Revisiting "public administration as a design science" for the twenty-first century public university. *Public Administration Review*, *76*(5), 762-763.

Cucinotta, D., & Vanelli, M. (2020). WHO declares COVID-19 a pandemic. *Acta Bio Medica: Atenei Parmensis*, *91*(1), 157-160.

Cutts, Q., Robertson, J., Donaldson, P., & O'Donnell, L. (2017). An evaluation of a professional learning network for computer science teachers. *Computer Science Education, 27*(1), 30-53.

Davis, K. (2015). Teachers' perceptions of Twitter for professional development. *Disability and Rehabilitation*, *37*(17), 1551-1558.

Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*, *43*(7), 960-974.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work, 14*(1), 133-156.

Engeström, Y. (2009). The future of activity theory: A rough draft. In A. Sannino, H. Daniels, & K. Gutierrez (Eds.), *Learning and expanding with activity theory* (pp. 303-328). Cambridge University Press.

Engeström, Y., & Sannino, A. (2021). From mediated actions to heterogenous coalitions: Four generations of activity-theoretical studies of work and learning. *Mind, Culture, and Activity*, *28*(1), 4–23.

Fucoloro, D. J. (2012). *Educators' perceptions and reported behaviors associated with participation in informal, online professional development networks* (Doctoral dissertation, Saint Louis University).

Funamori, M. (2017). The issues Japanese higher education face in the digital age—Are Japanese universities to blame for the slow progress towards an information-based society? *International Journal of Institutional Research and Management*, *1*(1), 37–51.

Gacs, A., Goertler, S., & Spasova, S. (2020). Planned online language education versus crisis-prompted online language teaching: Lessons for the future. *Foreign Language Annals*, *53*(2), 380–392.

Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social Problems*, *12*(4), 436–445. https://doi.org/10.2307/798843

Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., & Bond, M. A. (2020). *The difference between emergency remote teaching and online learning*.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Inoue 井上兼生. (2021). 「主体的・対話的で深い学び」を目指すオンライン授業の考察：コロナ禍での 「総合的な学習の時間の指導法」等の授業実践を踏まえて. *聖学院大学論叢*, *33*（第1・2号合併）, 123-137.

Lantz-Andersson, A., Lundin, M., & Selwyn, N. (2018). Twenty years of online teacher communities: A systematic review of formally-organized and informally-developed professional learning groups. *Teaching and Teacher Education*, *75*, 302–315. https://doi.org/10.1016/j.tate.2018.07.008

Lavicza, Z., Hohenwarter, M., Jones, K., Lu, A., & Dawes, M. (2010). Establishing a professional development network around dynamic mathematics software in England. *International Journal for Technology in Mathematics Education, 17*(4), 177-182.

Lawrence, A. T., & Dubetz, N. (2001). An urban collaboration: Improving student learning through a professional development network. *Action in Teacher Education, 22*(4), 1-14.

Leont'ev, A. N. (1978). *Activity, consciousness, and personality.* (Englewood Cliffs, Prentice-Hall).

Macià, M., & García, I. (2016). Informal online communities and networks as a source of teacher professional development: A review. *Teaching and Teacher Education*, *55*, 291–307.

Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland Journal of Higher Education*, *9*(3).

Maier, R., & Schmidt, A. (2015). Explaining organizational knowledge creation with a knowledge

maturing model. *Knowledge Management Research and Practice*, *13*(4), 361–381. https://doi.org/10.1057/kmrp.2013.56

Marinoni, G., Van't Land, H., & Jensen, T. (2020). The impact of Covid-19 on higher education around the world. *IAU Global Survey Report*, *23*.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook. 3rd.* Sage.

Ridings, C. M., Gefen, D., & Arinze, B. (2002). Some antecedents and effects of trust in virtual communities. *The Journal of Strategic Information Systems*, *11*(3-4), 271-295.

Rieckhoff, B. S., & Larsen, C. (2012). The impact of a professional development network on leadership development and school improvement goals. *School-University Partnerships, 5*(1), 57-73.

Roth, W.-M., & Lee, Y.-J. (2007). "Vygotsky's neglected legacy": Cultural-historical activity theory. *Review of Educational Research*, *77*(2), 186–232.

Sandoval-Reyes, J., Idrovo-Carlier, S., & Duque-Oliva, E. J. (2021). Remote work, work stress, and work–life during pandemic times: A Latin America situation. *International Journal of Environmental Research and Public Health*, *18*(13). https://doi.org/10.3390/ijerph18137069

Shangraw, R. F. Jr, & Crow, M. M. (1998). Public administration as a design science. *International Journal of Public Administration*, *21*(6-8), 1059-1077.

Siemens, G., & Matheos, K. (2010). Systemic changes in higher education. *In Education, 16*(1).

Sobotta, N. (2016). Why forwarded email threads are hard to read: The email format as an antecedent of email overload. *Communications of the Association for Information Systems*, *39*(1), 2.

Thomas, A. B. (2007). Supporting new visions for social justice teaching: The potential for professional development networks. *Penn GSE Perspectives on Urban Education, 5*(1), n1.

Trust, T. (2015). Deconstructing an online community of practice: Teachers' actions in the Edmodo math subject community. *Journal of Digital Learning in Teacher Education*, *31*(2), 73–81.

Trust, T. (2017). Using cultural historical activity theory to examine how teachers seek and share knowledge in a peer-to-peer professional development network. *Australasian Journal of Educational Technology*, *33*(1), 98–113.

Tu, C. H., Sujo-Montes, L., Yen, C. J., Chan, J. Y., & Blocher, M. (2012). The integration of personal learning environments & open network learning environments. *TechTrends*, *56*(3), 13-19.

Warschauer, M. (2020). Networking the Nile: Technology and professional development in Egypt. *In Technology and English Studie*s (pp. 163-172). Routledge.

Watkins, K. E., & Marsick, V. J. (1992). Towards a theory of informal and incidental learning in organizations. *International Journal of Lifelong Education*, *11*(4), 287–300. https://doi.org/10.1080/0260137920110403

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research*, *79*(2), 702–739.

Wenger, E. (2009). Communities of practice: The key to knowledge strategy. In *Knowledge and communities* (pp. 3-20). Routledge.

Young, M. L., & Tseng, F. C. (2008). Interplay between physical and virtual settings for online interpersonal trust formation in knowledge-sharing practice. *Cyberpsychology and behavior*, *11*(1), 55-64.

Zenkari – The Road to Online Classes [全カリ・オンライン授業への道]. (2020/09/24). *Zenkari Newsletter, 48*, 5. https://www.rikkyo.ac.jp/education/system/general/overview/publication/mknpps000001x3bw-att/letter48.pdf

【Research Article】

Teacher Agency in Teaching Debate: A Sociomaterial Perspective

Masakazu Mishima & Yuka Yamamoto

Abstract

Teacher agency has been well recognized as a critical theoretical notion in explaining teacher decision making and teaching practice. This institution-based case study examined one instructor's course development process and teaching practice in a pilot debate class based on fieldnotes, instructor interviews, and questionnaire analysis. The results showed that various decisions that the instructor made from the development to delivery of the pilot course were in line with the intersection of multiple elements, including the instructor's beliefs, professional experience, and various pedagogical conditions such as class size, students' level, institutional demands, and time constraints. The study demonstrated that while teacher agency functioned as an essential mediator in teacher decision-making, the instructor's various decisions were nonlinear and adaptive. Based on the findings, the researchers discuss teacher agency from a sociomaterial perspective, which focuses on social and material resources deeply embedded in the pedagogical environment. The study addresses a potential application of the sociomaterial approach to teacher agency studies and sheds light on how teacher agency can be reconceptualized.

Keywords: Teacher Agency, Sociomaterial appraoch, English language education

Introduction

The present study is a part of a three-staged longitudinal research project that examines teacher agency in curriculum development processes. Phase one of the research reported students' responses to a pilot English debate class (see Mishima & Yamamoto, 2020). This article presents results from the phase two study, which attempted to examine teacher agency in relation to the pilot English debate course taught at an urban university in Japan.

Teacher agency

The central conceptual notion of the present study is teacher agency—a highly elusive concept that has called for various debates and contentions over the years (Emirbayer & Mische, 1998; Priestley et al., 2015a). In education, teacher agency is often defined as their capacity to act (Priestley et al., 2015a). This particular definition of teacher agency has been the locus of a debate among scholars and philosophers, in which teacher agency is often assumed to be an attribute of the individual insomuch as other arguably more widely researched psychological constructs such as teacher motivation, beliefs, identity, and emotions (see Mercer & Kostoulas, 2018, for collective work on language teacher psychology). Overviewing the research landscape on teacher agency, studies specific to English language teaching are still scarce (White, 2018) though with some notable exceptions. For instance, one thread of teacher agency research focused on examining the relationship between teacher agency and identity (Ruohotie-Lythy & Moate, 2016; Kayi-Aydar, 2015). Some pursued the role of teacher agency in classroom settings (Kitade, 2015; White, 2018), while others investigated teacher agency concerning language planning and policy (Ng & Boucher-Yip,

2016).

This study adds to the growing body of research on teacher agency to explore it from an ecological perspective. While some scholars advocate the ecological view of English as a Foreign Language (EFL) contexts (e.g., Priestley et al., 2015b), ecological orientations in teacher agency studies in EFL contexts remain lacking. In addition, an investigation into teacher agency concerning teaching English debate at the tertiary level in Japan is, to the best of our knowledge, currently nonexistent.

To explore teacher agency in a situated manner, we adopted a sociomaterial approach, which defines our theoretical positionality in our investigation.

Sociomaterial approach

The sociomaterial approach has been used predominantly in work organization research, which seeks to understand *materiality* and its role in shaping various professional organizations (Fenwick & Landri, 2012). The approach is characterized by its fundamental focus on understanding phenomena/activities within social and material conditions (Fenwick & Landri, 2012). In this view, the sociomaterial approach treats all subjects of interest, including humans and nonhumans, on an equal plane. In addition, a phenomenon under investigation is assumed to emerge from the collective interactions of humans and non humans. Based on these theoretical tenets, we view agency as an emergent state that resides in the particular/temporary relationships between humans and nonhuman objects in a specific moment; it is unstable and ever-changing. We also presuppose that teacher agency is not an individual's exclusive attribute but distributed across social and material conditions that envelop teachers' decisions and actions.

Methods

Context

The current study was conducted in the process of developing a mandatory English course for first-year students, English Debate, at an urban university in Tokyo. The course aims to help students understand the nature and structures of debate in English and develop critical thinking and research skills. Students must analyze and formulate arguments logically on issues from multiple perspectives and respond to questions.

Participant

Lisa (pseudonym), the instructor participant, was a Japanese and English bilingual. She lived in the United States for five years and in Singapore for four years. She holds two M.A.s and a Ph.D. in Lowercase and is an experienced teacher trainer, curriculum developer, and textbook writer. In her 20 years of teaching experience, Lisa has designed and taught various English language classes. However, she had no prior experience with teaching debate.

Data collection

Data were collected in Spring 2019 from the pilot debate class instructor, who was also in charge

of developing the debate course curriculum. The researchers first administered an instructor questionnaire (Appendix I) to elicit the instructor's basic demographic information and her opinions and beliefs about teaching the pilot debate class. The researchers also observed pilot class lessons, took field notes, and collected all class materials from the instructor.

In addition to the above, two instructor interviews were conducted (Appendix II). The interviews were semi-structured to allow the participant to dialogically share her experience and perceptions about teaching the pilot debate class. All predetermined questions were open-ended, and follow-up questions were asked for clarification and elaboration. The interviews lasted for two hours and were recorded on an IC recorder. The recorded interview data were then transcribed for analysis.

Data analysis

To explore teacher agency from a sociomaterial perspective, the researchers analyzed all data focusing on Lisa's decisions and actions in teaching the pilot debate course. The interview data, field notes, and all class materials were open-coded. All coded data were then thematically labeled and organized. After which, the researchers identified illustrative excerpts from the interview data and the other data sources to present notable findings. Throughout the analysis, the researchers consulted each other to ensure the trustworthiness and truthfulness of the emerged codes, themes, and findings.

Results

The analysis identified three major themes and their associated subthemes as a nexus of sociomaterial conditions within which Lisa's decisions and actions are embedded. Table 1 presents the three major thematic categories of sociomaterial conditions: a. instructor attributes, b. teaching environment, and c. time constraint.

Table 1
Descriptors of three thematic categories of sociomaterial conditions

Sociomaterial Categories	Descriptors
Instructor Attributes	The category refers to the instructor participant's unique qualities defined by her teaching experience and beliefs.
Teaching Environment	The category refers to specific sociomaterial constraints as the instructor participant perceives within the instructional context.
Time Constraint	The category refers to a physical constraint placed by the availability of time as perceived by the instructor participant.

To demonstrate how her agentive decisions and actions emerged in consort with the sociomaterial conditions, select excerpts are presented in the sections that follow.

Instructor attributes

Teaching experience. Lisa encountered many challenges as she had never taught debate to English language learners. For instance, in the instructor questionnaire, Lisa mentioned difficulty in imagining the structure and timing of a debate for teaching purposes:

Excerpt 1

I never had the experience of teaching a debate class. So it was hard to imagine the structure and timing [of debate].

Lisa's lack of experience with teaching debate coincided with her perceived difficulty in designing an appropriate debate format to teach. She first searched and watched video recordings of debates on the Internet to remedy the issue.

What follows in Excerpt 2 shows that Lisa's decision to search and watch video recordings is a product of her lack of teaching experience and foregrounded forms of materiality: videos on the Internet and textbooks. Those resources provided the primary material conditions wherein Lisa's decision was made possible.

Excerpt 2

I first tried to design the end product of the debate by watching several video recordings on American elementary school students having a debate. I also went through several debate textbooks written for Japanese elementary school students. I looked through elementary school materials to get the overall structure [of a debate].

Teacher's beliefs. Teachers' beliefs are often discussed with teachers' decision-making and teacher agency (Biesta et al., 2015). Excerpt 3 below shows that Lisa firmly believes in developing students' curiosity and helping them to become more "inquisitive." However, as discussed elsewhere (Mishima, 2018), teachers' beliefs need to be translated into tangible forms of teaching practice in alignment with available material conditions.

Lisa mentions a means to give a practicable form to her beliefs, "by searching for information, especially on the Internet..." This particular account demonstrates that the availability of the Internet had been assumed in her pilot debate class, wherein Lisa repeatedly encouraged her students to conduct thorough research on the Internet to develop their arguments and find evidence to perform a good debate (fieldnotes).

Excerpt 3

My mission as a language educator is not simply to improve my students' English language skills. I want my students to take the initiative in their learning. My role as a teacher is facilitating and creating an environment where students feel comfortable speaking and asking questions. What I constantly feel lacking among students is their curiosity. I want them to be more inquisitive and question things. I want them to be curious about different cultures, values, beliefs, and practices. By searching for information, especially on the Internet, they can access multiple perspectives beyond what is provided in textbooks (Instructor Questionnaire).

Teaching environment

Instructional decisions and actions can never be separated from social conditions embedded in the specific teaching context (e.g., institutional culture, rule of conduct, and expectations). We collectively refer to this type of contextual condition as the teaching environment based on the social network theory proposed by Wellman (1988). The teaching environment includes three major conditional elements across two different levels of context. Class size and students' level were identified at the micro-level of context (i.e., the classroom), and institutional demand was identified at the meso-level of context (i.e., the language program). We found that these contextual backdrops encapsulated Lisa's instructional decisions.

Class size. Class size is an essential factor in planning and teaching a lesson (Russell & Curtis, 2013). Class roasters were all tentative at the research site before the first lesson. In other words,

Lisa was not sure how many students would sign up for her class until after she began teaching. Excerpt 4 indicates the uncertainty of her class size and how it might affect the team and time arrangements of debates in planning her lessons.

Excerpt 4

The main issue was the number of students. I wasn't sure how many students would sign up for the course. The team and time arrangements [for debates] would differ depending on that. In the worst case, I would end up with only one or two students in class. But in the end, 25 students signed up for the course (Follow-up interview).

Students' level. Planning lessons according to student language proficiency is vital for quality language education. However, students' level was another element that remained uncertain until Lisa started teaching her class. To safeguard against the unknown element, she devised strategic counter-measures. For example, Excerpt 5 presents two different ways Lisa implemented to prepare her lessons. One was to teach the debate skills in small chunks to help potentially weak students learn the necessary skills. The other was randomly changing students' groupings to balance students' differing levels.

Excerpt 5

Because this pilot course was conducted in an elective class, I didn't have a clear idea of the students' level. To solve the issue, I decided to introduce the debate skills step by step. So even less proficient students can learn. Also, I changed the groups randomly at the beginning of every lesson to ensure students were divided into different members and levels (Instructor Questionnaire).

In addition to the above, Excerpt 6 indicates that Lisa intentionally chose an introductory textbook with the premise that modifying the textbook for higher-level students would be much easier.

Excerpt 6

I wasn't quite sure about the level of my students, so I looked for a basic-level textbook. It's much easier to choose a more accessible textbook and adapt it by adding more challenging materials (Follow-up interview).

Even after Lisa began teaching the class, her reflective and adaptive decisions manifested differently. Excerpt 7 presents a notable example in which she made substantial changes to the initially prepared lecture slides by adding more complex tasks to meet the needs of varying levels of students in her classroom.

Excerpt 7

In reality, students' levels varied. So I had to change the PPT slides and add more challenging tasks for higher-level students (Follow-up interviews).

Institutional demand. Whether teaching or developing a curriculum, teachers' decisions are never free from the pedagogical context and its various influences (Owston, 2007). The debate class Lisa designed was to be delivered as one of the first-year mandatory English courses within the unified English curriculum. The curriculum serves approximately 4000 first-year students annually at the university (fieldnotes). The class size was 25 or fewer in many mandatory courses, including the debate class. The prospective number of instructors assigned to teach at least a section of the debate course was thus proportionately large. Given these contextual backgrounds, Lisa was highly conscious of how she designed the debate course and how it might impact other instructors once the course was officially launched. Accordingly, Lisa aimed to develop a course that would be simple and easy for instructors to understand. Consider Excerpt 8:

Excerpt 8

Because the course is to be offered in the unified curriculum, it had to be as simple as possible so that other teachers could also understand and follow the structure. The unified curriculum means there is a set of fixed goals and objectives. Within the framework, teachers can plan their lessons. I'm in a position to design the course, so I need to show them a model. Flexibility in the syllabus is suitable for experienced teachers, but most instructors have no experience teaching a debate course. Also, part-time instructors teach in other universities and don't have much time to plan their lessons. So what I always had in mind was to create a course that is simple and easy to understand (Follow-up interview).

Time constraints

Teachers face many constraints in planning and teaching lessons. One of such constraints is time availability, which is crucial in understanding teachers' decisions (Teig et al., 2019). At the initial stage of planning her lessons, Lisa searched for reading materials that could serve as an introduction to topics to be debated in class. However, finding suitable reading materials on a wide range of topics was not easy, primarily due to her limited time developing, preparing, and teaching the debate class. Excerpt 9 suggests that Lisa's perceived time limitations were an essential factor in searching for ready-made materials to teach debate skills and structure with minimal modifications.

Excerpt 9

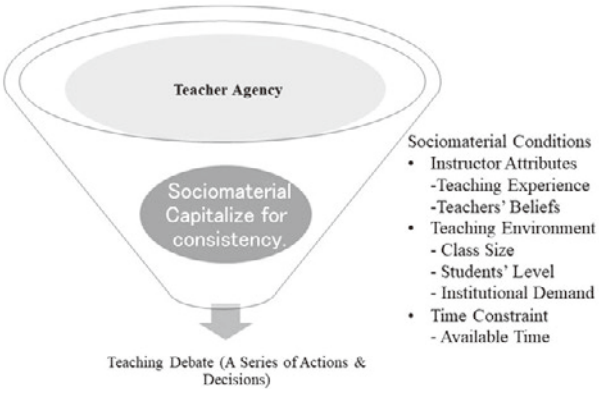
Due to the time limitations, I searched for a published textbook with familiar reading topics for Japanese university students, such as having schools on Saturdays and increasing consumption tax in Japan. It was a critical thinking textbook. So I adopted reading materials from it, and then created PPT slides to introduce the debate skills and structure. They are easy to make and modify later (Follow-up interview).

Discussion and Implications

Our results show that Lisa's various decisions were inseparably linked to specific social and material conditions in her work environment. Figure 1 below presents the teacher agency's overall structure and the identified sociomaterial conditions as a summative overview of the results.

Figure 1

Teacher Agency and Sociomaterial Conditions



A critical finding in this study is that teacher agency as manifested in agentive decisions and actions is embedded in multifaceted social and material limitations perceived by the agent in the field. The various manifestations of teacher agency are highly interactive and dialogic in that there is no clear separation between the self and the environment. The finding is contrastive to the pervasive conceptualization and representation of teacher agency as an individual characteristic that teachers possess and act upon (see Bandura, 2001 for more discussion on the individualistic approach to teacher agency).

In our investigation, we adopted the sociomaterial approach under the assumption that whatever forms of teacher agency one might exercise, their resulting decisions and actions are socially and materially constrained; social and material contexts need to be integratively investigated in understanding teacher agency. The sociomaterial approach enabled us to treat an individual (i.e., teacher) and their perceived material and social conditions as mutually constitutive to the agent's decisions and actions.

As we highlighted in the results, Lisa's pedagogical decisions are entangled with the various forms of materiality, such as the Internet, videos, and textbooks. Giddens (1984) rightly argued that teacher agency is constrained by available resources such as classroom equipment. Physical constraints in our study extend to the availability of time within which the instructor planned, developed, and conducted debate lessons. Furthermore, Lisa's perceived social constraints are present in the forms of unknown class size and student level as well as institutional expectations. The finding corresponds to to that of Hanson's (2003) that school conditions and processes mediate teacher agency. Our study indicates, however, that social conditions seem to go beyond the mediators of teacher agency as they are embedded in teachers' actions. In other words, without the social and material constraints, Lisa's adaptive and interactive decision-making and actions cannot be fully explained. This point is notable as social conditions are often external factors to teachers' actions (Luttenberg et al., 2013).

The hallmark of teacher agency is found in the ontologically performative sphere wherein the agents actively seek and find ways to execute socially and materially possible decisions and actions Etelapelto et al., 2013 Pappa et al., 2019). We found Lisa a highly active agent, as exemplified in her adaptive actions in planning and preparing lessons. This type of heightened involvement in professional work marks agents as active, and it requires them to be aware of various constraints placed upon them at personal and institutional levels. The importance of teacher agency studies lies in its recognition of teachers as active agents for professional development, curriculum reforms, and improvement (Eteläpelto et al., 2013; Priestly et al., 2012). The line of studies pursues the role of the agent and its impact on teachers in various contexts under the premise that teachers with an active agency should bring about positive educational outcomes. However, agents are not free from material and social limitations; *active* does not readily mean that positive changes are possible.

Furthermore, Biesta et al. (2015) argue that what matters in teacher agency studies is to examine the *quality* of actors' engagement with contexts to act. That is why exploring social and material constraints as perceived by teachers is necessary. In addition, the focus of such studies needs to be on agents' adaptive decisions and actions in and with context rather than on agents in isolation or factors affecting them. In such endeavors, the sociomaterial approach helps us shift our view of agency from human to phenomena, which encompasses human and nonhuman actors (i.e., sociomateriality) as equally important research subjects.

Conclusion

This study examined one instructor's teacher agency teaching an EFL debate class in a Japanese higher education context. By adopting a sociomaterial approach, the study found that the instructor's teacher agency manifested in her pedagogical decisions and actions encapsulated by various social and material conditions. It demonstrated that the exercise of teacher agency is context-bound, interactive, and adaptive; the instructor's enacted decisions and actions are part of social and material constraints.

Given the ongoing COVID-19 epidemic, English language teaching in Japan is facing radical changes on all fronts. Amid these changes, the role of teachers is becoming ever more critical as the quality of instructors is central to quality education. Teacher agency is an essential area of research in language education to understand how teachers adapt and respond to the constantly changing landscape of social and material conditions in their respective pedagogical contexts. As this study has shown, an ontologically individualistic approach to examining teacher agency may well be insufficient to represent the complexity of the construct and how it is exercised in tandem with emerging sociomateriality. This point echoes the proposition forwarded by Emirbayer and Mische (1998) to reconceptualize agency:

> Neither rational choice theory, norm-based approaches, nor any of the other sociological perspectives extant today provide a fully adequate understanding of its significance and constituent features. Nor do such perspectives satisfactorily answer the question as to how agency interpenetrates with and impacts upon the temporal relational contexts of action.
>
> (p. 1012)

Finally, our study is limited in its replicability and generalizability, given our methodological choice. We purposefully adopted a qualitative research design to investigate teacher agency in *situ* and focused on one instructor participant. While it allowed us to closely analyze various qualitative data sources, the extent of social and material conditions we identified was most likely far from exhaustive. For example, social conditions may extend to the instructor's social relationships and interactions with her colleagues as collegial support or collaboration is commonplace in developing and teaching a new curriculum. Bringing other instructors into research would have added another layer of complexity to the study.

For future research, we believe it is imperative to include multiple agents and their surrounding social material conditions within the same instructional context to present a fuller picture of teacher agency at work.

References

Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching, 21* (6), 624-640, doi: 10.1080/13540602.2015.1044325

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1-26.

Emirbayer, M., & Mische, A. (1998). What is agency? *American Journal of Sociology, 103, 962-1023*. doi:10.1086/231294

Eteläpelto, A., Vähäsantanen, K., Hökkä, P., & Paloniemi, S. (2013). What is agency? Conceptualizing professional agency at work. *Educational Research Review*, *10*, 45-65.

Fenwick, T., & Landri, P. (2012). Materialities, textures and pedagogies: Socio-material assemblages

in education. *Pedagogy, Culture & Society*, *20*, 1-7, doi: 10.1080/14681366.2012.649421

Giddens, A. (1984). *The constitution of society*. Berkeley, CA: University of California Press.

Hanson, E. (2003). *Educational administration and organizational behavior* (5th ed.) Boston, MA: Allyn and Bacon.

Imants, J., & Van Veen, K. (2010). Teacher learning as workplace learning. In: P. Peterson, E. Baker, & B. McGraw (Eds.). *International Encyclopedia of Education (3rd edition, vol. 7)* (pp. 503-510). Amsterdam: Elsevier.

Kayi-Aydar, H. (2015). Teacher agency, positioning, and English language learners: Voices of pre-service classroom teachers. *Teaching and Teacher Education, 45,* 94-103.

Kitade, K. (2015). Second language teacher development through CALL practice: The emergence of teachers' agency. *CALICO Journal*, *32*, 396-425. doi: 10.1558/cj.v32i3.26637

Luttenberg, J, Imants, J, & Van Veen, K. (2013). Reform as an ongoing positioning process: The positioning of a teacher in the context of reform. *Teachers and Teaching: Theory and Practice*, *19*(3), 293-310.

Mercer, S., & Kostoulas, A. (2018). *Language Teacher Psychology*. Bristol, Blue Ridge Summit: Multilingual Matters.

Mishima, M. (2018). Teacher cognition and L2 writing feedback: A case of an experienced L2 writing instructor. *Writing & Pedagogy*, *10* (1), 93-127.

Mishima, M., & Yamamoto, Y. (2020). Japanese English learners' response to English debate: Insights from a pilot class. *JAAL in JACET Proceedings*, *2*, 15-22.

Ng, P. C., & Boucher-Yip, E. (Eds.). (2016). *Teacher agency and policy response in English language teaching*. New York, NY: Routledge.

Owston, R. (2007). Contextual factors that sustain innovative pedagogical practice using technology: An international study. *Journal of Educational Change*, *8*, 61-77. doi: 10.1007/s10833-006-9006-6

Pappa, S., Moate, J., Ruohotie-Lyhty, M., & Eteläpelto, A. (2019). Teacher agency within the Finnish CLIL context: Tensions and resources, *International Journal of Bilingual Education and Bilingualism, 22*(5), 593-613, doi: 10.1080/13670050.2017.1286292

Priestley, M., Biesta, G. J. J., & Robinson, S. (2015a). Teacher agency: What is it and why does it matter? In R. Kneyber & J. Evers (Eds.), *Flip the system: Changing education from the bottom up*. London: Routledge.

Priestley, M., Biesta, G. J. J., & Robinson, S. (2015b). *Teacher agency: an ecological approach.* London: Bloomsbury.

Priestley, M., Edwards, R., Priestley, A., & Miller, K. (2012). Teacher agency in curriculum making: Agents of change and spaces for manoeuvre. *Curriculum Inquiry*, *42*(2), 191-214

Ruohotie-Lythy, M. & Moate, J. (2016). Who and how? Preservice teachers as active agents developing professional identities. *Teaching and Teacher Education*, *55*, 318-327. doi:10.1016/j.tate.2016.01.022

Russell, V., & Curtis, W. (2013). Comparing a large- and small-scale online language course: An examination of teacher and learner perceptions. *Internet and Higher Education*, 16, 1-13. doi:10.1016/j.iheduc.2012.07.002

Teig, N., Scherer, R., & Nilsen, T. (2019). I know I can, but do I have the time? The role of teachers' self-efficacy and perceived time constraints in implementing cognitive-activation strategies in science. *Frontiers in Psychology*, 1697.

Wellman, B. (1988). Structural analysis: from method and metaphor to theory and substance. In B. Wellman & S. D. Berkowits (Eds.). *Social Structures: A Network Approach* (pp. 19-61).

Cambridge: Cambridge University Press.
White, C. J. (2018). Language teacher agency. In S. Mercer & A. Kostoulas. (Eds.). *Language Teacher Psychology* (pp.196-210). Bristol, Blue Ridge Summit: Multilingual Matters.

Appendix I
Instructor Questionnaire

Q1. What is your sex?
Male/Female
Q2. How long have you taught English at the university level?
Q3. Have you ever taught debate in English?
Yes/No
Q4. In your opinion, what are essential things to remember in teaching debate?
Q5. Did you encounter any problems in planning lessons for the debate class? If yes, how did you solve them?
Q6. What kinds of resources would you need in planning future lessons for the debate class?

Appendix II
Semi-structured Interview Questions

Q1. Please describe the process when you were planning the course.
Q2. What were the challenging things when you were designing the course?
Q3. Please explain the process of selecting/creating the teaching materials.
Q4. What aspects did you try to emphasize the most when you were teaching the debate course? Why?
Q5. What were some of the difficulties that students faced during the course?
Q6. To what extent did the classroom facilities affect your teaching?

【Research Article】

Training Learners of Writing to Use Online Vocabulary Tools to Increase Lexical Richness

Tanya L. Erdelyi

Abstract

Lexical richness is often used as an indicator of productive language proficiency. Researchers regularly utilize a variety of digital tools (i.e., VocabProfile, RANGE) to analyze the lexical richness of a language learner's written text. However, many studies do not provide the opportunity for learners to use the same tools to analyze their own writing. The following longitudinal mixed-methods study analyzed the written essays of 36 Japanese lower English proficiency first-year university students, before and after they received training in the usage of VocabProfile, Corpus of Contemporary American English, and Google's Ngram Viewer. The first and final drafts of the essays were analyzed with RANGE and compared with a paired-samples *t*-test. Next, 65 randomly selected sentences containing low-frequency words were analyzed with a 4-point scale for errors by six human raters. Finally, a self-efficacy questionnaire completed by the students was compared with the statistical results. The results indicated a significant increase in the error-free lexical richness of the language learners' essays between the first and final draft.

Keywords: *Academic writing, Lexical richness, VocabProfile*

Learning vocabulary is a key component for developing language proficiency. Once the most common high-frequency words in English have been obtained (i.e., the first 2000 most frequent words), the next logical step is for language learners to continue increasing their academic and low-frequency vocabulary knowledge, especially for language learners at the tertiary level of education. Increasing a learner's lexical richness can assist the learner in the reading comprehension of academic texts, as well as enable them to mirror the language used in their field of study while writing.

Lexical richness is often used to determine the academic quality of a language learner's productive language. In terms of writing, generally speaking, lexical richness is determined by analyzing a language learner's written work for the density and variety of low-frequency words. Furthermore, the words should be accurate in meaning, grammar, and word combinations or frequent collocates (Nation & Webb, 2011). To measure lexical richness, some tools that are often used are Lexical Frequency Profiles such as Tom Cobb's VocabProfile for measuring lexical frequency (Abbasian & Shiri, 2011; Cobb, 2002.; Laufer & Nation, 1995) or RANGE (Heatley *et al.*, 2002; Kyle, 2019). Additionally, concordances found in the Corpus of Contemporary American English (COCA) and Google's Ngram Viewer are often used to check collocations and multiword units. By using these tools to examine these factors in a written text, the lexical richness of said text can be determined.

Data-driven learning (DDL) and Corpus-based learning (CBL) are common approaches for increasing a language learners' lexical richness. The focus of DDL is the use of computers as a tool for language learners. CBL involves corpus-based tools, such as a concordancer, for language learning. By providing learners with examples of linguistic performance through tools such as a keyword-in-context concordance, the learners are encouraged to use their brains to decipher the correct way that language is used (Johns, 1991).

Although there are a multitude of studies conducted on lexical richness, DDL, and CBL, there is a need for more empirical data on CBL and DDL (Gaskell & Cobb, 2004; Gries, 2015). Moreover, there are a few issues that should be explored more fully. Many of the studies have focused primarily on intermediate to advanced language learners (Cobb, 2010; Gaskell & Cobb, 2004; Granger, 2012; Henriksen & Danelund, 2015; Nesselhauf, 2005), with little focus on learners with lower language proficiency. Moreover, various studies have been performed in one sitting (Cobb, 2010; Gaskell & Cobb, 2004; Gilmore, 2009) as opposed to a more longitudinal design, such as across an entire semester. Additionally, most studies involve teacher-supplied resources or employ teacher-designed worksheet-based methods derived from some of the previously mentioned vocabulary tools (Cobb, 2010; Gaskell & Cobb, 2004; Granger, 2012) instead of allowing the students to explore their own writing by personally using these tools. Finally, although Lexical Frequency Profiles can measure lexical richness in terms of word frequency (Abbasian & Shiri, 2011), they cannot measure the accuracy of word use in terms of grammar and semantics. Therefore, more research is required to address these issues.

As mentioned above, many researchers have made an effort to examine the vocabulary use of language learners to acquire information about the complexity of their language. They employ a vast array of tools designed to analyze and evaluate the vocabulary used by their participants. Nevertheless, equipping these learners with the very tools used by researchers might yield interesting results. By allowing the learners to analyze and evaluate their own writing through explicit instruction in utilizing some of these vocabulary tools and discussing the benefits of the data they receive, it is hoped that the lexical richness of the learners' writing may improve.

The following study attempts to contribute to the research that has been conducted on lexical richness, DDL, and CBL. It was designed to explore the effects of teaching learners to use vocabulary-related tools to increase the percentage of academic and low-frequency words in their academic research papers, thus improving the paper's lexical richness. An additional focus of the current study is to check the effectiveness of providing tools to help reduce the number of errors when the learners increase the number of academic and low-frequency words.

In order to address these issues, the following three research questions were created and answered through document and statistical analysis:

1. What, if any, is the increase in coverage of academic words after receiving explicit instruction in the use of vocabulary tools for analyzing written work?
2. What percentage of the academic and low-frequency words used to replace high-frequency words was both grammatically and semantically correct?
3. How do errors in writing before and after the treatment instruction compare in both quantity and quality?

Literature Review

The following is a brief overview of literature related to the topic of study in this paper. A closer look will be given to studies conducted on lexical richness, DDL, and CBL.

Lexical Richness

Several studies have provided results on the existence of lexical richness in language learners' writing. Some studies have shown that lexical richness in undergrad students reflects their

pre-existing knowledge of vocabulary (Ha, 2019). However, much of the existing literature on lexical richness shows that there is a paucity of lexical richness in second language writing. Some studies (Henriksen & Danelund, 2015) have shown that language learners with a higher level of English proficiency tend to rely heavily on comfortable and easier to use high-frequency vocabulary rather than attempting to incorporate their low-frequency vocabulary.

One criticism of how lexical richness is measured is that Lexical Frequency Profiles such as VocabProfile and RANGE only provide measures of vocabulary quantity and do not account for errors in grammatical and semantic use. Therefore, although these Lexical Frequency Profiles indicate lexical richness to some extent, they should be used in conjunction with other forms of measure (Abbasian & Shiri, 2011). Although some studies (Stæhr, 2008) use human raters to analyze student writing holistically for errors in grammar and vocabulary, many studies on lexical richness focus solely on the number and types of lexical items (Laufer & Nation, 1995; Lei & Yang, 2020).

Data-Driven Learning (DDL) and Corpus-Based Learning (CBL)

Studies on DDL and CBL have uncovered many results concerning the use of online tools to help improve language learner writing and recognize errors. According to several researchers, DDL has successfully helped learners find their own solution to language problems using authentic resources and tools (Boulton, 2009; Cobb, 2010; Gilmore, 2009; Granger, 2012; Johns, 1991). Researchers have also discovered that CBL is useful for recognizing patterns in grammar and word use to help with error correction (Cobb, 2010; Gaskell & Cobb, 2004; Gilmore, 2009). Moreover, studies have shown that concordance information is useful for intermediate and advanced learners during writing activities (Cobb, 2010; Gaskell & Cobb, 2004).

Studies on DDL and CBL have employed a variety of methods. Given the inherent difficulty of using some of the available online tools (Lee & Lin, 2019), many of the studies have focused primarily on intermediate to advanced proficiency level language learners to lessen the cognitive load (Cobb, 2010; Gaskell & Cobb, 2004; Granger, 2012; Henriksen & Danelund, 2015; Nesselhauf, 2005). Often in response to the difficulty in using some of the tools, various studies have resorted to employing teacher-supplied resources or worksheets designed by teachers that had been derived from the online vocabulary tools (Cobb, 2010; Gaskell & Cobb, 2004; Granger, 2012) rather than asking the language learners to use the tools themselves. Additionally, regarding the implementation of many DDL and CBL studies, a common preference has been to conduct the research in one sitting (Cobb, 2010; Gaskell & Cobb, 2004; Gilmore, 2009) as opposed to a more longitudinal design, such as across an entire semester.

After reviewing the existing literature, several missing components were discovered. To begin with, more studies that evaluate lexical richness in terms of grammatical and semantic accuracy are needed. Additionally, there is a need for more studies on lexical richness with lower level proficiency language learners. Finally, there seems to be a need for more longitudinal studies conducted over a longer period of time that involve the use of authentic online vocabulary tools. Therefore, the current study has been designed to address these issues.

Methodology

The following section contains an explanation of the methodology for the current study. Included are details about the context, research design, procedure, treatment, and questionnaire. It ends with

an explanation of how the data were analyzed.

Context

Research was conducted on the written texts of 36 Japanese first-year university students, 25 women and 11 men between the ages of 18 and 20, learning academic English writing at a private Japanese university. This study was conducted in an academic reading and writing class that met thrice a week for 10 weeks during the third and final term of the learners' first year. At the beginning of the school year, the students wrote a TOEFL PBT for the purpose of placement. The scores of these learners range from 350 to 450, therefore the majority of the students in this study were in the low-intermediate level of English proficiency.

Research Design

This mixed-methods longitudinal study was designed to analyze the lexical richness of academic papers written by lower-proficiency English learners over one entire university semester. It was designed to address some of the issues found in the literature by providing empirical evidence on lexical richness, DDL, and CBL.

Procedure

The learners were asked to write an 8-paragraph secondary research paper comparing and contrasting a topic of their choice in bioethics. The paper was written in sections during the 10-week course. Two drafts of each section were written, with peer editing and revision conducted between each draft and section. At the beginning of the course, the students wrote two drafts of a single introduction paragraph. This was followed by two weeks devoted to the two paragraphs in Section 1. Another two weeks were spent focusing on the two paragraphs in Section 2. The final section, which comprised two main body paragraphs and the concluding paragraph, were written in two drafts over two weeks. Table 1 shows the timeline of the language learners' research paper.

Table 1
Timeline of Research Paper and Treatment

Time	What Was Finished
Week 1	Topic Selection and Outline
Week 2	Introduction Draft 1
Week 3	Introduction Draft 2
Week 4	Section 1 Draft 1
Week 5	Section 1 Draft 2
Week 6	Section 2 Draft 1
Week 7	Treatment Instruction
Week 8	Section 2 Draft 2
Week 9	Section 3 and Conclusion Draft 1
Week 10	Section 3 and Conclusion Draft 2

Treatment

As shown in Table 1, the treatment occurred after the first draft of the second section, at the midway point of the paper. This ensured that the first drafts of the introduction, Section 1, and Section 2 were not affected by the treatment.

For the treatment, the participants were introduced to the Academic Word List (AWL) and asked to begin self-study of the words using various vocabulary-learning strategies such as word cards and practice exercises such as cloze sentences available on the Internet. Then, the participants were taught how to use various vocabulary-related tools to analyze their own writing. Following Nation's (2009) guidelines for training students in learning strategies, each tool, over two lessons in computer-equipped classrooms, was first modeled by the teacher. Next, the participants practiced the different steps alone, and then with partners, reporting back to the teacher when problems arose. Occasional feedback was provided by the teacher and fellow classmates in the peer editing sessions. Finally, further consultation was provided to individual learners during tutorial sessions in the teacher's office.

The learners were asked to analyze their papers with the tools and replace high-frequency words with words from the AWL or low-frequency words. The AWL (Coxhead, 2000) is a list of academic words derived from a corpus consisting of over a million words from academic texts. It contains the most frequent academic English words after West and West's (1953) General Service List (GSL) comprising the 2000 most high-frequency English headwords. The AWL encompasses 570 word families divided into 10 sublists according to frequency. Coxhead (2000) recommends that the AWL should be taught explicitly, allowing for opportunities for the vocabulary to be met in meaning-focused reading and listening texts, and used productively in speaking and writing. Meant purely as a goal to help with motivation, the students were asked to aim for 10% AWL words because according to Coxhead (2000), "The AWL accounts for 10% of the tokens in the Academic Corpus" (p. 222). Following a few tutorial sessions where some of the participants showed concern on finding enough AWL words to boost the percentage, the learners were encouraged not to ignore the low-frequency alternatives that were not present in the AWL as well.

The first tool taught to the participants was Cobb's (2002) VocabProfile, conveniently located on his website, Lextutor. The version they used categorized the first 1000 and second 1000 words from the GSL, and the AWL words in their research papers. All other words were marked as "Off-list" (i.e., low-frequency words, proper nouns, non-English words, spelling errors). Once they had learned to identify the high-frequency words, the participants were shown how to use the right-click functions of both Google Docs and Microsoft Word that provide possible synonyms for the highlighted word in question. The participants were warned that, even though a potential replacement might have been located, the word had to match the sentence in both grammar and meaning. To help check the replacement words in the sentences, the participants were taught how to use the Google Ngram Viewer (Michel *et al.*, 2010) for checking which word combinations are used most often and the color-coded keyword-in-context concordance on the COCA website (Davies, 2010; Johns, 1991) for checking possible collocations and recognizing possible patterns and how others use the language. With these tools, the participants set out to improve the lexical richness of their research papers.

Questionnaire

On the final day of the course, when the final drafts of the research papers were due, the

participants were asked to complete a questionnaire about the tools they did or did not use while improving the lexical richness of their writing. The timing of the questionnaire completion was chosen for maximum effect, as the probability of the participants using the tools prior to the class in order to complete the assignment was high and therefore fresh in their minds. The questionnaire, as seen in Appendix A, consisted of seven multiple-choice questions and one open-ended question for qualitative purposes. The multiple-choice questions were asked to confirm which tools the students used and found most useful and easiest to use and to learn if the students used other tools that were not discussed in class. The open-ended question asked what was most difficult about changing the high-frequency words to more academic or low-frequency words. Thus, a complete picture of the tools that were used was obtained.

Data Analysis

There were three sources of data used in this study. The first was the research papers written by the participants, 72 in total, with 36 first drafts and 36 final drafts. The next available data were from the questionnaire. The last source of data was individual sentences randomly chosen from the papers that contained an AWL or low-frequency replacement word.

Using the steps involved in measuring lexical richness from Nation and Webb (2011, p. 256), the following decisions were made:

Steps Involved in Measuring Lexical Richness

1. Decide on the text to be analyzed (research papers)
2. Decide on the unit of counting (word families)
3. Decide what to do with errors (compare the original and replacements words)
4. Decide on how to measure lexical richness (multiple human raters) (p. 256)

For analyzing data to help answer the first research question concerning the coverage of AWL words, the first and final drafts of all the participants were compared. To ensure that only the participants' words were being analyzed, all direct quotes and proper nouns were removed from each draft. Word family counts for the AWL words were obtained from the RANGE program (Heatley, Nation, & Coxhead, 2002). According to Durrant and Schmitt (2009), much can be learned by examining and comparing individual scores, as these results are often hidden when only entire corpora are compared. Therefore, the first drafts of each participant were compared with their final drafts using a paired-samples *t*-test.

With regard to the data analysis used to answer the second research question regarding the grammatical and semantic fit of the replacement words, 65 sentences from 20 participants were randomly chosen from the fourth paragraph of the first and final drafts that contained academic or low-frequency replacements of high-frequency words. The fourth paragraph was chosen as it came from Section 2 of the paper, where the treatment occurred, thus quite possibly representing the best effort by the participants in using the tools learned during the treatment. Each word replacement was presented in its original sentence and rated on a four-point scale as seen below. The capitalized word was the low-frequency replacement word. The word in brackets was the high-frequency word from the first draft.

Example:

No prospect of UTILIZATION (using) alternative ways has yet emerged.

- both grammar and word choice are correct
- correct grammar, problems with word choice
- correct word choice, problems with grammar
- both grammar and word choice are incorrect

Each of the 65 sentences was rated by at least three of six human raters, all members of the same applied linguistics doctoral cohort as the researcher. The raters were trained as a group to help ensure inter-rater reliability. Inter-rater reliability was calculated by dividing the number of agreed-upon items with the total number of items.

For the final research question concerning the quantity and quality of errors between the two drafts, the same human raters and 65 sentences were used. Following each sentence, the raters were asked to compare and assess each original and replacement word to determine which was more correct using the following four-point scale.

Example of rating scale:

For the previous sentence, which word was most correct?

- Both are equally correct
- Word in ALL CAPS
- Word in (brackets)
- Neither

Results

The following are the results for this study. Discussion of the results is presented in the Conclusions section.

Questionnaire

Table 2 presents the results from the questionnaire that was completed by the participants. As can be seen, nearly all participants used VocabProfile to analyze the word frequency of their papers, as opposed to those few who reported using Ngram or the concordance.

Table 2
Data from Questionnaire

	Used	Most Useful	Easiest	More Instruction	Most Problematic
VocabProfile	97.2%	83.3%	47.2%	50.0%	Matching meaning
COCA	11.1%	2.8%	2.8%	38.8%	41.6%
Ngram	13.8%	2.8%	2.8%	22.2%	Matching grammar
MSWord Syn	66.7%	27.7%	41.6%	16.6%	16.6%
Google Syn	36.1%	13.8%	5.5%	2.8%	Collocations
Grammarly	88.8%	5.5%	13.8%	0.0%	5.5%
Dictionary	11.1%	13.8%	0.0%	0.0%	

Note. Grammarly and dictionary use were not taught by the teacher for this study.

Texts Analyzed

Table 3 presents the descriptive data of the texts that were analyzed. The numbers reflect the number of words after removing the direct quotes and proper nouns.

Table 3
Descriptive Data of Texts Analyzed (Adapted from Durrant & Schitt, 2009)

Description	Drafts	Number of Texts	Number of Writers	Total Words	Mean Words/ Text	Writers' L1
Academic argumentative	First	36	36	46,684	1,297	Japanese
secondary research papers written by first-year Japanese university students with lower-intermediate English proficiency studying English academic reading and writing Paper topic - bioethics	First	36	36	55,255	1,534	Japanese

Individual Scores Between the First and Final Drafts

A paired-samples *t*-test was conducted to evaluate whether the number of AWL word families would increase between the first and final drafts after a treatment on the use of vocabulary tools. The results indicated that the mean AWL count for the final draft (M = 61.86, SD = 25.68) was significantly greater than the mean AWL count for the first draft (M = 42.78, SD = 20.11), $t(35) = -3.51$, $p = .0004$. The standardized effect size index, d, was .83. The 95% confidence interval for the mean difference between the two ratings was 35.97-70.55 (Green & Salkind, 2013).

Percentage of Grammatically and Semantically Correct Replacements

Table 4 displays the percentages assigned to the replacements as determined by the human raters. The majority ruling for the rating of each replacement word was used in the analysis. The inter-rater reliability was 86%. It should be noted that the highest percentage was obtained by word choices that were both grammatically and semantically correct.

Table 4
Analysis of the Replacement Words

Grading Criteria	Percentages
Both grammar and word choices are correct	56.9%
Correct grammar, problems with word choice	35.4%
Correct word choice, problems with grammar	6.2%
Both grammar and word choice are incorrect	1.5%

Comparison of Errors

Finally, a comparison between the original high-frequency word used in the first draft was compared with the replacement academic or low-frequency word to determine which word, if any, was more correct. Table 5 exhibits the percentages of the error judgments made by the raters. The scores used for the analysis were the ratings that received the majority vote from the raters for each

pair of words that were analyzed. The rater reliability for this was also 86%. Of particular interest is that 90.9% of the replacements were considered to be more correct by the human raters.

Table 5
Analysis of the Errors

Grading Criteria	Percentages
Both are equally correct	55.4%
The academic replacement is correct	35.4%
The original high-frequency word is correct	7.7%
Neither is correct	1.5%

Conclusions

This discussion is presented to help analyze the results, situate the findings in the existing literature, and answer the following three research questions: 1) What, if any, is the increase in coverage of academic words after receiving explicit instruction in the use of vocabulary tools for analyzing written work? 2) What percentage of the academic and low-frequency words used to replace high-frequency words was both grammatically and semantically correct? 3) How do errors in the first and final drafts compare in both quantity and quality?

Findings

Although many of the new tools, such as Ngram and the COCA, were not used by the majority of the participants, as indicated in Table 2, all but one participant used the VocabProfile. The fact that the more difficult Ngram and COCA were underutilized coincides with the findings from Lee and Lin (2019). In addition, many students also used the synonym functions in Microsoft Word and Google Docs. Therefore, as shown by the results of the *t*-test, there was a significant increase in the coverage of academic words from the first to the final draft. As there was no control group, it cannot be definitely said that the increase in coverage was due solely to the explicit instruction of the vocabulary tools. However, the learners appear to have gained awareness of the frequency level of their vocabulary just by using VocabProfile. Should they continue to use the tools, it could potentially increase their lexical richness in future papers.

Additionally, the second research question was created to address issues about the lack of attention given to the accuracy of vocabulary use especially regarding grammar and meaning, as discussed in studies such as that of Abbasian and Shiri (2011). Human raters analyzed the language learners' written work, similarly to Stæhr (2008). Ultimately, 56.9% of the low-frequency replacements were deemed both grammatically and semantically correct by the raters. Furthermore, the results from the raters coincided with the qualitative data collected from the questionnaire, as can be seen in Table 2. At 41.6%, the participants reported that matching the meaning of the replacement word to the original was the most difficult, whereas only 16.6% deemed grammar as the most difficult factor. When compared with the results in Table 4, the sample replacement words had 35.4% errors in word choice as opposed to 6.2% grammar errors. This might indicate a need for further instruction in the Ngram and COCA concordance tools, and it also coincides with Lee and Lin's (2019) findings.

Concerning the third research question, the raters determined that the replacement words, on average, were more correct than the original high-frequency words. This might indicate that the

tools were useful for replacing the high-frequency words, as they possibly provided the learners with opportunities to analyze more closely the grammar and meaning of the words being used than they normally would have. These results are consistent with some of the previous studies (Boulton, 2009; Cobb, 2010; Gilmore, 2009; Granger, 2012; Johns, 1991).

Implications

This study provides more empirical data on lexical richness, CBL, and DDL that might help fill the gaps in the literature discussed earlier. It is also a longitudinal study across an entire semester on learners with lower English proficiency. However, the most important point about this study is that it could possibly show the potential involved in encouraging the learner to use some of the vocabulary tools generally used by teachers and researchers, especially VocabProfile. Because of the improvements in lexical richness that occurred through the use of these vocabulary tools, teachers might consider training their students to use vocabulary analysis tools more actively as part of their students' writing process.

Limitations and Future Study Recommendations

There are a few limitations and possible recommendations for future research that arose during this study and should be addressed.

First, the learners were also asked to use low-frequency words that were off-list from the GSL and AWL as replacements. With the removal of proper nouns from the participants' research papers, the remaining off-list words are presumably all low-frequency words. However, both RANGE and VocabProfile do not report the word family count for these off-list words. Moreover, according to the raw data, all 36 participants' off-list word counts increased, at least as was indicated by the word types percentage in RANGE. Thus, processing the two drafts of the papers through the British National Corpus (BNC) version of RANGE to check for the frequency of the words used between the 1st -14th 1000 words in English was considered. However, it should be noted that the GSL and the BNC are not completely analogous, thus introducing a limitation to the analysis of this study. The GSL/AWL were chosen for their ease of comprehension and processing for the participants. Future researchers might consider using the BNC for obtaining a more accurate account of the increase in low-frequency word replacements.

Additionally, because of the low percentage of use with the Ngram and COCA for checking the appropriateness of the replacement words, and as indicated on the questionnaire, more instruction should be given in the use of these valuable tools.

Furthermore, there is the possibility of distorted results from participants who did not use all of the tools (Gilmore, 2009). However, a counter to this was attempted by comparing individual scores (Durrant & Schmitt, 2009; Granger, 2012).

Moreover, further confounding variables, such as the effects of peer editing, and other writing tools not taught in the treatment, such as Grammarly, should be isolated or controlled for in future studies.

In addition, as Gilmore (2009) explained, there is a need for a control group. This study did not have a control group. Potentially, a similar writing assignment from the previous year's cohort at the same university from the same academic reading and writing course could have been used as a control group for this study. However, due to time constraints and the need for obtaining the

permission of the previous year's students, this could not be accomplished.

Finally, the placement of the treatment was handled acceptably for this particular research paper. However, as this particular research paper was written in sections, the first drafts of Section 3 and the concluding paragraphs quite possibly were affected by the treatment. Therefore, future studies might elicit purer results if the treatment is placed between the first and second drafts of an entire paper or essay.

References

Abbasian, G. R., & Shiri, P. M. (2011). Validation of lexical frequency profiles as a measure of lexical richness in written discourse. *Technology of Education Journal, 5*(2), 125-131.

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *Recall*, *21*(1), 37-54.

Cobb, T. (2010). Instructional uses of linguistic technologies. *Procedia-Social and Behavioral Sciences*, *3*, 14-23.

Cobb, T. (2002). Web VocabProfile (an adaptation of Heatley, Nation & Coxhead's (2002) Range). *Lextutor.* Retrieved on Feb 12, 2022 from http://www.lextutor.ca/vp/.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213-238.

Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing, 25*(4), 447-464.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL-International Review of Applied Linguistics in Language Teaching*, *47*(2), 157-177.

Kyle, K. (2019). Measuring lexical richness. In S. Webb (ed.), *The Routledge handbook of vocabulary studies* (pp. 454-476). Routledge.

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, *32*(3), 301-319.

Gilmore, A. (2009). Using online corpora to develop students' writing skills. *ELT Journal*, *63*(4), 363-372.

Granger, S. (2012). Learner corpora. In C. A. Chapelle (ed.), *The Encyclopedia of Applied Linguistics* (pp. 3235–3242). Wiley-Blackwell.

Green, S. B., & Salkind, N. J. (2013). *Using SPSS for Windows and Macintosh* (7 edition). Pearson.

Gries, S. (2015). Statistical methods in learner corpus research. In G. Gaëtanelle, S. Granger, & F. Meunier (eds.), *The Cambridge handbook of learner corpus research* (pp. 59-181). Cambridge University Press.

Ha, H. S. (2019). Lexical richness in EFL undergraduate students' academic writing. *English Teaching*, *74*(3), 3-28. doi:10.15858/engtea.74.3.201909.3

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. From http://www.victoria.ac.nz/lals/staff/paul-nation.aspx.

Henriksen, B., & Danelund, L. (2015). Chapter two studies of Danish L2 learners' vocabulary knowledge and the lexical richness of their written production in English. *Lexical Issues in L2 Writing*, *29*.

Johns, T. (1991). *Should you be persuaded: Two samples of data-driven learning materials* (pp. 1-13). PDF file. https://lexically.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*(3), 307-322.

Lee, P., & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, *81*, 14-25. https://doi.org/10.1016/j.system.2018.12.011

Lei, S., & Yang, R. (2020). Lexical richness in research articles: Corpus-based comparative study among advanced Chinese learners of English, English native beginner students and experts. *Journal of English for Academic Purposes*, *47*, 1-9. https://doi.org/10.1016/j.jeap.2020.100894

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Brockman, W., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2010). Quantitative analysis of culture using millions of digitized books. *Science, 331*(6014). doi:10.1126/science.1199644

Nation, I. S. P. (2009). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle Cengage Learning.

Nesselhauf, N. (2005). *Collocations in a learner corpus* (Vol. 14). John Benjamins Publishing.

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139-152. doi:10.1080/09571730802389975

West, M., & West, M. P. (Eds.). (1953). *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. Addison-Wesley Longman Limited.

Appendix A
Vocabulary Questionnaire

Please answer these questions HONESTLY about your research paper and vocabulary. You may check more than one answer if you need to:

1. Which vocabulary tools did you use?
   - ☐ VocabProfile
   - ☐ COCA concordance
   - ☐ Google Ngram Viewer
   - ☐ Google Docs synonyms (right click and define)
   - ☐ Microsoft Word synonyms (right click and synonym)

2. Did you use any other tools I didn't teach you?
   - ☐ Dictionary – what kind? _____
   - ☐ Grammarly
   - ☐ Only my brain
   - ☐ Other(s) _____

3. Which tool(s) were the most useful?
   - ☐ VocabProfile
   - ☐ COCA concordance
   - ☐ Google Ngram Viewer
   - ☐ Google Docs synonyms
   - ☐ Microsoft Word synonyms
   - ☐ Other(s) _____

4. Which tool(s) were the easiest to use?
   - ☐ VocabProfile
   - ☐ COCA concordance
   - ☐ Google Ngram Viewer
   - ☐ Google Docs synonyms
   - ☐ Microsoft Word synonyms
   - ☐ Other(s) _____

5. Which tool(s) would you have liked more instructions/directions from the teacher?
   - ☐ VocabProfile
   - ☐ COCA concordance
   - ☐ Google Ngram Viewer
   - ☐ Google Docs synonyms
   - ☐ Microsoft Word synonyms
   - ☐ Other(s) _____

6. Did your peer editors help with your academic words?

☐ Yes
☐ No
☐ I don't know

7. Did you help your writing group members with their academic words?

☐ Yes
☐ No

8. What was most difficult about changing the vocabulary to more academic words?

【研究論文】

中国語における「多＋形容詞」と「有＋多＋形容詞」

李　菲

要旨

「どのくらい～だ」を意味する中国語の「多＋形容詞」は、単独で述語になる以外、動詞「有」を述語に立てた「有＋多＋形容詞」の文を作ることができる。本稿は、「有」のない「多＋形容詞」文と「有＋多＋形容詞」文について、両者の統語と意味上の差異や使い分けについて考察を行った。「有」のない「多＋形容詞」文が成立するのは、「多＋形容詞」自体がイディオム化し、「多大」「多高」のように専ら「年齢」「身長」を聞くための表現として用いられるときである。一方、形容詞が表す状態の程度が通常より「高い」ことがわかっていて、その上で「程度の高さ」を尋ねる場合には、動詞「有」を述語に立てた「有＋多＋形容詞」の文しか用いることができず、「有」のない「多＋形容詞」文が成立しにくい。従来の文法記述は、「有」のない「多＋形容詞」を基本的な用法と考え、「有＋多＋形容詞」の文に対して、「「有」が入ることがある」と記してきたが、「多＋形容詞」のイディオム化に関係なく成り立つ「有＋多＋形容詞」の文は制限が少ないという点で、基本的用法と考えるべきである。

キーワード：*有　形容詞　イディオム　how*

1．問題提起

中国語では身長、年齢を尋ねるとき、次のように、「多＋形容詞」の表現を用いる。

（1）　你　　多　　高?
　　あなた　どのくらい　高い
　　「あなたは身長が何センチですか。」

（2）　你　　多　　大?
　　あなた　どのくらい　年上
　　「あなたはいくつですか。」

「多高」、「多大」は直訳すると、「どのくらい高い」「どのくらい年上」という意味であるが、日本語でいうところの「身長は？」「いくつ？」と同じ意味である。「多高」、「多大」は、意味と構造の面で英語の how tall、how old と似ている。すなわち、両方とも「身長」「年齢」を尋ねるのに、「身長は？」「いくつ？」のように名詞表現を使わずに、「どのくらい＋形容詞」の構造をした表現を用いる点で共通している[1]。一方、両者には大きな違いもある。英語の how tall、how old は be 動詞を伴うのに対し、中国語の「多高」「多大」は（1）（2）のようにそのまま述語となるか、または以下のように動詞の「有（ある）」を伴うことができる。

（1’）　你　　**有**　　多　　高?
　　あなた　ある　どのくらい　高い
　　「あなたは身長が何センチですか。」

1　日本語には、「年齢が上である」ことを意味する英語の old や中国語の「大」のような形容詞がない。「年齢」を尋ねるのに、日本語では「どのくらい＋形容詞」のように聞くことができず、専ら「年齢は？」「いくつ」「何歳」などの名詞表現を用いて表現する。

（2’） 你　　**有**　　多　　大?
　　あなた　ある　どのくらい　年上
　　「あなたはいくつですか。」

（1）と（1’）、（2）と（2’）の違いは「有」の有無だけで、同じ意味を表す表現としてとらえることができる。どちらも「身長、年齢」を尋ねることができる。つまり、「有」はあってもなくてもいい、オプショナルなものということができる。身長、年齢以外に、距離、物の大きさや寸法を尋ねる場合でも、動詞の「有」がオプショナルである。

（3） 这条铁路　**有**　　多　　长?（吕叔湘（1996））
　　この鉄道　ある　どのくらい　長い
　　「この鉄道はどのくらい長いですか。」

（3’） 这条铁路　多长[2]?

（4） 你家　　的　电视机　**有**　　多　　大?
　　あなたの家　の　テレビ　ある　どのくらい　大きい
　　「あなたの家のテレビはどのくらい大きいですか。」

（4’） 你家的电视机　多大?

（5） 那个柜子　**有**　　多　　高?
　　あの棚　　ある　どのくらい　高い
　　「あの棚はどのくらい高いですか。」

（5’） 那个柜子　多高?

一方、上の例に対し、「有」が必須の場合もある。次の例が示すように、難易度や影響の大きさを尋ねる場合には、「有＋多＋形容詞」しか用いることができず、「多＋形容詞」の文にすると表現自体の許容度が落ちる。

（6） 汉语　　**有**　　多　　难?
　　中国語　ある　どのくらい　難しい
　　「中国語はどのくらい難しいですか。」

（6’） ?汉语　多难?

（7） 影响　**有**　　多　　大?（CCL）
　　影響　ある　どのくらい　大きい
　　「影響はどのくらい大きいですか。」

（7’） ?影响　多大?

2 （3’）は吕叔湘（1980）の記述に基づく筆者の作例である。なお、本稿の例文で、出典が示されていないものはすべて筆者の作例となる。

吕叔湘（1980）は、「「多＋形容詞」の文では「有」が出現しやすい」という指摘をしている。しかし、上の２組の例文では、動詞の「有」が必須であるので、この指摘に問題があることがわかる。「「多＋形容詞」の文では「有」が出現しやすい」ならば、「多＋形容詞」のみが成立し、「有＋多＋形容詞」が成立しない例もなくてはならないが、そうした例が現段階ではあまりみられない。一方、（6）と（6’）、（7）と（7’）は「有＋多＋形容詞」のみが適格な文となることを示すものなので、この点から、むしろ「「有＋多＋形容詞」の文では、「有」がなくてもいい場合がある」という事実がみえてくる。

また、動詞の「有」の有無という観点とは別に、構文の観点から、「有＋多＋形容詞」と「多＋形容詞」の用法について比較し、その差異について考察する必要がある。上の一連の例文から、両者は部分的に使用場面（用法）が重なるものの、完全に同じように使えるわけではないことがみてとれる。ならば、どの場面では両方を同じように使うことができ、またどのような場面になると片方の構文しか用いることができないのかについて、もっと詳細な記述を行う必要がある。本稿ではこうした一連の問題をめぐって、用例を通して考察と記述を行う。

2 「多＋形容詞」疑問文

2.1 イディオムとしての「多＋形容詞」

「多＋形容詞」は直訳すると、「どのくらい〜ですか」という意味であり、生産性の高い表現である。一方、実際の言語運用の場面では、大きさ、長さ、高さ、距離、重さなどを尋ねるための定型フレーズとなっているものが多く、文というよりフレーズの印象が強い。

（1） 你多高?
「あなたは身長が何センチですか。」

（2） 你多大?
「あなたはいくつですか。」

（4’） 你家的电视机多大?
「あなたの家のテレビは何インチですか。」

（8） 离　这儿　多　远?
から　ここ　どのくらい　遠い
「ここからどのくらい離れていますか。」

（9） 你　的　行李　多　重?
あなた　の　荷物　どのくらい　重い
「あなたの荷物は何キロですか。」

身長・高さは「多高」、年齢とサイズは「多大」、距離は「多远」、重さは「多重」というように、これらの「多＋形容詞」の表現はほとんど一語化している。一語化しているということは、構文を構成している構成素自体の存在が感じられなくなり、構文の分析性と合成性が低くなることを意味する。それは同時に、構文のイディオム化を意味する[3]。イディオムの特徴は、構成素どうしの意味が「１＋１＝２」にならない点に

3 本稿でいう構文とは、「形式と意味の慣習的な結びつき」（斎藤・田口・西村（2015））を指す。よって、文レベルのものだけでなく、慣習的に結びついている短いフレーズやイディオムも構文となる。そして、構文は必ず、非合成的で、各構成素に還元できない「構文的な意味」を含んでいる。

ある。これが非合成的、非分析的な意味といわれている[4]。一語化した「多高」「多大」にもイディオム的な意味がみられる。たとえば、「多高」の直訳は「どのくらい高い」であるが、身長が高くない人にも「多高」で身長を聞くしかない。同様に、大きさを尋ねる「多大」は直訳すると、「どのくらい大きい」であるが、物の大きさに関係なく使うことができる。直訳とはいわば、構成素の意味を一つずつ足すことで合成的に全体の意味をとらえようとすることである。以上の「多＋形容詞」はいずれも、直訳が効かないイディオム的な意味をもっている。

「多高」などが本来の直訳的な意味を失っていることは、それを用いた質問への返答の仕方からもみてとれる。「多高」は直訳すると、「どのくらい高い」という意味で、これは「（何かが）高い」ということを前提とした上で、その高さの程度を尋ねる表現である。つまり、本来は単に身長を聞く質問ではなく、高さを示すような返答を要するものである。しかし、「多高」の答えにはこのような縛りがない。高いか否かに関係なく、単に身長を答えるだけで会話が成り立つ。

仮に「どのくらい高い」という文字通りの質問に対しては、「高い」という前提を根底から否定する返答も考えられ、「全然高くないんだ」と答えることができる。しかし、このように答えてしまうと、「質問と答えがかみ合わない」チグハグな会話を作ってしまう。

（10） a 他多高?
「彼は身長が何センチですか。」
b ?他一点儿也不高[5]。
「彼は少しも背が高くない。」

（10）b の不成立から、（10）a がもはや「どのくらい高い」という文字通りの意味をもっていないことがわかる。よって、（10）a における「多高」は単に身長、高さを聞くためのイディオムであって、高さの程度を聞くという意味機能はもう失われているといえる。これと同じことが、大きさや年齢を聞く「多大」、重さを聞く「多重」にも起きている。文字通りの「大きさ・重さの程度」を聞くという意味機能が失われているため、大きさや重さに関係なく数値だけを答えることが許され、逆に「大きくないんだ、重くはないんだ」とわざわざ否定する返答は変になる。

（11） a 你多重?
「あなたは何キロですか。」
b ? 我不重。
「私は重くないです。」

「多高」「多大」がイディオム化したことは、統語上の特徴からもみてとれる。これらは単独で述語になれる以外、次のように動詞の目的語成分になることができ、名詞と同じふるまいをしている。

（12） 长　　了　　多　　高
伸びる　ASP[6]　どのくらい　高い

4 認知言語学ではイディオムの例としてよく、英語の kick the bucket が取り上げられる。直訳すると「バケツを蹴る」だが、実際は「死ぬ、くたばる」を意味する。よって、構成素の意味を足しても全体の意味にたどりつけないのがイディオムとしての重要な特徴となる。こうしたイディオムは、どの言語にも数えきれないくらい大量に存在する。たとえば、日本語の「割りを食う」という表現。筆者がかつて村上春樹の「わりをくう男」というエッセイを読んだとき、「わりをくう」という日本語が最後まで理解できなかった。「わり」は「割合」「割引」などで「パーセンテージ」的な意味として理解していたが、「割りを食う」の意味には全く「歯が立たなかった」。

5 本稿では、考察対象である「多＋形容詞」「有＋多＋形容詞」の例文にはグロス（逐語訳）を付しているが、それ以外の例文は、紙幅の関係でグロスを省略した。

6 ASP は、アスペクトマーカーを指す。

「どのくらい伸びた」

（13） 活　　到　　多　　大（CCL）
生きる　まで　どのくらい　年上
「何歳まで生きた」

（12）のアスペクト「了」や、（13）の結果補語「到」の直後は目的語となる名詞句が現れる場所である。（12）（13）のような用例は、「多高」「多大」が名詞となったことを意味する。これは、日本語の数量を聞く疑問詞「どのくらい」「いくつ」「いくら」が名詞としてふるまうのと同じである。

こうした「多＋形容詞」フレーズの中で、もう一つ、数量を聞く「多少（どのくらい）」についても少し言及する必要がある。「多少」は他のフレーズと違って、単独では述語になるより、動詞の直後で目的語になることが多い。

（14） a 你　　**有**　　多少？
あなた　ある　どのくらい
「あなたはどのくらいありますか」
b？你　　多少？

（15） 你　　买　　了　　多少？
あなた　買う　ASP　どのくらい
「あなたはどのくらい買いましたか」

「有」のない（14）b は文法的に正しい文とはいえない。これは「多少」が「多高」「多大」と最も異なる点である。どちらもイディオム化しているものの、本稿では「多少」を通常の「多＋形容詞」フレーズと区別して考える[7]。

2.2　イディオム化していない「多＋形容詞」

以上は、イディオム化した「多＋形容詞」の意味、統語上の特徴についての考察である。では、イディオム化する前の、程度の高さを問う、「どのくらい〜だ」を意味する「多＋形容詞」の用法はどうだろうか。1章で取り上げた（6）と（6’）のペアを思い出されたい。「中国語が難しいことがすでにわかっていて、その上で、どのくらい難しいのか」を表すのは「有＋多＋形容詞」の文である（(6)）。「多＋形容詞」を用いた（6’）は自然な表現ではない。

（6）　汉语**有**多难？
(中国語はどのくらい難しいですか。)

（6’）　？汉语多难？

つまり、イディオム化していない「多＋形容詞」は単独で述語になれない。この問題は、従来の研究においてあまり指摘されていない。「多＋形容詞」は文字通りに訳すと、「どのくらい〜だ」を表し、形容詞が表

7 「多少」は「多＋形容詞」がイディオム化したものではない可能性がある。中国語には「高低（高さ）」「大小（大きさ）」といった、反義語の関係にある形容詞を並べて「程度」という意味の名詞を作り出す語形成法がある。よって、「多少」は「多＋形容詞」ではなく、「多＋少」の構造をもつとも考えられる。語の形成、構造がはっきりしない以上、さしあたり「多少」を考察対象から外す。

す状態の高さを聞く表現としてとらえられている。しかし、（6）と（6’）のペアで示したように、「難しさの程度」を問うことができるのは「多＋形容詞」ではなく、「有＋多＋形容詞」の方である。この場合、動詞「有」を述語としなければいけない。これは、イディオム化した「多高」「多大」が単独で述語になれるのと大きく異なる。このことについて、もう少し例文を通して検証したい。ここで、（7）と（7’）のペアをもう一度みてみよう。

（7） 影响**有**多大？（CCL）
「影響はどのくらい大きいですか。」

（7’） ?影响多大?

（7）は、ある事態の影響がいかに深刻なものかを聞く場面での文である。「深刻な影響があった」ということがその前の文脈ですでに語られていて、それをふまえ、「いかに深刻なのか、どれだけ深刻なのか」を聞いている。このような場面では、（7’）のように表現することができない。「多大」が単独で述語になるときは常にイディオムとして、「年齢は？」や「サイズは？」を意味する。本来の、「どのくらい大きい」という意味を表す場合、（7）のように動詞「有」を伴わなければならない。ここからも、「多＋形容詞」は単独では、程度の高さを聞く表現として自立できない（述語になれない）ことがわかる。

ここで、「多＋形容詞」と英語の「how ＋形容詞」を比較してみたい。両者は意味、構造の面で類似しているため、比較しやすく、その分違いもみえやすい。「多＋形容詞」と同様、「how ＋形容詞」の構造をもつ表現の中に、how much、how many、how old といったイディオム化したものが存在する。大雑把にいえば、how much は「いくら」、how many は「どのくらいや頻度」、how old は「いくつ」を意味する疑問詞である。一方、これら以外に、how は様々な形容詞と結びつくことができ、その形容詞が表す程度の高さを聞くことができる。よって、「how ＋形容詞」の形をもつ表現の中に、イディオム化したものとその都度生産的に作られているものとが混在しているといえる。前者は、イディオム的な意味をもつ。後者は、その都度組み立てられた一回限りの表現として、文字通りの「程度の高さ」を聞くという意味機能をもつ。混在しているものの、両者の意味に違いがみられる。次の（16）はイディオム how old の例である。（17）の C1 は、その都度組み立てられた一回限りの表現である。

（16） **How old** are you ?
（あなたはいくつですか。）

（17）
A1 There is an **angry** husband in your living room.
（**怒った夫**がリビングにいるぞ）
C1 **How angry?**
（**どのくらい激怒している？**）
A2 **……**
C2 Is he packing ?
（何かを所持している？）
A3 What ?
（どういうこと？）
C3 Does he have a weapon ?
（武器を持っているのか？）
A4 He has a stick.
（杖を持っているぞ） （Two and a Half Men, Season 2, Episode 24）

（17）における how angry（C1）の部分に注目されたい。how angry はイディオム化していない、「how ＋形容詞」の表現である。意味は文字通りの、「どのくらい怒っている」である。この場合、（17）の A1 で、「angry husband」が待っていることが告げられているので、「怒っている」ということがすでにわかっていて、その上で、C1 が「じゃ、どのくらい怒っているのか」という怒りの程度を確認している[8]。（17）の how angry を中国語に訳そうとするとやはり、「有＋多＋形容詞」の形を使わざるをえず、「有」のない「多＋形容詞」の表現を用いることができない。

（18） a 他　**有**　　多　　生气?
　　彼　ある　どのくらい　怒っている
　　「彼はどのくらい怒っていますか。」
　　b[?] 他　　多　生气?

ここで、「how ＋形容詞」と「多＋形容詞」の違いがようやく浮き彫りになった。両者は、どちらもイディオム化したものとそうでないものを含んでいるが、「how ＋形容詞」の場合、統語上の変化が起きていない。一方、「多＋形容詞」は述語になることができるのはイディオム化した表現のみで、本来の「程度の高さ」を問うという用法になると、動詞「有」を述語に立てる必要がある。こうした違いは、以下の対訳を比較するとより鮮明になる。

【イディオム】
（19） **How old are** you ?
　　你　多大?
【非イディオム】
（20） **How angry is** he ?
　　他 **有**　多 生气?

この比較から、「多＋形容詞」は単独で述語になれるのは、（19）のようなイディオム化した表現に限られることがわかる。「程度の高さ」を尋ねる場合、「有＋多＋形容詞」を用いなければならない。「how ＋形容詞」との比較が、「多＋形容詞」の思わぬ特徴を発見できた。（6’）の「[?]汉语多难？」や（7’）の「[?]影响多大？」が成立しにくい理由もここにある。難しさの程度、深刻さの程度を表す「多难」「多大」はイディオム化しておらず、ゆえに動詞の「有」を述語に立てる必要があり、「有＋多＋形容詞」の文を用いなければ表現できない。

3　構文からみる「多＋形容詞」と「有＋多＋形容詞」

２節で明らかになった「多＋形容詞」の特徴は、構文の観点からとらえ直すことができる。構文の観点に立つということは、構文全体の意味、統語的特徴をすべてそれぞれの構成素に還元しようとせずに、丸ごととらえることである。このようにみた場合、「多＋形容詞」と「有＋多＋形容詞」との関係は、単に動詞の「有」の有無にとどまらず、二つの、別個の構文であるという見方ができる。単独で述語になることができる「多＋形容詞」は①、「有＋多＋形容詞」は②のように示すことができる。

8　（17）はアメリカのテレビドラマ（シットコムシリーズ）『Two and a Half Men』（邦題：『チャーリー・シーンのハーパー★ボーイズ』）の台詞である。この会話は、主人公のチャーリー（C）が２階の自室で人妻との情事に勤しんでいるところへ、１階にいたチャーリーの弟アラン（A）が急に上がってきて、その人妻の夫が今家に来ているので、急いで１階に降りるよう忠告している場面である。「angry husband」とは、「妻の情事を発見したことで、カンカンになっている夫」を指し、それに対しチャーリーが「どのくらいカンカンなんだ」と怒りの度合いを確かめようとしている。

①　「名詞　＋　多＋形容詞？」
　ex 你　多大?
　　他　多高?

②　「名詞　＋　**有**　＋多＋形容詞？」:
　ex 汉语　**有**　多　难?
　　他　　**有**　多　生气?

①では、イディオム化した「多＋形容詞」が述語となって、年齢、身長を尋ねる文を作る。こうした質問に対して、具体的な数値を一つ挙げれば答えとなる。中国語では、年齢や身長といった具体的な数値はそのまま述語になることができる。

（21）a 你 多大?
　　「あなたはいくつですか。」
　　b 我 38 岁。
　　「私は 38 歳です。」

（22）a 你 多高?
　　「あなたは身長が何センチですか。」
　　b 我 一米七。
　　「私は 170 メートルです。」

（21）b、（22）b のような「名詞＋名詞」の文は、「名詞述語文」と呼ばれている。（21）a、（22）a は、「名詞述語文」と構造が同じである。これに加え、イディオム化した「多大」「多高」は、目的語の位置に現れることができるため（（13）と（12））、名詞としてとらえることができる。この 2 点から、「多＋形容詞」が作る①の構文は、その答えと同様、名詞述語文の一種ということができる。

①に対し、②では動詞の「有」が述語となっており、「有」の文の一種といえる。そして、「多难」「多 生气」は「多大」「多高」のようにイディオム化していないため、直ちに目的語または名詞句と断定することができない。また、「有」の文は通常、存在や所有を表すが、②の構文では動詞「有」の意味を具体的に示すことが難しく、文法化した用法であるといえる[9]。よって、名詞述語文である①の構文に比べて、②の構文は統語的な面で少し特異である。

形容詞が表す状態の程度がいかに高いかを問う②の構文は、実は意味の面でも少し特異である。というのは、②の問いに対し、具体的で明確な答えを提示することが①の問いほど簡単ではない。日本語でも、「いくつ？」「身長は？」と聞かれれば、瞬時に年齢と身長を答えることができるが、「中国語はどのくらい難しいですか」「彼はどのくらい怒っているのか」というような問いに対しては、「そうね、どのくらいだろう…」と、一瞬立ち止まって考えてから答えを出す人が多いと思う。このことは、（17）の会話にも表れている。弟のアランが兄のチャーリーに、「その怒った夫が一体どのくらい怒っているのか（how angry）」と聞かれ、一瞬「……」（A2）とことばに詰まっていた。アランが答えに困っていることを見抜いたチャーリーは続けて、「何かを持っているのか」と武器の有無を確認しようとしたが、それでもアランが「What ？」（A3）と質問の意図を汲み取ることができなかった。激怒している人がいて、その怒りがいかにすごいものかを描写するのは、簡単なことではないことがみてとれる場面である。

また、②の問いに対する返答は、必ずしも一つではないという点においても①の場合と異なる。①の問いは（21）b、（22）b のように数値で答えるのが一般的である。一方、②の問いは程度の高さを尋ねている

9 『现代汉语词典』（第 6 版）では、数量や形容詞を伴う動詞の「有」の意味を、「表示达到一定的数量或某种程度（一定の数量または程度まで達した）」としている。

ので、具体的な数値を挙げたり、数値化できない場合は実際の状況や様子を伝えたり、あるいは他の物に例えたり、比較を行うなどの手法によって伝える。

（23） a 汉语**有**多难?
「中国語はどのくらい難しいですか。」
b 比英语难学。
「英語より難しいです。」

（24） a 他**有**多生气?
「彼はどのくらい怒っていますか。」
b 他的脸都气红了。
「顔が怒りで真っ赤になっている。」

（23）b では、中国語を英語と比較して、難易度のレベルを伝えている。（24）b では、顔が真っ赤になったという様子にふれることで、怒りの激しさを伝えている。難易度、怒りはどちらも数値化できないものなので、（21）b、（22）b のように答えられず、その分色々な答えが考えうる。

4　構文①と②の関係性

「有＋多＋形容詞」の構文②は、構文①と、構造の面だけでなく、意味機能及び答え方において顕著な違いがあることが以上で明らかになった。次は、1 節の問題提起で取り上げた次の四つのペアについて考えてみたい。こうしたペアを通して、構文①と構文②の関係性、及びその意味の差異をとらえてみたい。

（1’） 你**有**多高?
（1） 你多高?

（2’） 你**有**多大?
（2） 你多大?

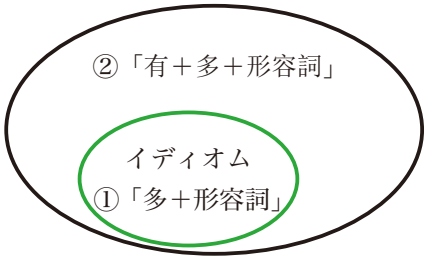
（4） 你家的电视机**有**多大?
（4’） 你家的电视机多大?

（5） 那个柜子**有**多高?
（5’） 那个柜子多高?

（1’）、（2’）、（4）、（5）は②の構文であるが、「多高」は身長や高さ、「多大」は年齢、大きさを表すイディオムである点で、難易度や怒りの激しさを問う他の例と異なる。こうした例からは、構文②にはイディオム化した「多＋形容詞」が現れる場合があることがわかる。イディオム化していない「多＋形容詞」は②、イディオム化した「多＋形容詞」は①と②の両方に現れるということは、①と②は相互排他的な関係ではなく、むしろ包含関係にあるといえる。つまり、②の中に、一部イディオム化した「多＋形容詞」だけが動詞「有」を落とした①の構文を作ることができる。

従来の文法記述では、「多＋形容詞」文に動詞「有」が現れることもあるとしているが、実際はむしろその逆で、「有＋多＋形容詞」の一部が動詞「有」を落とすことができる。形容詞の観点からみた場合、構文②は構文①よりも制限が少なく、生産性が高い。そして、「多＋形容詞」がイディオム化している場合、構文①と構文②の両方が可能となるので、（1’）と（1）、（2’）と（2）、（4）と（4’）、（5）と（5’）のよ

図1　②と①の関係



うなペアができる。

次は、こうしたペアの意味について考えてみたい。①と②は形、意味の面でそれぞれ異なる特徴を示していることはすでに明らかになった。こうした違いは、上の四つのペアにおける意味の差をとらえる上でヒントとなる。まず、答え方からそれぞれのペアにおける両構文を区別できる。つまり、構文①の（1）、(2)、(4’)、(5’）は通常は数値で答えるのに対し、構文②の形をした（1’)、(2’)、(4)、(5）は、具体的な数値に加え、それ以外の答え方もできる。

(1’)　你**有**多高?
　　a 我一米七。(170 センチです。)
　　b 我在我们班最高。(クラスで一番高いです。)

(2’)　你**有**多大?
　　a 我 20 岁。(二十歳です。)
　　b 我比你还大两岁。(あなたよりさらに二歳年上です)

(4)　你家的电视机**有**多大?
　　a 60 英寸。(60 インチです)
　　b 就是一个小影院。(まるでミニシアターのようだ)

(5)　那个柜子**有**多高?
　　a 1 米 8。(180 センチです。)
　　b 能碰到房顶。(天井に届くほど。)

各質問に対して、b の答え方が可能ということは、(1’)、(2’)、(4)、(5）はイディオム化した「多高」「多大」からなっているものの、動詞「有」及び構文全体の力によって、本来の「程度の高さ」を聞くという意味機能を取り戻しているということができる。「クラスで一番高い」「まるでミニシアター」という答えは、「程度の高さ」を表現したものである。例えば、身長が高いことがすでにわかっていて、「じゃ、どのくらい高いのか」という質問に対して、「クラスで一番高い」が最も自然な答えとなる。あるいは、家のテレビが大きいことはすでに知っていて、「どのくらい大きいのか」に対して、「まるでミニシアター」という答えからは「大きさ」が伝わる。

このように考えると、(1’)、(2’)、(4)、(5）は実は多義的（ambiguous）な表現であるという見方もできる。つまり、「多高」「多大」をイディオムとして解釈する場合、a のように数値で答える。一方、本来の「どのくらい高いのか」「どのくらい大きいのか」という「程度の高さ」を聞く文として理解すると、b の答え方が可能となる。イディオムとしてとらえるかどうかは、文脈に左右されるほか、個人の理解によっても異なりうる。これに対し、構文①の（1）、(2)、(4’)、(5’）は専ら「身長」「年齢」「サイズ」「高さ」を尋ねる質問で、程度の高さを聞くという意味が含まれていない。よって、(4’）の質問に対し、「まるでミニシアター」は自然な答え方とはいえず、唐突な感じを与える。

（4’）　你家的电视机多大?
（あなたの家のテレビは何インチですか。）
—? 就是一个小影院。（まるでミニシアターのようだ）

以上から、①と②の構文が重なり合うような場合でも、意味の差がみられることがわかる。イディオム化した「多＋形容詞」は①と②の両方の形をとることができるものの、意味が異なる以上、①と②を異なる構文として区別する必要がある。また、両者は包含関係にあり、構文①を構文②から派生した用法とみることができる。構文②が動詞「有」の文であるのに対し、構文①は名詞述語文の一種である。

5　今後の課題—「有」の文法化

以上、「多＋形容詞」と「有＋多＋形容詞」の成立条件、意味機能などについて考察を行ってきた。名詞述語文である①の文は、②の下位構文として位置づけることができ、「多＋形容詞」がイディオム化した場合のみ成り立つ。一方、②の「有＋多＋形容詞」の文は、イディオム化していない「多＋形容詞」がプロトタイプであり、「程度の高さ」を聞く表現として用いられる。

最後に、中国語の「有＋多＋形容詞」構文を英語の「how ＋形容詞」と比較し、動詞「有」の文法化の問題について少しふれておきたい。英語の「how ＋形容詞」はイディオム化に関係なく、be 動詞を伴う。

［how ＋形容詞＋ **be 動詞**＋名詞］

一方、中国語は「存在」「所有」を表す動詞「有」を用いる。また、「多＋形容詞」のイディオム化で新たな構文が生じている点でも、英語と異なる。

② ［主語＋**有**＋多＋形容詞］
↓イディオム化
［主語＋多＋形容詞］

英語は年齢と身長を尋ねたり、また程度の高さについて質問するとき、一つの構文によって表現できる。この構文において、be 動詞は「how ＋形容詞」の意味内容を発話の現在にグラウンディング（grounding）するという役割をもっている[10]。これは、「存在」や「所有」といった具体的な概念を超越した意味的、文法的機能である。そして、英語は中国語のように下位構文が生じていないのは、be 動詞の文法化と関係していると思われる。これに対し、対応する中国語の構文②における動詞「有」は、be 動詞のように、グラウンディング機能をもつほど文法化しているのだろうか。従来の研究では、②の「有」を「（程度に）達する」と解釈しているが、これもグラウンディン機能の一種と考えてよいか。また、「存在」「所有」「到達」という三つの概念がどのように関係しあっているのか。今後は、グラウンディング機能、文法化の観点から、「有＋多＋形容詞」における動詞「有」の働きについて引き続き考察していきたい。

＜参考文献＞

吕叔湘（主编）1980.『现代汉语八百词』。北京：商务印书馆。
斎藤純男，田口善久，西村義樹（編著）2015.『明解言語学辞典』。東京：三省堂。
Langacker, Ronald W.2008.*Cognitive Grammar:A Basic Introduction*. Oxford: Oxford University Press.

10　be 動詞のグラウンディングについては、Langacker（2008）を参照。

【研究論文】

あいさつをめぐる日・韓母語話者の意識の相違[1]
――ポライトネスからの一考察――

岡村佳奈

要旨

本研究では、インタビュー調査を通じ、①なぜ日本人は韓国人よりも定型的あいさつを多用するのか、②なぜ韓国人は相手との親疎関係によってあいさつを使い分けるのに日本人はそのような使い分けをしないのか、という2点に関する母語話者の意識を探りつつ、あいさつの日韓差が生じた原因を考察した。

その結果、日本語の定型的あいさつは主に無標ポライトネス――「あって当たり前で、それが現れないときに初めてそれがないことが意識され、ポライトではないと捉えられる（宇佐美 2008: 160）」もの――として機能していることがわかった。そのため、日本人はあいさつをあまり使い分けることなく、誰にでも定型的あいさつを頻用するようである。

一方、韓国人は、相手に対する関心・好意、礼儀の必要性の有無が親疎によって異なると思っており、親しくない人にはネガティブ・ポライトネスである定型的あいさつ、親しい人にはポジティブ・ポライトネスである非定型的あいさつを使っていることが明らかになった。親しくない人にのみ定型的あいさつを交わすため、全体としては、日本人よりも定型的あいさつの使用頻度が低くみえるのだと考えられる。

Keywords: *日韓対照、ポライトネス、定型性、定型的 / 非定型的あいさつ、インタビュー調査*

1. はじめに

金香来（2000）で行われた調査によれば、日本人[2]は、「밥　먹었어？（pap mekesse? ご飯食べた？[3]）」、「어디 가세요？（eti kaseyyo? どこ行かれますか？）」のような韓国語のあいさつに抵抗を感じる一方、韓国人は、「すみません」、「お疲れ様」といった日本語のあいさつに抵抗を感じるという。そして、日本人・韓国人が抵抗を感じるあいさつの多くは、それぞれの母語では使われない表現だとも述べられている。母語の影響のために、コミュニケーション上で問題が生じているということになるが、日韓のあいさつにはどのような違いがあり、その違いが生まれた原因は何なのだろうか。

先行研究による言及をいくつかに分類するならば、あいさつの日韓差については、第1に、韓国語よりも日本語の方が定型性が高い、つまり、日本人は定型的あいさつを多用するのに、韓国人はそれを用いる頻度が低いと、金香来（2000）、서정수（ソ・ジョンス 1998）などで頻繁に指摘されてきた。

第2に、韓国では相手との関係に応じて用いる表現を変えるのに対し、日本では誰にでも定型的あいさつが用いられる傾向があると報告されてきた。例を挙げると、朴英順（2003）は、韓国語では相手によって異なる表現が用いられたり表現の待遇形式が変わったりするが、日本語では定型化された単一語が主に使われると述べている。岡村（2015）では、日本人は親疎にかかわりなく非定型的表現を用いない一方、韓国人は親しい人には「누구랑　왔어？（nwukwulang wasse? 誰と来たの？）」などの非定型的表現、親しくない人には「안녕하세요？（annyenghaseyyo? 安寧ですか？）」などの定型的表現を用いると言及されている。韓国語における親疎関係の影響については、김선정・김예지（キム・ソンジョン & キム・イェジ

1　本論文は博士論文である「談話における「あいさつ」の日韓対照研究　――対面会話開始部と終結部の様相から――（東京大学大学院総合文化研究科言語情報科学専攻、2022）」の第5章に加筆・修正を加えたものである。

2　ある個人の国籍と母語は常に一致するものではないため、国籍が日本である日本人すべてが日本語母語話者とは限らない。しかし、本研究では日本語母語話者を日本人、韓国語母語話者を韓国人と称することとする。

3　韓国語のローマ字表記はイェール式、和訳は筆者によるものである。

2011)、박수란（パク・スラン 2005）などでも述べられており、例えば、前者によると、韓国人は相手と親しければ、定型的でない言葉を用い、親密度を維持・強化するのに対し、親しくなければ、定型的なあいさつを使用する比率が高いという。しかし、日本語に関する先行研究では、親疎によってあいさつを使い分けることはほとんど報告されていない。相手との関係、特に親疎による使い分けは、日本語よりも韓国語において顕著だといえるだろう。

また、上記のうち、第１の日韓差（定型性に関する差異）が生じた原因については、名倉（2005）、任・井出（2004）などで論じられているが、使用している用語は違うものの、あいさつにおける日韓差は、両国の「コミュニケーションに対する志向性」によるものだと結論づけられているように見受けられる。例えば、名倉（2005）ではあいさつの日韓差は、両国の文化コンテクストとコミュニケーション・ルールが異なるからだと述べられている。具体的には、韓国語は「表意言語」であり、あいさつする際にも伝えたいことを明確に表現するのに対し、日本語は「隠没言語」であり、あいさつでも伝えたい部分を省略し、定型化しながら、言葉の裏側を察するよう促すと言及している。任・井出（2004）は、Brown & Levinson（1987［1978］）のポライトネス理論による説明を試み、日本で好まれる定型的あいさつはネガティブ・ポライトネス、韓国語で用いられる「예쁘게　입으세요（yeyppukey ipuseyyo きれいに着て下さい)」など命令形の表現や意味が曖昧でない多様な表現はポジティブ・ポライトネスに則ったものだと述べながら、ポライトネスに関する好みの違いがあいさつにおける日韓差を生じさせたと示唆している。

だが、これらの先行研究は、定型性に関する日韓差を、理論的な枠組みや研究者の主観によって解釈しているにとどまっており、データを示しながら「あいさつの日韓差が『コミュニケーションに対する志向性』に起因している」という根拠を示すことができていない。そのほか、なぜ韓国人は日本人に比べ、相手との関係、特に親疎に応じてあいさつを使い分けるのかという第２の日韓差については全く触れられていないという問題点も指摘できる。

そこで、本研究では、インタビュー調査を実施し、①なぜ日本人は韓国人よりも定型的あいさつを多用するのか、②なぜ韓国人は相手との親疎関係によってあいさつを使い分けるのに日本人はそのような使い分けをしないのか、という点に関する母語話者の意識を探りつつ、あいさつの日韓差が生じた原因を考察してみることとする。

2. 分析の前提

2.1.　インタビュー調査の概要

前章で述べた研究目的のため、内省的方法の１つであるインタビュー調査を行うこととした。ネウストプニー J.V.（2002: 28）でも指摘されている通り、文法外コミュニケーションに関するルールや使用様相は、文法ほど自動化されておらず、その使用には、何かしらの言語管理意識が影響している。そのため、内省的方法を用い、あいさつに関する日本人・韓国人の意識を探れば、日韓差が生じた原因を明らかにできると考えた。

研究対象とするインタビュー・データは、あいさつの日韓様相を明らかにしようと実施した口頭言語産出アンケートに対するフォローアップ・インタビューを通じて収集されたものである。口頭言語産出アンケートでは、＜表１＞のような８つの場面をインフォーマントの母語で提示した後、それぞれの場面で親しい目上・親しい同年輩・親しい目下・親しくない目上・親しくない同年輩・親しくない目下（計６パターンの聞き手）にどうあいさつするか口頭で自由に答えてもらうよう依頼した。つまり、アンケートは 48 項目（8 場面 ×6 名）で構成されていたということである。

＜表1＞
口頭言語産出アンケートで提示した場面と質問[4]

場面		提示した質問
出会い	場面1 日常的に出会った時	学校に着いたところ、先に来ていた次のような日本人 / 韓国人と目があいました。何とあいさつしますか。
	場面2 偶然出会った時	東京都内 / ソウル市内にある有名デパートで次のような日本人 / 韓国人に偶然会いました。何とあいさつしますか。
	場面3 久しぶりに出会った時	夏休みの後、次のような日本人 / 韓国人と久しぶりに会いました。何とあいさつしますか。
別れ	場面4 日常的に別れる時	次のような日本人 / 韓国人と一緒に食事をしました。別れ際、何とあいさつしますか。
	場面5 相手が先に帰る時	学校の人と集まっていた席で次のような日本人 / 韓国人が先に帰ることになりました。別れ際、何とあいさつしますか。
	場面6 長らく会えなくなった時	次のような日本人 / 韓国人が田舎に引っ越すことになり、しばらく会えなくなってしまいました。別れ際、何とあいさつしますか。
特殊	場面7 訪問する時	引越しパーティーに招待され、次のような人の家を訪問することになりました。何とあいさつしますか。
	場面8 お見舞いする時	盲腸の手術をした次のような人をお見舞いに病院に行きました。何とあいさつしますか。

インタビューは、このような調査を実施した直後、アンケート調査の回答者全員に対し、インフォーマントの母語を使用して行った。アンケートに対する回答をみながら、「どうして○○とあいさつしましたか。」、「この場面で○○のような表現を使うのはどうですか。」などと主に尋ねたが、インフォーマントの反応にあわせて質問を随時変更するという半構造化インタビューの形式を採った。

なお、この2つの調査に協力したインフォーマントは、首都圏在住の日本人25名（男性：12名、女性：13名、平均年齢：28.9歳）、および、ソウル・京畿道在住の韓国人25名（男性：13名、女性：12名、平均年齢：29.6歳）であり、調査期間は2012年7～9月（韓国人を対象）、2014年8～12月（日本人を対象）であった。

2.2. あいさつの種類、および、ポライトネス理論との関連性

あいさつは、「人と人が出会ったときや別れるときに社交・儀礼的に交わす言葉や動作で、相手に敬意や情意を示して、好意的な関係を結んだり維持したりする行為」と、名倉（2005: 69-70）では定義づけられている。本研究では、このうちの「言葉」のみを研究対象とするが、서정수（ソ・ジョンス 1998: 14）では、あいさつ言葉の大部分は定型化していると述べられている。だが、私たちは常に定型化したあいさつのみを使っているわけではない。そのため、あいさつは、①「あいさつする場面において決まり文句として用いられている表現であり、言語形式が固定化しているあまり元来持っていた命題、意味が失われたり希薄化している定型的表現、および、その派生形と変種[5]」である「定型的あいさつ」と、②それ以外の言語表現である「非定型的あいさつ」に大別される。それぞれに属する表現の例は、次の通りである。

4 韓国人には韓国語に翻訳した場面や質問を提示した。

5 土屋（1998: 59）による「決まり言葉的なあいさつの言語形式とその変種」、小林（1981: 89）による「一種の符帳的合図、または極端な省略表現で、意味の上からはいわゆる「非命題的」（non-propositional）とされる表現である。」などの先行研究を参考に定義づけを行った。

定型的あいさつ：
日本語：おはようございます、お疲れ様です、おす、うぃーす[6]、バイバイ、等
韓国語：안녕？（安寧？）、오랜만이야（久しぶり）、잘　가（よく行け）、等

非定型的あいさつ：
日本語：元気だった？、どうしたの？、また連絡します、気をつけて、等
韓国語：어디　가세요？（どこ行かれますか？）、가지　마（行かないで）、等

前述の通り、任・井出（2004）は、ポライトネス理論によってこの２種類のあいさつが持つ発話効果を示し、定型的あいさつはネガティブ・ポライトネス（以下、NP）に相当し、多種多様な表現や命令形の表現、つまり、非定型的あいさつはポジティブ・ポライトネス（以下、PP）に相当すると言及したが、NP / PP とは、Brown & Levinson（1987［1978］）で提唱されたポライトネス理論における概念であり、人間が持つ以下のような基本的欲求をそれぞれ配慮する行為のことを指す。

ネガティブ・フェイス：
すべての「能力ある成人構成員」（competent adult member）が持っている、自分の行動を他者から邪魔されたくないという欲求

ポジティブ・フェイス：
すべての構成員が持っている、自分の欲求が少なくとも何人かの他者にとって好ましいものであってほしいという欲求

（Brown & Levinson 1987［1978］、田中典子監訳 2011: 80）

彼らによれば、すべての成人構成員は、互いのフェイスに配慮する必要がある。そして、他者を配慮する方法は、どちらのフェイスを考慮するかによって２つに分かれるという。その１つがネガティブ・フェイスを守る NP であり、相手と距離を置きたい時や、相手に配慮や敬意を示したい時に有用な方法である。もう１つは、ポジティブ・フェイスに対応する PP であり、相手との距離を縮めたり親密さを表明したい時に現れるものである。ところが、この２つは対極的な関係にあり、NP はネガティブ・フェイスを満たす代わりにポジティブ・フェイスを、PP は反対にネガティブ・フェイスを侵害する危険性を伴っているといわれている。

さて、あいさつに話を戻すと、定型的あいさつが NP として機能すると任・井出（2004）で解釈されている理由は、Brown & Levinson（1987［1978］）による＜表２＞のようなポライトネス・ストラテジーによって「慣習に基づき間接的であれ」、つまり、習慣的な表現は NP だとされているためだと考えられる。それに対して、「元気だった？」のような非定型的あいさつが PP だとされるのは、多くの非定型的あいさつが＜表２＞で示した PP ストラテジーと対応していることに起因しているのだろう。

6　定型的あいさつに「定型的表現、および、その派生形と変種」を含めたため、「おす」や「うぃーす」も定型的あいさつだとみなした。これらを定型的あいさつだとみなすことについては、異論もあるだろう。だが、「おはよう」などから派生した表現であるという理由により、「どうしたの？」、「また連絡します」といった非定型的あいさつと同じように扱うべきではないのではないかと考える。

＜表２＞
Brown & Levinson（1987［1978］）のポライトネス・ストラテジー[7]

NP ストラテジー	慣習に基づき間接的であれ
PP ストラテジー	H（の興味、欲求、ニーズ、持ち物）に気づき、注意を向けよ
	仲間ウチであることを示す指標を用いよ
	共通基盤を想定・喚起・主張せよ
	冗談を言え
	H に贈り物をせよ（品物、共感、理解、協力）

だが、「定型的あいさつ＝ NP」、「非定型的あいさつ＝ PP」と単純に結論づけて良いかは不明であるため、あいさつに対する母語話者の意識をインタビュー調査によって聴取し、任・井出（2004）で用いられていたポライトネス理論を援用しつつ、そのインタビュー内容を考察することとした。

3. 韓国人によるあいさつへの意識

3.1. インタビュー調査の結果

前述の通り、韓国人は、①定型的あいさつを用いる頻度が低い、②親しい人には非定型的あいさつ、親しくない人には定型的あいさつを多用する、と先行研究では言及されてきた。しかし、この２つは関連しており、「親しくない人にのみ定型的あいさつを好んで用いるため、全体的には定型的あいさつの使用頻度が低い」ようにみえるのだと考えられる。

では、韓国人はなぜ親疎関係によって、あいさつを使い分けるのだろうか。

筆者が実施した口頭言語産出アンケートでも韓国人は、親疎によって定型的 / 非定型的あいさつを選択的に用いていたが、その使い分けの理由に対するインタビュー結果が＜表３＞である。

＜表３＞
韓国人によるあいさつの使用理由（複数回答含）

親しくない人 → 定型的あいさつ使用	親しい人 → 非定型的あいさつ使用
関心・好意がないから（18 名） 礼儀を守りたいから（14 名） 会話が続かないから、会話を続ける必要がないから（4 名） 定型的あいさつは言うべきだから（3 名） よく知らない間柄だから（2 名） 見せかけだけになるかもしれないから（2 名）	関心・好意があるから（22 名） 礼儀を守らない、もしくは、定型的あいさつを言わなくても理解しあえるから（14 名） 会話が続くから、会話を続けたいから（5 名） 正直に言いあえるから（2 名） 雰囲気を和やかにしたいから（2 名）

このうちもっとも回答者数が多かったのは、相手に対する「関心・好意がないから」、「関心・好意があるから」だった。相手への関心・好意の有無が親疎によって異なり、それが定型的 / 非定型的あいさつに影響しているということである。

例えば、口頭言語産出アンケートでは、久しぶりに出会う場面３において、親しい人には「여름　휴가 잘　다녀오셨나요？（yelum hyuka cal tanyeosyessnayo? 夏の休暇、よく行ってきましたか？）」などの非定型的あいさつが頻用されていた一方、親しくない人には「안녕하세요？（annyenghaseyyo? 安寧ですか？）」という定型的あいさつがよく観察されたが、前者には持っている関心・好意を、後者には持ってい

7　NP / PP ストラテジーのうち、あいさつと関連があると思われるもののみ提示した。

ないため、簡単にあいさつしたいという以下のような回答が多くみられた。

> 친한 사람들하고는 반가우니까 더 많은 대화를 나누고 싶고 특히 여름 방학이 끝나고 봤으니까 궁금한 거죠. 그 사람이 여름 방학에 어디를 다녀왔는지 무엇을 했는지, 아니면 아팠는지 건강했는지 재미있게 잘 지냈는지 궁금하니까 뒷말을 계속 이어 가기 위해서 내가 질문을 던지는 거고. 친하지 않은 사람은 굳이 궁금하지 않으니까 말을 길게 할 필요성이 없으니까 인사만 하고 끝나는 거죠.（親しい人とは会えて嬉しいから、もっとたくさん対話を分かちあいたくて、特に夏休みが終わってから会ったので気になったんです。その人が夏休みどこに行ってきたのか何をしたのか、あるいは、具合が悪かったのか健康だったのか楽しく過ごしたのかが気になるから、次の言葉をずっと繋いでいくために私が質問を投げかけました。親しくない人はあえて気にならないから、言葉を長く言う必要性がないから、あいさつだけして終わります。）[8]
>
> 〔K24 インタビュー内容より抜粋〕

また、聞き手が先に帰る場面５でも上記と類似したインタビュー内容が多数得られた。この場面では、親しい人には「무슨 일이야? 왜 먼저 가는데?（mwusun iliya? way mence kanuntey? どうしたの？どうして先に帰るの？）」といった表現、親しくない人には「잘 가（cal ka よく行け）」などの定型的あいさつがよく使用されていた。そこで、アンケート調査後に、あいさつを使い分ける理由を質問したところ、前者には名残惜しさが生じる一方、後者にはそのような気持ちがないため、定型的あいさつを使うほかなかったという意見が多く得られた。

> 回答者：친한 친구가 먼저 가면 나랑 떨어지니까 그거에 대한 섭섭함이 있는 것 같아요. ……（親しい友だちが先に行ったら、私と離れるのだから、それに対する寂しさがあるような気がします。……）
> 筆 者：그러면 왜 친하지 않은 사람한테는 그냥 '잘 가'라고 하면서 그냥 보내 줬는지, 왜 붙잡지 않으셨나요？（じゃあ、どうして親しくない人にはただ「잘 가（cal ka よく行け）」と言いながらただ送ってあげたのか、どうして引き留められなかったんですか。）
> 回答者：친하지 않기 때문에 같이 있어도 그만, 같이 없어도 그만이니까요.（親しくないから一緒にいてもいいし、一緒にいなくてもいいからです。）
>
> 〔K3 インタビュー内容より抜粋〕

２番目に多かった回答は、「礼儀を守りたいから」と「礼儀を守らない、もしくは、定型的あいさつを言わなくても理解しあえるから」だった。以下は、その回答例である。

> (친하지 않은 사람에게는)[9] 우선 맨 처음에 '안녕?', '안녕하세요?' 안 하면 내가 그 사람한테 무례해 보일 수 있을 것 같아서. 우선 맨 처음에 인사를 하고 그 다음에 얘기를 해야 할 것 같아요. 예절을 지키는 그런…… (친한 사람에게는) 내가 굳이 그렇게 인사를 하지 않아도 그 사람은 나랑 친한 걸 서로 아니까 그거에 대해서 크게 생각을 하지 않을 것 같고, 나도 편하고 그 사람도, 상대방도 편해서 그거를 생략한다고 해서 서로 마음 상하거나 그런 걸 알기 때문에 (정형적 인사말이) 안 나오는 것 같아요.（(親しくない人には) まず最初に「안녕 (annyeng? 安寧？)」、「안녕하세요？(annyenghaseyyo? 安寧ですか？)」と言わなければ、私がその人に無礼にみえるかもしれないから。まず最初にあいさつをして、その後に話をしなければならないと思います。礼節を守る、そういう……（親しい人には）

8 インタビュー内容の文字起こしは、読みやすさを考慮して、フィラーや言い間違いなどを省く「ケバ取り」で行った。
9 括弧内はインタビューの内容がわかりやすくなるよう、筆者が注意書きした部分である。

私があえてそうあいさつをしなくても、その人は私と親しいことを互いに知っているから、それについて大きく考えないだろうし、私も楽で、その人も、相手も楽だから、それを省略したからといって互いに心が傷ついたりそういうことを知っているから（定型的あいさつが）出てこないんだと思います。）
〔K12 インタビュー内容より抜粋〕

親しくない人には無礼にみえないよう礼儀を示せる定型的あいさつを用いたいが、親しい人には礼儀を示さなくても構わないため、その使用が省略できるということであり、「礼儀の必要性 / 不要性」がそのまま「定型的あいさつの必要性 / 不要性」に直結していることが窺えた。

さらに、韓国人のなかには、「定型的あいさつ＝ 礼儀・格式」だと明言している次のような人もみられた。

回答者：친한 사람한테는 격이 없이 대하지만 친하지 않은 사람한테는 마음이 없지만 예의를 갖추는 것 같아요.（親しい人には隔たりなく接しますが、親しくない人には心はないけど礼儀を正すように感じました。）
筆　者：마음은 없어요？（心はありませんか。）
回答者：그 사람을 좋아하지 않지만 예의를 갖춘다고 할까요？……격식이라는 것은 상대방과의 거리를 더 멀게 만드는 역할을 하는 것 같애요.（その人が好きではないけど、礼儀を正すとでもいいましょうか。……格式というものは、相手との距離をもっと遠ざける役割をすると思います。）
筆　者：격식이요？（格式ですか。）
回答者：다양한 인사가 아니라 한국사람이 좋아하는 인사예절은 어느 정도 사람과 사람과의 간의 거리를 유지하는 역할을 하는 것 같애요.（多様なあいさつではなく、韓国人が好むあいさつの礼節は、ある程度、人と人との間の距離を維持する役割をしていると思います。）
〔K10 インタビュー内容より抜粋〕

このインフォーマントによれば、礼儀・格式は相手との距離を維持したり遠ざけるものであり、「あいさつの礼節」である定型的あいさつは、礼儀・格式であるという。インタビューのなかでは、ほかの韓国人と同様、親疎関係によって礼儀の必要性 / 不要性を判断するとも述べていたが、「定型的あいさつ＝ 礼儀・格式＝ 距離感の表示」であるため、それが必要だと思われる親しくない人にのみ定型的あいさつを使用するのだろう。

3.2.　韓国人の意識に対するポライトネスからの考察

以上、「関心・好意」、「礼儀」という２つの観点からポライトネス理論に照らしあわせて考えてみると、韓国人は、相手との距離を見積もりつつ、ポライトネスを適切に使い分けていると解釈される。

このようなポライトネスの使い分けは、韓国語以外の言語でもみられる現象だろうが、非定型的あいさつによって実現される PP は、相手に親近感を与える一方、聞き手との距離を小さく見積もる踏み込みが要求される。そのため、滝浦（2008: 37）で指摘されている通り、相手が相応の距離を維持しておきたいと感じている時には「図々しい」、「馴れ馴れしい」という印象を持たれる。ポジティブ・フェイスと相反する概念であるネガティブ・フェイスを脅かすためである。よって、相手との距離が離れている親しくない人に対しては、そのような発話をすることができず、無礼にみえないよう NP に相当する定型的あいさつを選択するのだろう。

一方、親しい人に対しては、相互間の距離を互いに小さくみなしているだろうという信頼があり、ネガティブ・フェイスを侵害しても無礼にみえる憂慮がないと判断していることがインタビューから明らかになった。そのため、礼儀に直結する NP を避けることができたのである。また、次の２つのインタビュー内容は、「非常に親しい人が自分に『안녕？（annyeng? 安寧？）』とあいさつしてきたらどう思うか」という質問に

対する回答を示したものである。これらの回答からも窺える通り、NP に相当する定型的あいさつは、相手への関心や好意の欠如を示してしまうため、仮に親しい人にこのような表現を用いた場合、「よそよそしさ」や意図していない誤解を与えてしまうこともあるようである。

> 지금 질문자께서 질문을 해 줬을 때 딱 드는 느낌은 '어, 나랑 안 친한가?' 라는 생각이 좀 들기도 해요. (今、質問者が質問してくれた時、パッと思った感じは「お、僕と親しくないのかな?」という気持ちがちょっとしたりします。)
>
> 〔K1 インタビュー内容より抜粋〕

> 상대가 피곤하다든지 자기 상태가 안 좋다든지 나한테 언짢은 게 있다든지 그렇게 느낄 거 같애요. (相手が疲れているとか、自分 (相手) の状態が良くないとか、私に不快に思っていることがあるとか、そう感じると思います。)
>
> 〔K3 インタビュー内容より抜粋〕

ここまで、韓国人は親疎によってあいさつを使い分けているという前提で議論を進めてきたが、親しい人に対しても定型的あいさつを用いていたインフォーマントの意識も紹介したいと思う。

以下は、親しい人にも「오랜만이야 (olaynmaniya 久しぶり)」、「잘 가 (cal ka よく行け)」など定型的あいさつを用いていた人から得られた内容である。

> 친한 사람들한테는 '안녕' 이라는 뜻을 내포하는 여러 가지, 외래어가 될 수도 있고 한국말이 될 수도 있는데, 그 말을 친한 사람들끼리는 예를 들면 억양을 딱 정식 억양을 써서 말하지 않게 되는 거 같아요. 그 억양으로도 친밀함을 표현하게 되는 거 같아요. 예를 들면 똑같이 '안녕하세요?' 라고 해도 친한 사람들한테는 '(낮은 목소리로) 안녕하세요-?' 이런 식으로. 그리고 친하지 않은 사람들한테는 그냥 정식으로 '안녕하세요?' 이렇게. (친하지 않은 사람에게는) 예의에 벗어나지 않아야 된다라는 그런 마음이 있는 거 같아요. (親しい人たちには「안녕 (annyeng 安寧)」という意味を含む色々な、外来語かもしれないし、韓国語かもしれないけど、その言葉を親しい人同士では、例えば、抑揚を正式な抑揚で言わないと思います。その抑揚でも親密さを表すんだと思います。例えば、同じ「안녕하세요? (annyenghaseyyo? 安寧ですか?)」だとしても、親しい人たちには「(低い声で) 안녕하세요- ? (annyenghaseyyo - ? 安寧ですかー?)」こんな風に。そして、親しくない人たちには、ただ正式に「안녕하세요? (annyenghaseyyo? 安寧ですか?)」こうやって。(親しくない人には) 礼儀から外れてはならないという、そんな気持ちがあるんだと思います。)
>
> 〔K4 インタビュー内容より抜粋〕

親しい人にも「안녕? (annyeng? 安寧?)」のような定型的あいさつが使用可能ではあるものの、そのような場合には、礼儀に反しない正式な抑揚ではなく、普通ではない抑揚でそれを言うということである。定型的あいさつを正式ではない抑揚で発することによっても、礼儀から外れ、親しみを表すことができるのだと考えられる。何かしらの方法で「礼儀・格式」から外れさえすれば、親しみを表すことができ、その手段として定型的あいさつの不使用ではなく、通常ではない抑揚という方法を選び、あいさつしようとしたということなのだろう。

韓国人にとっての定型的あいさつは「礼儀・格式」であり、NP として基本的には機能しているが、「通常、定型的あいさつはこのような言葉を、このような言い方でするもの」というようなイメージが、韓国人の間で共有されているのかもしれない。そして、正式ではない抑揚を用いたりしながら、そのイメージから外れさえすれば、親しみを帯びると認識されているのではないだろうか。

4. 日本人によるあいさつへの意識

4.1. インタビュー調査の結果

先行研究によれば、日本人は、定型的あいさつを多用するといわれている。筆者が行った口頭言語産出アンケートを通じても、日本人は定型的あいさつの使用頻度が高いということが確認されたほか、非定型的あいさつを用いる場合でも、それを単独で用いるのではなく、「おはよう。最近どう？」のように定型的あいさつを伴ってよく使用されていた。そこで、インタビューを実施したところ、25 名中 15 名から、会話の前後、もしくは、出会い頭や別れ際には、誰に対してでも必ず定型的あいさつを使うという次のような回答が得られ、なかには、定型的あいさつがないと、不快感や驚きなど否定的な感情を覚えるインフォーマントもいることがわかった。

筆　者：絶対「久しぶり」っていうのを言ってから、「何してたの？」とか「元気？」とかおっしゃっていたんですけど、「久しぶり」なしで、突然会って「あ、夏休み何してたの？」とか「元気だった？」っていうのは大丈夫そうですか。

回答者：大丈夫かもしれないけど、たぶんかな、絶対的ぐらいに久しぶりに会ったら「久しぶりだね」って言っちゃうかな。……（定型的あいさつを）つけなきゃいけないとは思ってないけど、でも、いきなりその人の近況を聞いたりとか何かする前に、とりあえず「久しぶり」のあいさつはするのが習慣。「おはよう」も「久しぶり」もたぶん会った場合もメールとかでも、ちょっとチャットとかになると時々用件だけ入れちゃう時、チャットとかラインとかだとあるけど、基本的に会社のメールとかでも何でも、まずあいさつしてから、用件とか話したいことに入っていく。

筆　者：……誰かが J9 さんに会って突然「久しぶり」とか言わないで、「夏休み何してたんですか？」みたいな感じだったらどういう風に思われますか。不快な感じとか不思議な感じとかされますか。

回答者：不快は感じないけど、親しかったら全然自然に「何やってたの？」って言われたら、あの「こうこうしてたよ。何やってたの？」みたいな。そしたら言わないかもね、「久しぶり」って、あっちからも「わー」って来てたら。

筆　者：特に嫌な感じとか、そういうのは？

回答者：ない。ただ親しくない人にいきなり、あの、会った瞬間に「何してたの？夏休み」とか言われたら、「え、なんか関係ありますか？」みたいな、なるかもしれない。

〔J9 インタビュー内容より抜粋〕

では、なぜ日本人は、相手を問わず定型的あいさつを使用するのだろうか。

その理由を探るため、インタビューでは、「なぜそのようなあいさつを使ったのか」、「定型的あいさつを除いて非定型的な表現のみを言うのはどうか」という質問を投げかけたほか、「別の表現を使うのはどうか」と特定の非定型的あいさつ[10]を例に挙げながら尋ねてみた。非定型的あいさつの使用を避けたいという気持ちから、定型的あいさつを頻用している可能性もあると考えたためである。

＜表４＞は、そのような質問に対する回答をまとめたものである。まず、定型的あいさつを使用した理由からみてみると、もっとも回答者数が多かった答えは、「そうあいさつするのが習慣・自然だから」であった。

10　例えば、学校で友人に出会ったという場面 1 では、韓国人が「일찍　왔네（ilccik wassney 早く来たね）」などをよく用いていたため、「早いね」などという表現はどうかと尋ねた。

＜表４＞
日本人によるあいさつの使用 / 非使用理由（複数回答含）

定型的あいさつ使用	非定型的あいさつ非使用
そうあいさつするのが習慣・自然だから（14 名） その場面に適したあいさつだから（９名） 特に理由はない（５名） あいさつで気づいてもらいたから（３名） 会話の前後にあいさつはするから（２名） 周りもそう言うから、同伴者がいると思ったから（各１名）	会話の途中に言う表現だから（12 名） 立ち入ったり詮索したりしたくないから（８名） 恥ずかしい、気持ち悪いから（５名） 違う状況や、詳しい状況を想定していなかったから（４名） そう言わなくても通じるから（４名） 関心がないから（４名） 自分の意図などを誤解されるかもしれないから（２名） 言わない理由はわからない（３名）

以下は、その回答例である。相手が先に帰る場面５で「お疲れ様。気をつけて。」、「お疲れ様。また今度ね。」のように発していた J3 は、必ずしも定型的あいさつが必要ではないが、「お疲れ様」と言うのが自分の癖・習慣であり、定型的あいさつを交わすことが自分にとっての「普通」であると述べていた。

筆　者：「お疲れ様」抜かして「気をつけてね」だけとか「また今後ね」だけはどうですか。
回答者：「また今度ね」だけの時もあるんですけど、集まりの時って、結構何でも遊びとか、真面目なことで集まってたりとか、両方別れるとか、先に帰るっていう時には「お疲れ」って言うのが癖になってるっていうか、それが今、私の普段の生活のなかで普通になっちゃってるので言いました。

〔J3 インタビュー内容より抜粋〕

２番目に多かった定型的あいさつの使用理由は「その場面に適したあいさつだから」、３番目に多かったのは「特に理由はない」であり、以下は、その回答例である。この２つは、定型的あいさつを使う理由を深く考えていないという点では類似していた。だが、「その場面に適したあいさつだから」に分類される J25 による言及は、出会ったり別れたりする時間帯や場面を考慮しながら、その場面にふさわしい定型的あいさつを選択して使っているという点で、J19 の意見と違いがみられた。

筆　者：「おはよう」とかを言いたい理由って何かありますか。
回答者：朝っていう設定？
筆　者：例えば朝だったら。
回答者：そのへんは深く考えず、朝だから「おはよう」、昼だから「こんにちは」、夜だから「こんばんは」みたいな感じで、そうですね、はい。

〔J25 インタビュー内容より抜粋〕

筆　者：さっきと同じ質問なんですが、「久しぶり」を入れる理由とかはありますか。
回答者：理由は特にないです。

〔J19 インタビュー内容より抜粋〕

続いて、非定型的あいさつを用いなかった理由のうち、１番多かった回答は、以下のような「会話の途中に言う表現だから」であった。「こんにちは」などの定型的あいさつではなく、「どうしたの？」のような非定型的あいさつを用いることもあるものの、そのような表現は会話をしている最中に使いたいという回答が 25 名中 12 名から得られた。「会話の前後、もしくは出会い頭や別れ際に必ず定型的あいさつを使っている」という本節の冒頭で紹介した意見とも一致するものであるが、以下からも「定型的あいさつ（出会い頭）→

非定型的あいさつ」、もしくは、「非定型的あいさつ → 定型的あいさつ（別れ際）」という順序性を日本人が重要視しており、定型的あいさつによる明示的な会話の開始表示・終了表示が求められていることが確認された。

筆　者：2番目の状況で、偶然会ったので、全員に「こんにちは」だったんですけど、これも「どうしたの？」とか「何買いに来たの？」、「誰と来たの？」とかそういうあいさつは、どうですか。
回答者：「こんにちは」の後に続く言葉で、そういう言葉は出ると思ったんですけど、最初に何て声かけるかっていうと、あいさつからかなと思ってそう（「こんにちは」と）言いました。
〔J19 インタビュー内容より抜粋〕

次に多かった非定型的あいさつを使用しない理由は、「立ち入ったり詮索したりしたくないから」であった。例えば、偶然に出会ったという場面2に関して「何買いに来たの？」、「誰と来たの？」といった非定型的あいさつを使うのはどうかと尋ねたところ、そのような表現は詮索しているようにも捉えられるため、親しい間柄でも使用を控えたいという回答が8名から得られた。

筆　者：こういった場面で「何しに来たの？」とか「何買いに来たの？」、「誰と来たの？」とかそういうあいさつはどうですか。
回答者：デパートなんで買い物に来ているのは当然として、そこまで「何買ったの？」とかまでは、詮索しているみたいで、僕はしないです。
（中略）
筆　者：詮索したくないとか、詮索すると嫌だろうっていうのは……親しい人でもそうですか。
回答者：親しい人、それでもちょっと考えます。親しくても聞かれたくないことだったらどうしようとか考えてしまうので。
〔J25 インタビュー内容より抜粋〕

4.2.　日本人の意識に対するポライトネスからの考察

任・井出（2004）では、日本で多用される定型的あいさつは、NP に則したものだと言及されていた。だが、ここまでみてきたように、インタビュー調査からは「礼儀を考えて」、「相手を配慮して」などネガティブ・フェイスを考慮したために定型的あいさつを使用しているという意見は1例も得ることができなかった。したがって、「日本人による定型的あいさつの使用＝ NP」と短絡的に考えるのはやや早計かもしれない。

2.2. でも述べたように「慣習に基づき間接的であれ」が NP ストラテジーとして挙げられており、慣習的な定型的あいさつを用いること自体が、NP に則っている証だとも捉えられる。また、非定型的あいさつを用いない理由として、「立ち入ったり詮索したりしたくないから」という点を挙げたインフォーマントもいたことから、ネガティブ・フェイスを考慮しながら非定型的あいさつを回避し、その代替手段として定型的あいさつを用いている日本人も実は一定数いるのではないかとも推測される。定型的あいさつを使用している理由として「そうあいさつするのが習慣・自然だから」などと答えた人のなかには、相手のネガティブ・フェイスを侵害してしまう非定型的あいさつを避けたいという潜在的な意識によって、NP に当たる定型的あいさつを頻用しているにもかかわらず、そのような自分の言語使用を意識化、あるいは、言語化できなかった人が含まれているのではないか、ということである。

しかし、インタビュー内容をみた限りでは、NP / PP という物差しのみで日本人の定型的あいさつ使用を解釈しようとするのではなく、宇佐美（2001; 2008）で論じられている「無標ポライトネス（以下、無標P）」も援用しながら考察した方がいいのではないかと判断される。

ここでいう無標Pとは、ディスコース・ポライトネス理論[11]のなかで述べられている概念であるが、そのなかで宇佐美（2001; 2008）は、Brown & Levinson（1987［1978］）によるポライトネスは、相手のフェイスを侵害する行為（依頼・褒めなど）を行う時、そのフェイス侵害度を軽減させるために採るストラテジーだと説明しつつ、日常会話にはフェイス侵害行為が生じていないポライトネスもあると言及している[12]。そして、そのような「あって当たり前で、それが現れないときに初めてそれがないことが意識され、ポライトではないと捉えられる（宇佐美2008: 160）」ものを無標Pと呼んでおり、そのような無標の状態、つまり、もっとも自然で最適な状態から逸脱した場合に、①親しみを表すなどのプラス・ポライトネス効果、②命題内容を強調するなどの言語的談話効果、③不快感を与えるマイナス・ポライトネス効果、のような何らかの特別な効果を生むとしている。

このような無標Pも念頭に置きながら、日本人によるインタビュー内容を再度みてみると、「そうあいさつするのが習慣・自然だから」というもっとも回答者が多かった定型的あいさつの使用理由は、定型的あいさつがあるのが当然という観点から無標Pと相通じるところがあり、「特に理由はない」も定型的あいさつを交わすのが無標であるために、なぜそう言うのか理由がわからないという意味でそう答えていたと推察される。また、「その場面に適したあいさつだから」という回答も、場面設定が考慮されてはいるものの、以下に再掲した通り、「（あいさつを使う理由を）深く考えず」と述べられていたため、無意識のうちに定型的あいさつが使われるという無標Pと一致している。

筆　者：「おはよう」とかを言いたい理由って何かありますか。
回答者：朝っていう設定？
筆　者：例えば朝だったら。
回答者：そのへんは深く考えず、朝だから「おはよう」、昼だから「こんにちは」、夜だから「こんばんは」みたいな感じで、そうですね、はい。

〔J25インタビュー内容より抜粋（再掲）〕

したがって、無標Pという概念を分析の枠組みに含めた上で、改めて考えてみると、結局のところ、これらの回答は、「会話の前後に定型的あいさつによって、言葉を交わすことが無標Pだから定型的あいさつを使用した」と解釈できる。日本人にとっての定型的あいさつとは、NPとしての性格を帯びつつも、無標Pとして強く機能しているものであるため、相手との親疎を問わず、定型的あいさつを交わしていると考えられるだろう。

ここまで日本人が定型的あいさつを誰にでも用いている原因を探ってきたが、はたして、日本人はあいさつする際、相手との関係を全く考慮していないのだろうか。

インタビューの結果、相手との親しさによっては、あいさつを使い分けていないことが、アンケートを通じて自己分析できたと25名中7名が述べており、特に次のようなインフォーマントは、親しい友人に対してでも砕けた表現を用いないのは、あいさつを交わす際に言う内容を重要視していないからだと言及していた。あいさつすること自体が重要なのであって、何を言うかには注意を払っていないということなのだろう。

11　ディスコース・ポライトネス理論は、Brown & Levinson（1987［1978］）によるポライトネス理論を修正したものである。①文レベル・発話行為レベルでポライトネスを捉えているため、各言語の構造の違いが大きく影響してしまう、②「特にポライトでもないが、失礼でもない」言語行動やインポライトネスの位置づけが明確に示されていない、③言語行動に対する聞き手の観点、および、話者同士の相互作用の観点がポライトネス理論には組み込まれていない、などの問題点を克服すべく、談話における基本（無標）状態を同定し、そこから逸脱しているかどうかという「動き」をみることによって、ポライトネス効果を相対的に分析することが提案されている。

12　ただし、滝浦（2008: 29-30）は、Brown & Levinson（1987［1978］）のポライトネスについて「会話の場において表現・伝達される、主として相手のフェイスを侵害することに対する軽減的・補償的な言語的配慮のこと」と説明しながら、事実をただ述べているようにみえる場合でも、状況や相手の受け取り方次第では、相手への非難だとみなされることもあるため、ほぼすべての言語行為は、潜在的なフェイスリスク（フェイスを侵害する可能性）を持っていると言及している。

筆　者：今、1から8までアンケートに答えていただいたんですけど、アンケートに対して、全体的な感想をお願いします。
回答者：あんまりボキャブラリーがないなと思いました、自分の。あと、あんまり親しいか親しくないかっていうことに関して分けてないなあと思う。っていうのは、親しい人に対して、すごく砕けた言い方をしないかなあ、あいさつでは。
筆　者：あいさつでは砕けた言い方をしない理由とかってありそうですか。
回答者：えー。
筆　者：普段の会話ではそうでもないですか。
回答者：たぶん中身とかが砕けた話をしていて、あいさつに関しては、特に重要視をしていないというか、そうですね。

〔J16 インタビュー内容より抜粋〕

しかし、なかには、誰にでも定型的あいさつであいさつをしてはいるものの、どのような種類の定型的あいさつを使うかという点に関しては、相手との関係を考慮している人も見受けられた。

例えば、次は、親しい目下と同年輩には「おす」もしくは「うぃーす」、親しい目上や親しくない人には「おはよう」、「おはようございます」と言っていた人から得られた内容である。このインフォーマントは、このように定型的あいさつを使い分けていた理由として、前者には「ちゃんとあいさつしなきゃいけない」、定型的あいさつを使って失礼がないようにしたい気持ちを持っている一方、後者にはそれを使わなくても失礼になる危険性がないことを挙げていた。

筆　者：ほかの人には「おはようございます」とか「おはよう」とか言っていて、AとB（親しい目下と親しい同年輩）だけは「おす」とか「うぃーす」って言ってたんですけど、この人たちに「おはよう」って言わなかった理由とかありますか。
回答者：何だろう。雰囲気で伝わるかな、みたいな。
筆　者：雰囲気で伝わるかな？「おはよう」とかは言いたくないですか、こういう人には。
回答者：いや、言うは言うんですけど、「おはよう」って改めて言うほどじゃないかな、みたいな。
筆　者：じゃあ、違う人には「おはよう」って改めて言わなきゃいけない理由とかあったんですかね。
回答者：目上はちゃんとあいさつしなきゃいけないかな、みたいなのはあるし、親しくない人は失礼がない程度に。

〔J23 インタビュー内容より抜粋〕

また、別れの場面で1度も「さようなら」を用いていなかったJ11は、「さようなら」は「今生の別れ」という意味も暗に含んでいるほか、学生と教師という間柄で用いられる表現であるため、日常生活ではあまり交わさないと答えていた。

筆　者：「さようなら」とかそういった教科書で出るようなあいさつ表現が1回も出てこなかったんですけれども、そういった言葉は普段は使われませんか。
回答者：あんまり使わないですね。院生室にいる時も「お疲れ様でーす」とか「お先に失礼します」。「さようなら」は違うな。「お先に失礼します」、「お疲れ様です」ですね。
筆　者：その言葉を使わない何か理由とか言いづらい理由とか。
回答者：なんか学校の先生みたい。
筆　者：学校の先生みたい？
回答者：学校の先生と生徒みたい。
筆　者：学校の先生と生徒みたい。
回答者：で、あと、言葉を教えるお仕事もするので、そういう時にはあいさつとして「さようなら」

は教えるんだけれども、どちらかというと「さようなら」って言うと今生の別れみたいな。

筆　者：今生の別れ。

回答者：今生の別れみたいな、そんな感じもする。けれど、学生には「さようなら」って言って、「先生、さようなら」、「うん、さようなら」とかって。だから学校みたい。そういうのもあって学校的だったり、今生の別れみたいに思ったりして、たぶんそれで使わないのかも。

〔J11 インタビュー内容より抜粋〕

このように日本人もあいさつを用いる時、様々な要因を考えていないわけではないようだが、そのような意識も定型的あいさつと非定型的あいさつの使い分けには結びつかないようである。その代わり、定型的あいさつを使う際に、どのような種類の定型的あいさつを使うのか、例えば「おっす」を使うのか、「おはよう」を使うのか、などに関しては、相手との関係など自分を取り巻く状況によって表現を選択しているのだと考えられる。

それでは、前述したように「日本人にとっての定型的あいさつは無標 P として強く機能している」のであれば、なぜ様々な要因による使い分けが行われるのだろうか。

宇佐美（2001: 32）は、談話の基本状態、つまり、無標の状態というものは、「個々の会話ごとに、話者間で交渉されつつ形成されていった「談話」を分析することによって、はじめて同定できるもの」だと言及している。つまり、無標の状態とは一般化できるものではなく、個々の談話をつぶさに調べながら、その 1 つの談話における無標とは何なのかを同定しなければならないものだということである。例えば、仲の良くないクラスメートがおり、その人とは毎日、何のあいさつも交わしていないとしよう。「あいさつしないこと」が無標の状態だといえる。しかし、そのクラスメートに対して何かしらの誤解をしていたことに気づき、ある日、「これからは仲良くしたい」という気持ちを持って、笑顔で「おはよう」とあいさつしたとしたら、その「おはよう」という定型的あいさつは、この場合に限り、PP として働く。また、ある職場の同僚 2 人によるあいさつについても考えてみたい。この 2 人は仲が良く、私的に会った際には「あー、疲れた」などと非定型的あいさつで会話を始めるが、会社で会った時は、「公的な場面」という要因が作用して、「おはようございます」と、いつもあいさつしているとしよう。この場合、私的な場面では非定型的あいさつ、公的な場面では定型的あいさつを使うことが無標の状態となるため、私的な場面で「おはようございます」と言ったら、悪ふざけをしている、のような発話効果が生じると推測される。どのような状態が無標だと捉えられるかは、人間関係によっても異なるだろうし、場面の制約なども受けて変わることが予想されるため、1 つ 1 つの談話における無標の状態を同定してからでないと、その発話が生み出す効果を正しく判断することはできない。

したがって、本研究では「日本人にとっての定型的あいさつは無標 P として強く機能している」と主張したが、日本人によって発せられるすべての定型的あいさつが無標 P だという意味ではなく、日本人に産出される多くの談話では、定型的あいさつが無標のものとしてみなされているという意味でそう言及したのである。

定型的あいさつの発話効果は、性別、年齢、地域、場面などによっても、ある定型的あいさつに対する意識が異なる可能性がある。「日本人によって産出される多くの談話では、定型的あいさつが無標のものとしてみなされている」ということを念頭に置きながらも、1 つ 1 つの談話をみながら、その基本状態（無標の状態）を同定し、それぞれの定型的あいさつの発話効果を判断すべきだと考える。

5. まとめと今後の課題

本研究では、インタビュー調査によってあいさつに対する日本人・韓国人の意識を探り、あいさつの日韓差が生じた原因を明らかすることを試みた。

その結果、まず「なぜ日本人は韓国人よりも定型的あいさつを多用するのか」という問いについては、日本語の定型的あいさつが主に無標 P として機能するためだということを明らかにすることができた。会話の前後に定型的あいさつを交わすこと自体が、日本人にとっては当然であるため、誰に対してでも定型的あい

さつを用いる結果、自然とその使用頻度が高くなるのだと推測される。

一方、韓国人は、親しくない人にのみ定型的あいさつを用いるため、全体としては、日本人よりも定型的あいさつを使用していないようにみえるのだと考えられる。

「なぜ韓国人は相手との親疎関係によってあいさつを使い分けるのに日本人はそのような使い分けをしないのか」という２つ目の問いについては、相手に対する関心・好意、礼儀の必要性の有無が親疎関係によって異なると韓国人が思っており、親しくない人にはNPである定型的あいさつ、親しい人にはPPである非定型的あいさつを使っていることが明らかになった。親しい人に定型的あいさつを用いていた人も見受けられたが、そのようなインフォーマントにインタビューを行ったところ、韓国人にとっての定型的あいさつは「礼儀・格式、NP」として基本的には機能しているが、「通常、定型的あいさつはこのような言葉を、このような言い方でするもの」というようなイメージが韓国人同士で共有されているため、定型的あいさつを交わしたとしても、それが、そのイメージから外れるものだった場合――例えば、非常に低い声で「안녕？（annyeng? 安寧？）」と言うなどの場合――には、「礼儀・格式」から逸脱したとみなされ、親しみを帯びる可能性があることも示すことができた。

また、日本人は、日本語における定型的あいさつが無標Pとして働いているため、韓国人のように、相手に応じたあいさつの使い分けをしていないと分析された。しかし、日本人も相手との関係を全く考慮していないわけではなく、どのような種類の定型的あいさつを使うかという点に関しては、相手との関係を考慮する人もいることがわかった。

つまり、韓国人の場合は相手との関係が「定型的あいさつと非定型的あいさつのどちらを使うのか」に影響を及ぼすが、日本人の場合は相手との関係が「定型的あいさつを使うには使うが、どのような種類の定型的あいさつを用いるのか」に反映されるという違いがあるということになる。

本研究では、インタビュー・データをBrown & Levinson（1987〔1978〕）によるポライトネス理論とその修正理論から解釈し、あいさつの日韓差が生まれた原因を明らかにしようと試みた。だが、数十名の母語話者からしか聴取を行えず、「データからこのような考え方ができる」ということを提示することしかできなかった。今後、量的調査を行い、本研究における言及を立証していきたい。また、日韓のあいさつ様相を左右する要因は１つとは限らず、いくつかが複雑に絡みあいながら日韓差が生じた可能性もあるため、多方面から原因を探っていくことも今後の課題としたい。

参考文献

任栄哲・井出里咲子（2004）『箸とチョッカラク ―ことばと文化の日韓比較』大修館書店

宇佐美まゆみ（2001）「談話のポライトネス ―ポライトネスの談話理論構想―」『談話のポライトネス』（第7回国立国語研究所国際シンポジウム報告書），国立国語研究所，9-58

宇佐美まゆみ（2008）「相互作用と学習 ―ディスコース・ポライトネス理論の観点から」西原鈴子・西郡仁朗（編）『講座社会言語科学 第4巻 教育・学習』ひつじ書房，150-181

岡村佳奈（2015）「出会いの場面における日韓挨拶表現の対照」『韓国学のフロンティア』1，4-25

金香来（2000）「あいさつ行動の日韓比較研究」奈良女子大学大学院博士学位論文

小林祐子（1981）「日本人とアメリカ人の挨拶行動 ―出会いの挨拶―」『東京女子大學附屬比較文化研究所紀要』42，87-110

滝浦真人（2008）『ポライトネス入門』研究社

土屋頼子（1998）「言語行動を構成する要素とその機能 ―出会いのあいさつを中心に―」『筑波応用言語学研究』5，57-69

名倉康司（2005）「日本語教育におけるコミュニケーション・ルールの教え方 ―日韓挨拶表現のギャップに関する考察に基づいて―」『愛知産業大学日本語教育研究所紀要』2，69-80

ネウストプニー J.V.（2002）「データをどう集めるか」ネウストプニー J.V.・宮崎里司（編）『言語研究の方法 言語学・日本語学・日本語教育学に携わる人のために』くろしお出版，15-33

김선정・김예지［キム・ソンジョン & キム・イェジ］（2011）‘중국인 한국어 학습자의 인사 화행 연구

―언어적・비언어적 접근―’“언어와 문화”7（2），55-79
박수란 [パク・スラン] (2005)‘한국어 교육을 위한 한국어 인사 표현 연구’이화여자대학교 대학원 석사학위논문
朴英順（2003）‘한일양국인의 인사행동에 관한 사회언어학적 연구’계명대학교 대학원 박사학위논문
서정수 [ソ・ジョンス] (1998)‘한국어, 일본어, 영어 및 중국어의 인사말 비교 연구’“비교한국학”4, 13-36
Brown, P. & S. C. Levinson（1987［1978］）*Politeness: Some Universals in Language Usage*. Cambridge University Press（田中典子（監訳）2011『ポライトネス 言語使用における、ある普遍現象』研究社）

【Research Brief】

Student Perception of Online Homework in a Japanese University EFL Course

Alex Blumenstock

Abstract

A study regarding student perceptions of online homework was conducted with an "English through Movies" course that had 25 students in a Japanese university. After two lessons, the course shifted from face-to-face to online lessons due to the COVID-19 pandemic. When classes were face-to-face, students recorded their homework answers in their textbooks, but when classes were online for the remainder of the semester, students instead submitted textbook answers online via Google Forms. At the end of the semester, a survey was conducted in which students indicated overall positive feelings about online homework. Student responses indicated that the system improved their motivation and understanding of the materials; they cited feedback as a positive factor, both the automatic feedback they received after submission as well as the targeted feedback given in class. However, although students preferred online homework in an online context, students preferred recording homework answers in their textbooks in a face-to-face context. Student responses underscore the importance of carefully choosing when and how to implement online learning tools to maximize learning outcomes and minimize dissatisfaction.

Keywords: *online homework, computer-based learning, feedback, motivation, assessment*

Introduction

Online homework is a relatively new tool in education, but the many different types of online homework make it difficult to understand its efficacy and how students actually feel about being asked to do it. This study aims to provide insight into student perceptions of completing their textbook homework online via Google Forms.

This study was conducted with an "English through Movies" course that had 25 students, who met for 100 minutes once a week for 14 weeks and used the textbook *Social Issues in the Movies – Discussion & Role-Play: Level 4* (Dyer, 2021). Although the first two lessons of the course were conducted face-to-face in a classroom, the remaining 12 lessons were conducted online in real-time via Zoom due to the COVID-19 pandemic. As a result, the original plan to have students complete homework within the textbook was altered. As soon as classes shifted online, students were likewise required to submit their textbook homework online instead of simply recording their answers in the textbooks.

Method

The textbook homework consisted of three sections that served as a review of the previous lesson and a preview of the next lesson. The first section contained five sentences split in half that were to be matched together. The second section was fill in the blanks with 10 sentences; a synonym for each blank was provided in the textbook, for example, "It is not possible to estimate its ___. (worth)" (p. 41). An appropriate response might be "value." The third section supplied a list of 5 idioms to be inserted into 10 fill-in-the-blank sentences. Completing these exercises in the book was

a relatively straightforward task. However, shifting the task online had the potential to make it more burdensome for students, as students had to both complete the textbook task and contend with an additional step of inputting their answers online.

Thus, when creating the Google Forms for students to input their homework, special consideration was given to making the system as easy to use as possible. The first section, matching, became a small grid, with numbers ascending vertically and letter answers forming the columns. The second section, synonyms, changed from open-ended to close-ended so that students could verify their answers. Each of these 10 questions had a dropdown menu of answer options. The third section, idioms, used a grid format similar to the style of the first matching section.

As soon as a student completed and submitted their homework, they received automated feedback that revealed the correct answers within Google Forms. In this way, they had the opportunity to verify whether they had understood the vocabulary and language points of the preceding lesson before completing further practice with the same points in the next lesson. In addition, they received an email copy of their responses, which served as both an acknowledgement that their work had been received as well as a document they could review when desired.

At the end of the final lesson, students were asked to complete a survey via Google Forms about the textbook homework they completed for the course (see Appendix A). Four binary items on the survey asked preferences between recording homework answers online and in the textbook regarding motivation, understanding, and within the contexts of online and face-to-face classes. There were also two Likert-type questions modeled after the questions in another study about online homework by Wooten and Dillard-Eggers (2013). One asked students whether online homework was much better or worse for understanding, and the other asked students to rate their experience using online homework. After each survey item, an open-ended "Why?" question provided students an opportunity to explain their answers. The means and standard deviations of the quantitative items were then collated with corresponding qualitative items to provide the analysis, as described below.

Results and Discussion

Of the 25 students in the class, 17 responded to the survey (see Appendix A). Student responses to binary survey items are presented in Table 1 below.

Table 1: Student responses to the binary items on the survey (n = 17)

No.	Item	M	SD
1	I felt more motivated to do my homework when... 私は…の時、宿題に対してやる気がより出ました。	0.76	0.44
2	I best understood the textbook's content when... 私は…の時、テキストの内容を最も理解できました。	0.59	0.51
3	In an online class, which do you think is better? オンライン授業において、どちらがより良いと思いますか。	0.82	0.39
4	In a face-to-face class, which do you think is better? 対面授業において、どちらがより良いと思いますか。	0.41	0.51

Note: 0 = Recording my answers for textbook exercises in the textbook. テキストの演習に対する自分の解答をテキストに記入した/記入する。
1 = Submitting my answers for textbook exercises online. テキストの演習に対する自分の解答をオンラインで提出した。

Motivation

Because the majority of the course was online, student responses may be skewed toward that context. Nonetheless, when responding to a binary survey item (see Table 1), 13 students indicated that they were more motivated to complete their homework when submitting it online (M= 0.76[0.44]).

If students are not held accountable for completing their homework, some may lack motivation to complete it. In an online context, there are few ways to effectively hold students accountable without requiring online homework. Several student comments expressed the students' understanding that the online system held them accountable:

- A teacher can check whether students did their homework.
- I feel that I receive a grade for doing my homework.（評価に含まれていると感じるから。）
- The teacher can properly see what I did. The teacher can tell whether the students did the homework.（やったかちゃんと見てもらえるから。さぼることが出来ない（やったかやっていないかがすぐ先生にばれる）から。)

In terms of the online system adding an extra step to homework completion, 13 students felt that the system either increased or did not decrease their motivation to do homework, but they expressed various ideas:

- It is unlikely to forget submitting the homework
- Personally, it's easier to answer on a computer than to write.（個人的に、書くよりもパソコンで答える方が楽だから。）
- It's quicker to enter the answer on the keyboard.（キーボードで答えを入力する方が手早くできるから。）
- I feel like I'm doing my homework more.（宿題をやっている実感がよりあったと感じるから。）

Some students mentioned the homework being easier online, possibly because the second section of the textbook homework, the open-ended "synonym" questions, became close-ended to make answer verification reliable. Several students indicated positive feelings about close-ended questions:

- [I like online homework] because there are choices to select answers.
- The answer choices were limited, so it was easy to choose.
- For synonym quiz, online homework is better because it has options

However, not all students felt motivated to complete homework online. Overall, four students indicated a preference for recording their homework answers in textbooks. One student felt that the first homework section, matching, was easier in the textbook, possibly because the textbook allowed students to draw lines when matching items, whereas the online system required choosing an answer in a multiple-choice grid. Another student noted the additional steps required to complete online homework:

> Because what I have to do is just writing down the answers on my textbook. In the case of submitting it online, I had to repeat to look the textbook and the screen many times. It was a little bit tiring for me. However, we have no choice but doing so due to the virus.

Despite only a few responses indicating that submitting homework online was demotivating, the context of online classes during the COVID-19 pandemic may have impacted students' feelings as a whole because they felt that they had "no choice" except to submit homework in this way.

Understanding Textbook Content

In terms of understanding the content of textbook, 10 students indicated a preference for online homework (0.59[0.51]) on a binary survey item (see Table 1). Likewise, when responding to a 1–5 Likert-scale item (see Table 2), students also indicated a preference for online homework (3.94[1.20]).

Table 2: Student responses to the first Likert-type item on the survey (n = 17)

No.	Item	M	SD
1	In terms of understanding the textbook's content, how does submitting the homework online compare with recording your answers in the textbook? テキストの内容を理解することに関して、宿題をオンラインで提出することは、解答をテキストに記入することと比べてどうですか。	3.94	1.20

Note: The scale was 1–5, with labels on each side of the scale (see Appendix A).
1 = Online homework is much worse オンラインの宿題の方が悪い。
5 = Online homework is much better オンラインの宿題の方が良い。

On the binary item, 10 students indicated submitting answers online was better for learning; their open-ended responses expressed two advantages of online feedback. First, students can receive feedback on whether their answers are correct and and can view the correct answer. Several comments expressed this viewpoint:

- You can quickly find out what you don't understand.（自分でわからないものをすぐ調べられるから。）
- It's easy to review because you can work on it quickly and you can see the answer immediately.（手早く取り組める上に、回答がすぐにわかるため、復習しやすいから。）
- It is so good because I can see my own mistakes

Second, because of the way the online homework submission system functioned, students' answers were automatically collated. As a result, in-class feedback on the homework could be efficiently weighted toward language items that were widely misunderstood by the class:

- I think it doesn't matter if I carefully check the individual answers when recording them in the textbook, and also if many people made a mistake, then this is mentioned in class.（テキストでもきちんと個々人の回答がチェックされたり、多くの人が間違えていた問題について触れてもらえるならどちらでもよいと思う。）
- I receive feedback on my answers in class.（回答のフィードバックを授業。）

Nonetheless, online homework has disadvantages. Due to the close-ended questions introduced in the online adaptation of the textbook homework, one student expressed that recording answers in the textbook is better because "You will also be careful about details such as the tense of verbs."（動詞の時制などの細かいところにも気を付けることになるから）. Another disadvantage of submitting answers online is that it fails to benefit kinesthetic learning, a complaint widely expressed in student comments:

- I like writing text with my hands.
- I can deepen my understanding by actually writing.
- I can understand the contents of the textbook more when I use my hands to memorize information.
- There is no big difference difference between these two, but I can memorize better when I write down some vocabularies. In addition, recording answers gives me an opportunity to revise by myself. It improves my English skills more.

- It's easier to answer, but it doesn't promote understanding compared with recording for me.

Online and Face-to-Face Contexts

Because the course shifted from face-to-face to online after two lessons, students were able to experience both recording homework answers in their textbook and submitting their homework answers online. Given their experiences, students indicated a preference that the method of homework submission match the learning context. In binary survey items (see Table 1), 14 students expressed preferences that in an online class, online homework was best (0.82[0.39]). Meanwhile, 10 students felt that in a face-to-face class, recording answers in the textbook was best (0.41[0.51]).

Although this class did not submit textbook homework online while classes were face-to-face, the students expressed concerns about this possibility. One concern that the students expressed was that they may need to do the work twice:

- If I record my answers for textbook and send it online, it will be troublesome because the work will increase.
- During lessons, some students might want to review vocabularies, so it is useful for such students to record their answers [in the textbook] in advance as homework.

These comments highlight the importance of having learning materials physically available during face-to-face lessons, which may become more difficult if students' homework answers are only accessible digitally or more burdensome if student answers must be duplicated to be both physical and digital. Although these students' opinions are based on a hypothetical situation rather than an actual experience, consideration should be given to minimizing such issues when implementing online homework in a face-to-face context.

Students' Overall Feelings about Online Homework

Overall, on a 1–5 scale Likert item (see Table 3), students expressed a favorable impression of online homework (4.59[0.62]). The implementation of the online homework was well received, according to student comments:

- Although I prefer recording [answers in the textbook], online homework was convenient to use.
- It was easy to answer.

Table 3: Student responses to the second Likert-type item on the survey (n = 17)

No.	Item	M	SD
2	Overall, how would you rate your experience using online homework? 全体的に、オンラインの宿題に対するあなたの評価は	4.59	0.62

Note: The scale was 1-5, with labels on each side of the scale (see Appendix A).
1 = Very bad とても悪い
5 = Very good とても良い

Despite the generally positive feelings expressed by students, survey items and student comments suggest that context and implementation are vital in shaping students' opinions about online homework. For example, the Likert items (Tables 2 and 3) in this study roughly correspond to the Wooten and Dillard-Eggers (2013) study of online homework, but the results differ. In Wooten and Dillard Eggers' (2013) study, 453 accounting students in face-to-face classes were asked to

submit their textbook homework via the publisher's software. In answering the question, "With respect to learning, how does the use of online homework compare with 'pencil and paper' homework?" as a 1–5 scale Likert item (1 = "Much Better" and 5="Much Worse"), students indicated almost neutral feelings toward it, with a score of 2.21 from the 225 students required to do the homework and 2.90 from the 228 students for whom the homework was optional. Similarly, in responding to a second question, "Overall, how would you rate your experience using online homework?" as a 1–5 scale Likert item (1 = "Very Good" and 5="Very Bad"), students again indicated neutral feelings, with a score of 2.07 from the 225 students required to do the homework and 2.60 from the 228 students for whom the homework was optional. The contrast between students' perceptions of online homework in Wooten and Dillard Eggers' study (2013) and this study underscores the need to adapt homework to best suit the context and students' preferences.

Limitations

A limitation of this study is the small and relatively nondiverse sample size of 17 respondents of predominately upper-intermediate proficiency Japanese university students from which data were collected. Additionally, the rather unique context of the course, that is, changing from face-to-face to online, may further limit the range of applicability of the results.

Conclusion

Overall, students had positive feelings about online homework, but it is possible that responses were biased, as online homework was a logical option within the context of an online class. Even students with positive feelings about online homework did not always prefer it over doing homework in their textbooks. Context is a vital aspect for instructors to consider when deciding whether and how to use online homework.

More research needs to conducted on ways to improve student learning outcomes and students' feelings of satisfaction as a result of completing online homework. In a meta-analysis of formative assessment and feedback in various higher education settings, Morris et al. (2021) found "a rather mixed picture" with regard to online assessment and feedback systems and observed that it is difficult to empirically measure the effectiveness of such systems without funding carefully implemented wide-scale studies. Although large studies are necessary for affirming the efficacy, or lack thereof, of online learning tools, these tools have only very recently come into popular use within higher education. As a result, instructors who implement online assessment and feedback systems should do so with a careful concern for student needs and preferences to maximize learning outcomes within their context. In addition, the creators of online learning tools and the instructors using them should be willing to make continuous incremental adjustments to improve their efficacy.

References

Dyer, M. (2021). *Social Issues in the Movies – Discussion & Role-Play: Level 4*. DTP出版.

Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education 9(3)*. https://doi.org/10.1002/rev3.3292

Wooten, T., & Dillard-Eggers, J. (2013). An investigation of online homework: Required or not required? *Contemporary Issues in Education Research 6*(2), 189–198. https://doi.org/10.19030/cier.v6i2.7728

APPENDIX A
Survey Questions as Answered by Students via Google Form

I felt more motivated to do my homework when... 私は...の時、宿題に対してやる気がより出ました。 *

- ○ Recording my answers for textbook exercises in the textbook. テキストの演習に対する自分の解答をテキストに記入した。
- ○ Submitting my answers for textbook exercises online. テキストの演習に対する自分の解答をオンラインで提出した。

Why did this make you feel more motivated to do your homework? なぜそれによってやる気がより出たのですか。

Your answer

I best understood the textbook's content when... 私は...の時、テキストの内容を最も理解できました。 *

- ○ Recording my answers for textbook exercises in the textbook. テキストの演習に対する自分の解答をテキストに記入した。
- ○ Submitting my answers for textbook exercises online. テキストの演習に対する自分の解答をオンラインで提出した。

How did this help you better understand the textbook's content? なぜそれによってテキストの内容をより理解できたのですか。

Your answer

In an online class, which do you think is better? オンライン授業において、どちらがより良いと思いますか。 *

- ○ Recording my answers for textbook exercises in the textbook. テキストの演習に対する自分の解答をテキストに記入する。
- ○ Submitting my answers for textbook exercises online. テキストの演習に対する自分の解答をオンラインで提出する。

Why do you think this way is better for online classes? なぜオンライン授業においてその方法がより良いと思いますか。

Your answer

In a face-to-face class, which do you think is better? 対面授業において、どちらが * より良いと思いますか。

- ◯ Recording my answers for textbook exercises in the textbook. テキストの演習に対する自分の解答をテキストに記入する。
- ◯ Submitting my answers for textbook exercises online. テキストの演習に対する自分の解答をオンラインで提出する。

Why do you think this way is better for face-to-face classes? なぜ対面授業においてその方法がより良いと思いますか。

Your answer

In terms of understanding the textbook's content, how does submitting the homework online compare with recording your answers in the textbook? テキストの内容を理解することに関して、宿題をオンラインで提出することは、解答をテキストに記入することと比べてどうですか。 *

	1	2	3	4	5	
Online homework is much worse オンラインの宿題の方が悪い。	◯	◯	◯	◯	◯	Online homework is much better オンラインの宿題の方が良い。

Why did you give this rating? なぜそう評価しましたか。

Your answer

Overall, how would you rate your experience using online homework? 全体的に、オンラインの宿題に対するあなたの評価は *

	1	2	3	4	5	
Very bad とても悪い	◯	◯	◯	◯	◯	Very good とても良い

Why did you give this rating? なぜそう評価しましたか。

Your answer

【Research Brief】

Learners' Perspectives of Using Their Video Recordings to Aid in Performing Assessments

Devon Arthurson

Abstract

With the increased availability of technology, video-making and -sharing for brevity. Applications and online tools allow learners to record their performance of lesson tasks and permit multiple viewings unlike real-time assessments. This exploratory study focuses on students' opinions about self-recording's effects on their ability to self- and peer-assess, in addition to their perceptions of the level of difficulty required to assess. The project participants were 13 first-year university students from English debate classes at a liberal arts university in Tokyo. Students were required to video record homework tasks. In the middle of the semester, they were asked to self-assess by comparing their performances from a previous homework task with a later lesson's task. Throughout the semester, students also peer-assessed debates. The quantitative data were collected from two surveys. The results were that over half of the participants felt that recording their own performance helped them to assess both themselves and others. The students' perceptions about assessment's level of difficulty depended on the skills being assessed and how long the students had been using the skills.

Keywords: *assessment, online tools, peer-assessment, self-assessment, video recordings*

1. Introduction

In 2020, the global pandemic forced those in education to seek out new ways of connecting with their learners through the utilization of online learning tools such as sharing platforms applications and online tools, As conditions for many returned to in-person settings, some, specifically those in language learning, saw the benefits of continuing to use those tools. Having students record their performance can be useful for not only the instructor's assessment but also the students' self- and peer-assessments. Some benefits of student-led assessment are that it can allow the learners to critique their performance, give them more ownership of their learning, and provide more meaningful discussion with the instructor about assessment (Sebba et al., 2008, p. 16). According to Sumardi, Adzima, and Wijaya (2020), video recording can transform foreign-language speaking assessments due to its effectiveness (p. 67). Instructors and students can view performances multiple times for a better understanding of the performances and can provide more accurate grading. Indeed, tasks requiring students to record themselves existed before 2020, as shown in Christianson, Hoskins, and Watanabe's (2010) study. Nonetheless, other educators may have used the recordings solely for their own assessment purposes. However, online video tools allow both students and teachers to easily create recordings anywhere at any time and upload these recordings for the whole class' viewing. These tools also let students decide which recordings to share. This study will explore EFL students' opinions about how making videos affected their ability to assess.

2. Literature Review

The benefits and implementation of self- and peer-assessment have proven to be useful tools in language-learning classrooms. Baleghizadeh and Masoun (2013) stated that foreign language teachers who did incorporate self-assessment in the lessons were in favor of using this type of assessment in their practice (p. 53). In a review of empirical studies about self- and peer-assessments, Joo (2016) asserted that students felt that increased language skills, more critical awareness, and a deeper understanding of their abilities were the advantages of these types of assessments (p. 76). For a successful implementation, it is important that both the students and their instructor view assessment in a different way. Sebba et al. (2008) contented that these assessments change the relationship between the instructor and learner from from being hierarchal to parallel, consequently shaping instructor's practice based on how their students respond (p. 2). Moreover, when instructors model how to assess for their learners, thus providing them with more autonomy by sharing the responsibility of assessments, the assessment tasks may result in more success. For those instructors seeking a more egalitarian classroom, allowing students to assess can be a meaningful exercise.

Communicative tasks can be challenging for those learning a foreign language. Many may feel reticent and anxious to speak in front of others, especially when formal assessment is also occurring. Giving assessment tasks to these learners can also be demanding as it is difficult to perform the task, process the performance, and then assess the performance. Students must have a solid awareness of the grading tools as this affects how successfully they can assess themselves or their classmates (Joo, 2016, p. 69). Additionally, utilizing achievable assessment tools is a part of self-assessment (Benson, 2011, p. 168). Therefore, providing students with awareness of grading criteria and opportunities to practice speaking tasks, such as presentations, debates, and speeches using self-recording tools, can lessen negative feelings and provide other positive benefits such as better performances. Self-recording tools let the learners record their performance in an environment of their choosing and allow them to record multiple times before uploading the performance they want to share with others. Sharing rubrics with students to practice their own self-assessment as well as giving instructor feedback about their performance can deepen their understanding of the grading requirements. Furthermore, having students make videos can provide more tangible recognition of grading expectations through the practice of self-assessment, and may make peer-assessment tasks less overwhelming. Having learners upload their recordings of tasks so that the instructor can comment by either written or oral responses can allow for students to review their performance in conjunction with instructor feedback to more concretely understand ways to improve their skills. The advantages of students recording videos allows their own and peer viewing of their performance multiple times, resulting in their ability to more consciously evaluate their delivery and recognize their skill development (Qureshi et al. 2019, p. 21). Not only can students' assessment skills increase through self-recordings, but their feelings about the language being learned and their relationship toward the language may also become more positive. The promotion of self-confidence is another benefit of students using the video recordings of their performance (Sumardi et al., 2020, p. 69). Göktürk (2016) reported that the learners felt making the recordings not only aided with the improvement of their speaking ability but also increased their motivation to use English and resulted in a feeling of satisfaction (2016). Even though it may seem that self-recording done in the foreign language under study could cause stress and be intimidating, the experience can be beneficial to some learners by developing more confidence and providing enjoyment in communicative language tasks.

The literature on the effects of student video recording and students practice of assessment in EFL or ESL settings is scant. However, Christianson et al.'s (2010) study with students at a university in Tokyo used webcams, and the recordings were done by students in groups at the instructor's office for peer- and self-analysis for an academic speaking course. They emphasized that the use of recordings lets the students control the assessment (p. 2). This means that students using their video recordings for assessment may result in more learner autonomy. Christianson et al.' study holds similarities to this study, yet with the changes in technology during these past 10 years, for instance the increased use of laptops and smartphones in language learning, along with online tools and applications that allow for the freedom to video record anywhere and at any time, further exploring student opinions about video recording and student-guided assessment tools is an important area for examination.

3. Research Methodology

This exploratory study used a quantitative approach to analyze the data from the 13 participants. They were first-year students from three mandatory debate classes at a Tokyo liberal arts university. They had TOEIC scores from 480 to 699, with two classes having in-person lessons and one class having online lessons. Though the number of students in all classes was much higher than the number of participants, these participants were selected as they made videos regularly as part of the homework tasks and completed both surveys. The surveys were delivered using a Google Forms link via email and the online learning platform. The first survey was implemented during the middle of the semester and the second survey during the end of the semester. This study analyzes the surveys' three-level Likert scale questions' responses.

The research questions are as follows:

1. Did making videos aid in students' ability to self- and peer-assess?
2. How do students perceive the level of difficulty in assessing?

Context of the Study

Over the 14-week semester in fall of 2021, students were required to make eight videos with a computer online tools as part of their homework tasks for weeks 1–6, 10, and 11. The video tasks were based on the current lesson activities and/or to prepare for the next lesson. Below are two examples of the homework tasks:

> *Lesson 2's task: For the proposition on page 22: "Skipping breakfast is good for your health." Share your arguments for both Affirmative and Negative with "Useful Expressions for Constructing Arguments" on pages 17 and 18. Use the reflection journal to help you.*

> *Lesson 10's task: Please use your Reflection Journal to help you. Use the summary your team made in the lesson. Remember to include three points from your team, the other team's points, the proposition, and your team's position. Also use Appendix A "Useful Expressions for Debate Skills" on page 58.*

Students were given individual and private feedback on the video-making and -sharing online tools by the instructor about their performance based on the appropriate section of the course's rubric. The assigned textbook *Up For Debate* (Mishima et al., 2021) that contained the rubric and skills referenced in the above homework example prompts. It was the instructor's expectation that the students would review the comments about their videos to gain a greater understanding of how to use the rubric. They would be expected to give feedback to their peers as judges during the mid-term and final-term test debates when their team was not debating.

4. Results and Discussion

The students' perspectives about making videos' effectiveness on their ability to assess will be shared. In response to the survey question, "Did watching the two videos help you to learn how to judge your performance?," 7 students agreed, 5 were not sure, and 1 disagreed. The process of making a video was viewed as having value to aid in self-assessment for over half of these students. Regarding the question, "Do you think watching your own video helps you to judge other students' performances?," 7 students agreed, 4 were not sure, and 2 disagreed. It appears that watching their own video-recorded performances was beneficial to these students during peer assessment. Potential reasons for agreement could be the level of exposure that they had with the rubric on a weekly basis and how the instructor gave written feedback for each video based on the rubric. This echoes what both Joo (2016) and Benson (2011) maintained that a good understanding of an achievable rubric will aid in students' ability to achieve assessment tasks. However, the instructor cannot confirm if the students reviewed the written feedback of their video performances. It was hoped that if the students could view their videos again after reading the feedback, they would better understand how the rubric was used for assessments. In future uses of video recording tasks when written feedback is given, it will be important to establish how learners review it? Activities such as reflections tasks or discussions about the feedback could be used to confirm this.

The level of difficulty in assessing given by students will be examined. As different skills needed to be assessed, students' opinions about the level of difficulty in assessing, both themselves and their classmates, also differed. Self-assessing more concrete skills such as a debate's organization that followed a clear pattern was viewed as having a moderate level of difficulty by nearly all students. Furthermore, when asked about generally assessing organization skills of debates, such as in peer-assessment settings, over half of the students responded that the task was easy. A possible reason is that students felt that assessing themselves was a task with less pressure as regards accuracy than assessing others, possibly making self-assessment easier than peer-assessment. With more abstract skills, such as when self-assessing a debate's strength, over half of the students felt it was a moderate to difficult task. Yet, when asked about generally assessing a debate's strength, 10 students agreed that it was an easy task. There could be a few factors for this. Data were collected about self-assessment in the middle of the semester, while the data about general assessment were collected at the end of the semester. This means students had more exposure to the skills as well as familiarity with assessments. Additionally, students assessed their peers with their debate team members during the test lessons at the middle and end of the semester. It could be that students had more confidence to assess with their teammates than by themselves, potentially lessening the perceived level of difficulty. In addition, as Sumardi et al. (2020) write about increased confidence in relation to recordings, it could be that as students completed more self-recorded videos for homework tasks by the end of the semester, they felt more sure of in their assessment abilities.

In future uses of video-recording and assessment tasks, it will be beneficial to have students, either individually or in the teams, submit a written peer-evaluation form, which would be collected, instead of only giving oral comments to their classmates about their test performance. Then the instructor can better understand how students assess their peers. This will be helpful to know more about students' opinions concerning the assessment task's level of difficulty. Then the instructor can share with the students their assessments' strengths and weaknesses, providing students with more certainty of their ability to assess and possibly decreasing the level of difficulty to perform assessments.

5. Conclusion

Video-making and sharing tools and applications are becoming increasingly accessible and can be a useful tool for students to practice their own self- and peer-assessments. Having students perform assessments can allow for students to have more autonomy in their learning. Unlike real-time assessments, video recordings can allow for multiple viewings. Using video-making and online -sharing applications tools also gives students more control control over deciding/choosing which performances to share share with others as well as the freedom to record at any time and anywhere. Instructors can provide feedback on the learner's performance with more certainty, and learners can view their performance again along with the feedback. This could allow for the learner to have a better understanding of how their performance was graded, possibly giving them more understanding of the grading tools.

In this exploratory study, more than half of the participants agreed that making videos helped them learn how to assess their own performance and their classmates' performances. The participants responded that the level of difficulty required to assess varies. Possible factors on the level of difficulty might be the type of skills to be assessed and their familiarity with the skills. Furthermore, assessment done by a learner individually versus assessing as part of a team may also influence differences in assessment's level of difficulty. A better understanding of how students assess compared with the instructor could aid in more effectively using student-recorded videos and student-assessment activities. More research needs to be done about having students use their own self-recorded videos and perform assessments as this is a timely area for exploration.

References

Baleghizadeh, S., & Masoun A. (2013). The effect of self-assessment on EFL learners' self-efficacy. *TESL Canada Journal /Revue TESL Du Canada (31)*1, 42-58. https://doi.org/10.18806/tesl.v31i1.1166

Benson, P. (2011). *Teaching and researching autonomy (2nd ed.)*. Pearson Education Limited.

Christianson, M., Hoskins, C., & Watanabe, A. (2010). Evaluating the effectiveness of a video-recording based self-assessment system for academic speaking. *Language Research Bulletin, 24,* 1-15. https://core.ac.uk/download/pdf/234717241.pdf

Göktürk, N. (2016). Examining the effectiveness of digital video recordings on oral performance of EFL learners, *Teaching English with Technology, 16(*2), 71-96. https://www.researchgate.net/publication/305445319_Examining_the_Effectiveness_of_Digital_Video_Recordings_on_Oral_Performance_of_EFL_Learners

Joo, H. S. (2016). Self- and peer-assessment of speaking. *Studies in Applied Linguistics & TESOL*

(SALT) 16(2). 68-83. https://doi.org/10.7916/salt.v16i2.1257

Mishima, M., Kita, S., Donnelly, M., Hartley, M., Iwai, K., & Vaughan, R. (2021). *Up for Debate* (*2nd ed.*). DTP Publishing.

Qureshi, B., Roseberry, J., & Qureshi, Z. (2019). Boosting intersubjectivity by digital video recording project in EFL classes. *Scope: Journal of English Language Teaching 4* (1). https://doi.org/10.30998/scope.v4i01.4977

Sebba, J., Crick, R. D., Yu, G., Lawson, H., Harlen, W., & Durant, K. (2008, January 1). *Systematic review of research evidence of the impact on students in Secondary Schools of self and peer assessment*. ORA. https://ora.ouls.ox.ac.uk/objects/uuid:f5df1fd5-2bcb-47b1-8226-b38ba2431019

Sumardi, Adzima. R., & Wijaya. A. N. (2020). Digital video project: An authentic assessment to assess students' speaking skills. *Indonesian Journal of EFL and Linguistics 5* (1), 52-72. https://doi.org/10.21462/ijefl.v5i1.217

【Research Brief】

Designing a Mobile Application to Track Spoken Fluency Development

Heather Woodward & Charu Gupta

Abstract

English discussion class (EDC) course designers state that the most important language-based objective of EDC is to improve students' spoken fluency (Hurling, 2012). In general, the students do seem to be increasing their speaking speed; however, designers have not developed a method to document their progress across the semester. Meanwhile, mobile-assisted language learning (MALL) applications can enhance personalization and feedback, but these applications have been traditionally designed to build students' second language (L2) vocabulary rather than their L2 fluency (Heil et al., 2016). In Spring 2022, we repurposed an existing mobile application that analyzes speech rate (words per minute) to provide more personalized feedback. The problem with the application is that it does not have features that reflect current spoken fluency research. Therefore, we review the literature on spoken fluency research, then design an application that collects and records students' speech rate (syllables per minute), mid-clause pause frequency, self-assessment, and peer-assessment.

Keywords: *EFL, MALL, speaking, 4/3/2*

Introduction

The goal of language learning for many Japanese students is the ability to speak fluently; however, achieving fluency is challenging in the context of Japan where there are few chances for students to use the foreign language owing in part to the washback effects from university entrance exams, which have historically not included a speaking component (Garside, 2020). Japanese students also strive for accuracy and perfection, so they spend time correcting their syntactic or phonetic mistakes and this extra time can decrease their spoken fluency (Watanabe & Long, 2019). Without incorporating fluency-building activities into the L2 classroom, communicative language teaching is limited in scope (Gatbonton & Segalowitz, 1988), and as a result, students' learning might not be available during the typical demands of real-world communication (Hurling, 2012). Not meeting typical demands of real-world communication means that students are unable to hold listeners' attention or save face (Lennon, 2000).

Applied linguistics researchers analyze three aspects of spoken fluency – speed, breakdown, and repair. The first aspect, speed, equates to the density and flow of speech. The second aspect, breakdown, consists of hesitations and pauses, and the last aspect, repair, comprises corrections, reformations, and repetitions (Skehan, 2003; Tavakoli & Skehan, 2005). For example, researchers might calculate the speed of the L2 spoken performance by using the average number of syllables per minute, then analyze the breakdowns of an L2 spoken performance by calculating the mean length of pauses. The process of transcribing and analyzing students' L2 spoken performances is too time consuming to be widely incorporated by second language (L2) instructors in the classroom. In addition, conditions such as background noises of other students speaking make collecting and analyzing data very challenging.

However, quantitative measures such as the average number of syllables per minute are limited

in describing spoken fluency performance because they are unable to account for other mediating factors such as students' first language (L1) speaking styles, personalities, or socio-pragmatic considerations (Tavakoli & Wright, 2020). One possibility is to create a mobile application that incorporates both quantitative measurements and qualitative judgments while also adding instructional features to enhance students' metalinguistic awareness of spoken fluency to provide more personalized feedback. By using a filter, the background noises of other students in the classroom might be reduced enough to collect information about students' spoken fluency performance. In this research brief, we first review literature on the goals of English discussion class (EDC). Afterward, we investigate research on applied linguistics to determine the most appropriate objective measurements to use to measure spoken fluency. We also discuss gamification and personalization, and then present a mobile application design that includes these features.

Literature Review

English Discussion Class

EDC is required for first-year students at Rikkyo University to improve their communication skills, academic discussion skills, and spoken fluency (Hurling, 2012). Examples of communication skills along with a corresponding formulaic sequence include clarification (Can you repeat that?), comprehension (Do you understand?), and paraphrasing (In other words, do you mean…?) (Kita et al., 2022). Academic discussion skills include joining a discussion (Can I start?), viewpoints (How about from the perspective of...?), and sources of information (How do you know about that?) (Kita et al., 2022). Students are separated into four EDC levels according to their test of English for international communication (TOEIC) scores. The four EDC levels are: Level 1 (TOEIC scores 680 or above; common European framework of reference for languages (CEFR): B2 and above), Level 2 (480-679; B1-B2), Level 3 (280-479; A2-B1), and Level 4 (279 or below; A1-A2). Additionally, the course is for 14 weeks, has 10 students per class, and students meet once a week for 100 minutes.

During each lesson, five students participate in a 20-minute discussion while instructors assess students' participation and use of formulaic sequences (English Discussion Committee Handbook, 2022). On the rubric, students receive a 4, or superior, for academic discussion skills if they use them quickly, appropriately, and without looking in the textbook. Quickness relates to the spoken fluency's aspects of speed and breakdown. They receive a 3, or good, if students use discussion skills, but cannot do so from memory (English Discussion Committee Syllabus, 2022). The other scale descriptors of none, poor, and fair, do not explicitly refer to students' spoken fluency. In addition, the two other constructs assessed—communication skills and participation—also do not explicitly refer to spoken fluency. However, speaking quickly during the discussion performance test gives others time to participate and use discussion and communication skills. The rubric is the same for all EDC levels (i.e., from Level 1 to 4). Although spoken fluency is the most important language-based goal of EDC according to course developers, the only fluency assessment of students consists of the extent to which instructors perceive their fluency of discussion skills in the context of group discussion performance.

4/3/2 Activity and Other Fluency Practices in EDC

A modified version of Maurice's (1983) 4/3/2 activity called 3/2/1 is used during every EDC

lesson. 3/2/1 incorporates task features of repetition and time pressure, and these task features have been shown to boost students' spoken fluency (e.g., Arevat & Nation, 1991; Boers, 2014; Garside, 2020; de Jong & Perfetti, 2011; Molina Barriga & Briesmaster, 2017; Thai & Boers, 2016; Tran & Saito, 2021). For this modified version of 4/3/2, students deliver three iterations of their monologues. Speakers change listeners after each iteration. 3/2/1 takes 15 minutes and with 10 students per class, 5 students speak simultaneously. For theoretical support, course designers cite Schmidt (1992) who uses Anderson's (1989) adaptive control of thought (ACT) theory (Hurling, 2012). In ACT, fluency equates to automatic processing of proceduralized knowledge (i.e., knowing how to use the L2 with effortless efficiency) and after repeated, meaningful practice, in theory, students can retrieve knowledge more quickly and smoothly than before.

Not only do students use 4/3/2 in EDC to improve spoken fluency, but course designers have also adopted other practices to help students boost spoken fluency. EDC students a) use formulaic sequences in communicative contexts; b) spend time pre-task planning to conceptualize and formulate their ideas; c) repeat the task; and d) learn about spoken fluency with metalinguistic awareness-raising activities (See Tavakoli & Hunter, 2018). For a) use formulaic sequences, students incorporate discussion and communication skill phrases into their discussions to boost fluency. For b) pre-tasking planning, students have preparation time before their discussions to think of their ideas. For c) task repetition, students repeat the discussion with slightly different topics for the practice and perform two discussions per lesson. For d) metalinguistic awareness-raising activities, Curran (2019), an EDC instructor, uses self-reflection activities after the 3/2/1. For instance, he asks, "Did you speak more quickly?" and "Did you have to pause?" Awareness-raising activities can help students achieve a greater understanding of spoken fluency and the ways to improve it (Curran, 2019). EDC course designers incorporate these types of activities to help students improve their spoken fluency during the course, but there is no way to measure students' progress across the semester so instructors might find it challenging to provide personalized feedback or support for students. Next, we explain spoken fluency research findings to design a mobile application that can track its development.

Spoken Fluency Research Findings

Spoken fluency is multidimensional so to explain research findings more precisely, Segalowitz (2010, 2016) separates spoken fluency into three interconnected dimensions: cognitive, utterance, and perceived. Cognitive fluency is the fluidity of underlying cognitive mechanisms that cause L2 speech acts (Segalowitz, 2016) and is operationalized as reaction time speed and reaction time stability (Segalowitz & Segalowitz, 1993). Utterance fluency refers to observable speech production (i.e., speed, breakdown, and repair measurements). Perceived fluency is raters' subjective judgments of overall spoken fluency (Segalowitz, 2010, 2016). His model emphasizes that fluency is not only a psycholinguistic construct, but also an interactional one, which means that listeners also play an important role in determining spoken fluency (Tavakoli & Wright, 2020). Fluency is regarded by examiners and raters as the most difficult aspect of L2 spoken performance to judge (Kang et al., 2019). Therefore, determining which aspects of utterance and perceived fluency should be evaluated and which features of utterance fluency should distinguish descriptor levels are key decisions for both human and automated rating systems (Tavakoli & Wright, 2020). In the next sections, we investigate the applied linguistics research to investigate these key decisions for designing our mobile application.

What Aspects of Utterance Fluency Should be Evaluated on the Application?

A meta-analysis of 22 studies by Suzuki et al. (2021) investigates the relationship between aspects of utterance fluency compared with listener-based perceived ratings of monologues. Utterance fluency measurements should be strongly associated with perceived fluency to be a reliable source of information for assessments. Results show that listeners' perceived fluency ratings are strongly associated with pause frequency and speed, moderately associated with pause duration, and weakly associated with repair fluency (Suzuki et al., 2021). Repair fluency (corrections, reformulations, and repetitions) can strongly relate to speaking style preferences (Kahng, 2014; Tavakoli & Skehan, 2005). Meanwhile, composite measurements show the strongest effect sizes compared with any pure speed, repair, or breakdown measures (Suzuki et al., 2021). Composite measurements incorporate two or more pure measurements. For example, one composite measurement is speech rate, which is calculated by dividing the total number of syllables in a given speech sample by the entire time it takes to create the speech sample in seconds, including pause time, and then multiplying by 60 (Kormos & Dénes, 2004). A pure speed measurement is articulation rate, which is similar to speech rate, but does not include pauses.

Pause location strongly affects perceived fluency as well. Suzuki and Kormos (2020) examine L2 argumentative speaking performance judged by 10 native English-speakers inexperienced raters. For this task, perceived fluency is more strongly associated with mid-clause pausing frequency than with other aspects (Suzuki & Kormos, 2020). For instance, between-clause pausing is "She doesn't like doing dishes, (pause) but you don't either," which can indicate topic change or an idea shift. Mid-clause pausing is "She doesn't like doing (pause) dishes." L2 speakers pause mid-clause more frequently as compared with L1 speakers, which suggests that at least some L2 speakers' mid-clause pausing reflects L2 proficiency gaps associated with online planning, reformulation, and replacement (Skehan et al., 2016; Tavakoli, 2010; Tavakoli & Skehan, 2005). In Suzuki et al.'s (2021) meta-analysis, researchers find that mid-clause pausing is even more strongly associated with perceived fluency than pure speed measurements, but slightly less associated than composite fluency measurements such as speech rate. This suggests that speech rate and mid-clause pausing are some of the best utterance fluency measurements for tracking students' spoken fluency progress across the semester.

What Utterance Fluency Features Should Distinguish Descriptor Levels?

EDC does not distinguish spoken fluency features for Levels 1-4, so one question we ask is, if we incorporate descriptor levels on the spoken fluency application, which features should distinguish levels? The problem is as Bradlow et al. (2017) conclude "individual variability in L2 spoken language production may be best understood within the context of individual variability in L1 spoken language production." For instance, repairs (repetition and pausing) can be indicative of L1 speaking style rather than L2-specific disfluencies (Bosker et al., 2012; Duran-Karaoz & Tavakoli, 2020). Additionally, Bradlow et al. (2017) write that in absolute terms, students speak their L2 slower than their L1, and L1 speaking rate can significantly predict L2 speaking rate. That is, faster L1 speakers are also faster L2 speakers. Shrosbree (2020) also finds that participants with TOEIC scores over 900 have positive correlations for 7 of the 10 spoken fluency measures; however, in absolute terms, they speak their L2 slower than L1 (Shrosbree, 2020). Yet, test makers seldomly ask raters to consider students' L1 speaking style when assessing (Segalowitz, 2016).

In addition to internal factors, there are external factors that influence spoken fluency. External

factors include background noise from others, ease or familiarity with topics, and interlocutors (Tavakoli & Wright, 2020). Segalowitz argued for the importance of assessing sources of variability that are not related specifically to L2 disfluencies but that characterize a person's general performance in the given testing conditions (Segalowitz, 2010). For these reasons, zero-stakes assessments, conducted for formative purposes, rather than summative, can include individualized baselines to account for students' speaking style, and this kind of assessment can also take external factors into consideration that might be sources of variability. One example is students journaling about their experiences to discuss different external factors. Therefore, instead of descriptors, an alternative solution is for the mobile application to suggest a small, incremental mid-clause pause frequency and speech rate goals that take into consideration internal factors such as students' L1 speaking performance and external factors such as topic familiarity.

Should Human Raters be Included With a Computer Automated Rating System?

Another design decision for a mobile application that tracks spoken fluency is whether to have the application fully automated or to have an element of human ratings. With regard to human raters, de Jong (2018) states that the specificity and amount of instruction can enable human raters to focus on certain fluency features such as pause frequency, pause location, and speed. Moreover, raters might even be innately sensitive to pause location with an understanding that mid-clause pausing is more likely a reflection of decrease in cognitive fluency (Kahng, 2018). Automated speech evaluation alone might not be suitable for spoken fluency assessments because the goal of measuring students' L2 ability is to determine the quality of test takers' verbal communication with humans, not with machines (Ginther et al., 2010). For these reasons, listener-based responsibilities during 3/2/1 such as tracking mid-clause pausing and peer-evaluation of comprehensibility are also important. Human raters might increase the reliability of fluency judgments because students perceive other reasons for pausing such as socio-pragmatic considerations (e.g., if students are telling a sad story, they might pause for non-L2 specific disfluent reasons).

Saeki et al. (2021) have designed an AI conversational agent called InteLLA that adapts its interview to assess L2 spoken proficiency. InteLLA's fluency aspect detects speakers' pauses, annotates pause locations (e.g., of mid-clause pauses), and also notes disfluency markers such as fillers and false starts. To mimic human ratings, researchers assign different weights to the fluency's temporal features based on associations between utterance fluency and perceived fluency. InteLLA's accuracy in replicating L2 perceived fluency is 60% using the CEFR level classification system (Saeki et al., 2021). Some of these features might be also useful for a mobile application. The technology can note disfluency markers such as fillers, but for the annotations of pauses, we wonder whether listener-based judgments of mid-clause pausing might be more reliable than an automated system because listeners might have greater insight into the reason for the pause than an automated system. As the mobile application's purpose is different from InteLLA's purpose, we can also add other human-rating judgments such as self-assessment and instructor-based assessments.

Gamification and Personalization

We have also reviewed literature on gamification and personalization. Implementing technology-driven play and competition is becoming increasingly popular in the classroom (Chen et al., 2021; Dehghanzadeh et al., 2021; Flores, 2015; Huang & Soman, 2013). Researchers have found that

students are more motivated and engaged, while also feeling less anxious about using their L2 when gaming elements such as points, medals, badges, or progress tracking are introduced (Dehghanzadeh et al., 2021; Flores, 2015; Huang & Soman, 2013). Students are more inclined to compete with their classmates and themselves to earn medals and increase their ranking (Arce & Valdivia, 2020). Rego (2015) specifies important elements of gamification such as having

a. clear goals
b. rules
c. an accessible and aesthetic user interface
d. collaboration and interaction aspects
e. rewards (e.g., medals or ranking)
f. tracking of students' progress
g. performance feedback

We have included features a)–g) in the mobile application design and we discuss these elements in the next section. In addition to Rego's (2015) advantages, gamification has also been lauded for its ability to incorporate personalized learning (Chen et al., 2021), as personalized activities are usually more meaningful, relevant, and self-initiated. For technology-based personalization, Kukulska-Hulme (2016) emphasizes that with continued L2 practice, personalization can help students to a) identify their needs, b) develop greater awareness of the learning process, and c) monitor their progress. For monitoring their progress students can reflect on and the analyze their performance (Dehghanzadeh et al., 2021). Heil et al. (2016) review 50 language learning-related applications and find that while mobile applications track progress and adjust difficulty levels, most adjustments are minor and not always implemented by students. The feedback given also does not provide reasons for incorrect responses or how to improve performance (Heil et al., 2016). To navigate the aforementioned limitations, instructors should include their feedback (Chen et al., 2021).

Application Design

In this section, we describe the functions of the application and connect it to the literature review. To measure students' spoken fluency, they record themselves for each round of 3/2/1. While they are speaking, the application is collecting data on average number of syllables untrimmed, maximum speed, total number of syllable, and fillers. Their partners listen to their monologues while tallying their mid-clause pauses for each round. If students reach or exceed their fluency goal, they have a "Congrats" message. Once speakers finish their monologues, their listening partners rate the speaker on comprehensibility (i.e., what percentage of the speakers' talk could they understand) and rate overall impressions. As students in each class are grouped according to their proficiency, we believe that listeners might be able to provide some useful feedback for speakers on their mid-clause pausing and overall comprehensibility. Below are screenshots of the application login page, homepage, and congrats message.

After students log in screen, the home screen is where students record their speech rate (See Figure 1 above). The green star above the speedometer on the second and third smartphone screens from left represent students' fluency goal. The speedometers have the features of average syllables per minute, maximum speed, total number of syllables, and total number of fillers. The design is inspired by Salomatin's (2019) speech rate screen, but has the addition of a timer, fluency goal, and

Figure 1

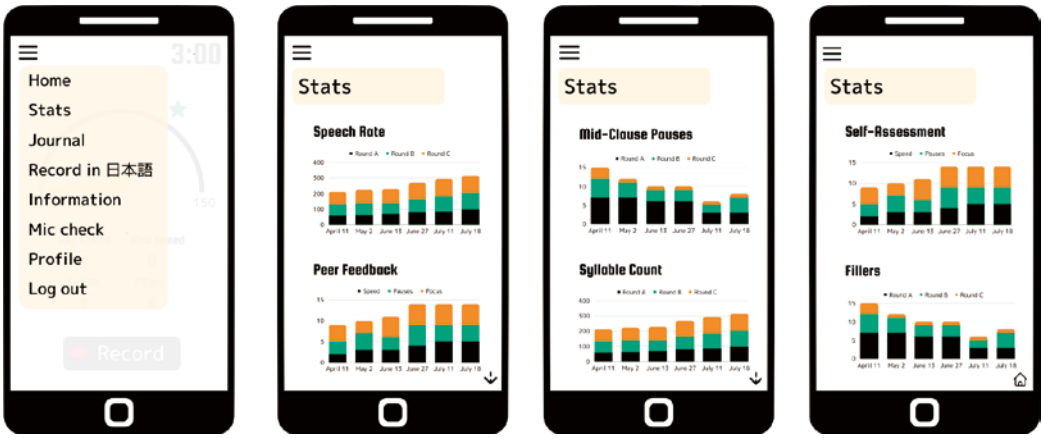
Application Login, Homepage, and Congrats Message

Note. Designed by primary author using Canva Pro with images from Red-Hawk, Jaruka, and Kerismaker

syllable counter. If students receive a speech rate at or above their goal, then the "Congrats" message appears on the fourth screen. The speech rate feature shown in the speedometer is untrimmed which means that disfluencies such as "um" and "uh" have not deleted. Other features include a) a noise suppression filter to reduce background noise so that multiple students can perform the activity at the same time, b) linking students' accounts to Rikkyo University Gmail accounts for safety purposes and ease of access, and c) programming the application for both Android and iOS so that it is accessible to all students with mobile devices. Figure 2 below shows the screens for menu and the statistics.

Figure 2

Application Menu and Stats



Note. Designed by primary author using Canva Pro with images from Red-Hawk, Jaruka, and Kerismaker

The menu screen is the first image on the left of Figure 2. The menu items include home, stats, classmates, journal, record in Japanese, information, mic check, profile, and logout. When students tap on "Stats," the second screen from the left appears, and students can scroll down to view different graphs (See Figure 2 above). These bar graphs include speech rate, peer feedback, mid-clause pauses, syllable count, self-assessment, and fillers. Figure 3 below shows screens for the listener responsibilities.

Figure 3

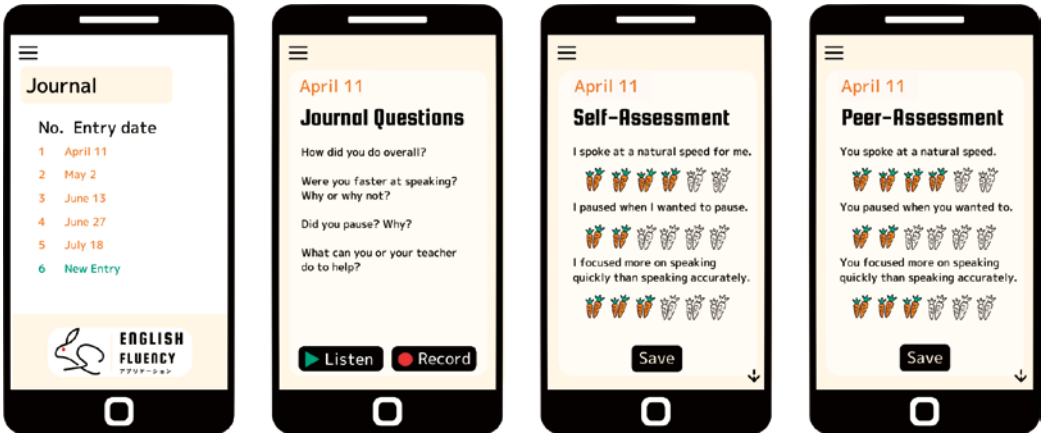
Peer-Assessment Features

Note. Designed by primary author using Canva Pro with images from Red-Hawk, Jaruka, and Kerismaker

On the first screen from the left, there is a list of the students' classmates. Before the 3/2/1 activity, students select their speaking partners. Once they select their partner, they can see the second screen to the left. Listeners press the green plus sign when they hear a mid-clause pause from their speaking partners. If listeners make a mistake, they can press the negative (orange circle) to erase the pause that they mistakenly added to their partners' data. Students need to be taught to detect mid-clause pauses, but we believe that teaching students to be sensitive to mid-clause pausing is feasible given research findings of de Jong (2018) and Kahng (2018). After, listeners can rate their speaking partners on comprehensibility. They can slide the green circle to the percentage that they can comprehend of the speaker's speech. On the last screen from the left, listeners answer the questions about the speakers' performance. The more carrots they select indicates greater agreement with the statements. Figure 4 below shows the journal entry sections for self-assessment and instructors' feedback.

Figure 4

Journal, Self-assessment, and Instructor Feedback Pages



Note. Designed by primary author using Canva Pro with images from Red-Hawk, Jaruka, and Kerismaker

On the first screen to the left of Figure 4, students review previous journal entries or select new journal entries. When they select a date (e.g., April 11), they see the adjacent screen on the right with metalinguistic questions to reflect on their performance. When they scroll down the journal section,

there is a self-assessment screen. There are three self-reflection statements: a) I spoke at a natural speed for me; b) I paused when I wanted to pause; c) I focused more on speaking quickly than speaking accurately. Selecting more carrots equates to a greater agreement with the corresponding statements a)–c). We have decided to use these types of reflective activities so that there is a place for students to discuss other mediating factors such as emotions, topic familiarity, or socio-pragmatic considerations affecting their performance. Figure 5 below shows the L1 mid-clause pause fluency, L1 speech rate, and information pages.

Figure 5

L1 Fluency and Information

Note. Designed by primary author using Canva Pro with images from Red-Hawk, Jaruka, and Kerismaker

On the first screen from the left, students can determine their L1 speech rate. Shrosbree (2020) states that ideally Japanese speech rate is calculated by counting morae for long, spontaneous speeches in part because Japanese is a mora-timed language rather than a syllable-timed one such as English. However, as the application directly compares speech rates, it needs to use the same syllable units. L1 pauses and L1 speech rate data can be used for stats graphs and speedometer as a baseline. For goals, we imagine students continually update their goals to be a little faster than their previous speech rate and they can use the information from their L1 speaking performance to set realistic goals that are personalized. The third screen is the information page, which can help students increase their metalinguistic awareness of spoken fluency. When they tap one of the questions, answers appear. Figure 6 below show the profile, fluency goal, and logout pages.

On the first screen, there is the students' profile page with the language goal. They can change their name, icon, language, and fluency goal. On the second and third screens, students can choose their fluency goal by moving the green circle or the green star. The orange circle and star represent goal recommendations based on students' progress. The last screen is the logout page.

Another aspect is the application design itself. Rego (2015) stresses the importance of an aesthetically pleasing and accessible user interface. Memon (2019) reinforces this idea of designing an application that meets the user's needs, contains consistent elements throughout, and is uncluttered. Although there is some technical terminology that might make the interface less user friendly, we include a language option if students want to use the application in Japanese and we also have a feature whereby students can double tap any technical words to retrieve definition and translation so that they can understand the technical words more easily. For instance, if they double tap on the words "Syllable Count," a definition appears. In the next section, we discuss possible limitations of the application.

Figure 6

Profile, Fluency Goal, Logout Pages

Note. Designed by primary author using Canva Pro with images from Red-Hawk, Jaruka, and Kerismaker

Limitations

Monologic and Dialogic Fluency

Studies have investigated differences between dialogic and monologic fluency. Monologic speaking tasks include Maurice's 4/3/2 and presentations whereas dialogic speaking tasks equate to discussions with another partner. By extension, dialogic speaking tasks share more characteristics in common with group discussions than monologic speaking tasks (e.g., turn-taking, and greater online planning time). Researchers find that students' spoken fluency increases significantly more for dialogic speaking tasks than for monologic speaking tasks (Ferrari, 2012; Michel, 2011; Michel et al., 2007; Tavakoli, 2016); During dialogic speaking tasks, students produce less disfluencies of filled pauses, replacements, reformulations, and repetitions (Michel, 2011; Michel et al., 2007), have less pauses and hesitations (Ferrari, 2012); and have shorter length of pauses, faster articulation rates, longer fluent runs, and higher phonation time ratio (Tavakoli, 2016) than monologic speaking tasks.

Researchers cite Pickering and Garrod's (2004) interactive alignment hypothesis, according to which, interlocutors imitate others' spoken production. They create semi-fixed expressions or routines and reduce cognitive demand of making decisions by streamlining language comprehension and production. Compared with dialogues, monologues have increased cognitive demands for spoken production with no partners to rely on (e.g., less time for online planning because listeners' attention wanes) (Tavakoli & Foster, 2008). The limitation is that applications cannot differentiate speakers' voices during discussions so even though dialogic fluency is more relevant to discussion, we can only assess monologic fluency. Even with differentiation, determining who owns the pauses during dialogic speech is challenging. As EDC only assesses spoken fluency during group discussions, there is a limitation insofar as the mobile application measures only monologic fluency and researchers have shown that students tend to be less fluent for monologic fluency than for dialogic fluency.

Highly Proficient Students

Speed fluency increases with L2 proficiency, but a ceiling effect exists around levels B2 and C1

of CEFR whereby students' speed usually stays constant (Tavakoli et al., 2020). Additionally, fluency judgments no longer become a determining factor of L2 speaking performance assessment for levels C1 and above as raters consider other aspects such as speakers' accuracy or complexity (Tavakoli & Wright, 2020). Therefore, upper EDC Level 1 students might not increase their speed fluency and even if they do increase it, it might not be a determining factor for overall L2 speaking performance ratings. From our own observations, Level 1 students' pausing behavior seems natural rather than a result of possible L2 disfluencies. Building spoken fluency and using a mobile application to track fluency might not be as justifiable for these students; however, other course goals are still important such as building confidence in their ability to use English for communication, to better express their ideas, and respect others' opinions.

Characteristics of Fluency Development

Larsen-Freeman (2020) stresses the importance of acknowledging that language and language learning is complex and dynamic. Spoken fluency, one aspect of language, is also characterized by these attributes as its development is dynamic and nonlinear with complex factors such as individual differences impacting it. For example, influencing factors can be internal such as motivation, personality, language learning aptitude, and L1 speech rate whereas external factors can be background noise from other students or ease or familiarity with topics, as well as the speaker's interlocutor (Tavakoli & Wright, 2020). There are many factors that affect students' fluency, and because of this, students might not always see increases in speed and decreases in pause frequency. Although strong positive correlation exists between objective measurements of L2 overall proficiency and perceived fluency (Bosker et al., 2012; Derwing et al., 2004), EDC is only for one semester (i.e., four months); thus, raising students' awareness of these internal and external factors, and their ability to affect students' spoken fluency, is important because students might not understand why working hard does not always translate into increases in spoken fluency.

Research Project

The application design was preceded by an ongoing study using an existing application, SpeechRate by Yuri Salomatin (2019), to record and measure the words spoken per minute by students during 3/2/1. We chose this application after comparing applications and software; SpeechRate was the most accurate for calculating the words spoken per minute while also having benefits of being accessible offline and free on iOS and Android. Participants were 20 first-year university students at Rikkyo University, and they recorded each round of their 3/2/1 fluency activity in the beginning of the lesson. SpeechRate did not store data, so students had to enter word counts on Google Sheet. To account for the additional time to conduct this activity, 3/2/1 had to be reduced to 2/1.5/1. After the activity, they could write notes about their performance in English or Japanese. At the end of the semester, students were asked to fill out a survey discussing their perceptions of using technology to assess their fluency. From our experience, participants asked more questions about developing spoken fluency and stated that they prefer having specific spoken fluency goals.

Conclusion

EDC course designers write that spoken fluency is the most important language-based objective

for the class (Hurling, 2012). Currently, assessing spoken fluency equates to instructors' perceived quickness of students' academic discussion skill use during group discussion performance. Having an application that can monitor students' monologic spoken fluency progress might help students and instructors to better achieve EDC's language-based objective of improving spoken fluency by increasing metalinguistic awareness and providing feedback on performances. Applied linguists characterize spoken fluency as complex and dynamic insofar as it develops nonlinearly, is multi-dimensional, and has a multitude of factors that influence it such as students' L2 proficiency, L1 speech rate, and topic familiarity. Fluency is also considered to be the most difficult aspect of L2 speaking performance to assess according to Kang et al. (2019), but using mobile applications to track fluency might help. There is more work that needs to be done in designing, developing, and implementing the application, but we look forward to the process and value the feedback that we receive from students and other interested parties.

References

Anderson, J. R. (1989). Practice, working memory, and the ACT* theory of skill acquisition: A comment on Carlson, Sullivan, and Schneider (1989). *Journal of Experimental Psychology: Learning Memory, and Cognition, 15*(3), 527-530. https://doi.org/10.1037/0278-7393.15.3.527

Arce, N. P. H., & Valdivia, A. M. C. (2020). Adapting competitiveness and gamification to a digital platform for foreign language learning. *International Journal of Emerging Technologies, 15*(20), 194-209. https://doi.org/10.3991/ijet.v15i20.16135

Arevat, S., & Nation, I. S. P. (1991). Fluency improvement in second language. *RELC Journal, 22*(1), 84-94. https://doi.org/10.1177/003368829102200106

Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal, 45*(3), 221-235. https://doi.org/10.1177/0033688214546964

Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., & de Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed, and repairs. *Language Testing, 30*(2), 159-175. https://doi.org/10.1177/0265532212455394

Bradlow, A. R., Kim, M., & Blasingame, M. (2017). Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. *Journal of the Acoustical Society of America, 141*(2), 886-889. https://doi.org/10.1121/1.4976044

Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning: Topic modeling and knowledge mapping. *Educational Technology and Society, 24*(1), 205-222. https://www.jstor.org/stable/26977868

Curran, P. (2019). Raising fluency awareness. *New Directions in Teaching and Learning English Discussion, 7,* 14-18.

Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaee, E., & Noroozi, O. (2021). Using gamification to support learning English as a second language: A systematic review. *Computer Assisted Language Learning, 34*(7), 934-957. https://doi.org/10.1080/09588221.2019.1648298

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning, 54*(4), 655-679. https://doi.org/10.1111/j.1467-9922.2004.00282.x

Duran-Karaoz, Z., & Tavakoli, P. (2020). Predicting L2 fluency from L1 fluency behavior: The case of L1 Turkish and L2 English speakers. *Studies in Second Language Acquisition*, 42(4), 671-695.

https://doi.org/10.1017/S0272263119000755

English Discussion Committee Handbook. (2022). *English discussion class handbook*. Tokyo, Japan: Rikkyo University.

English Discussion Committee Syllabus. (2022). *English discussion class syllabus*. Tokyo, Japan: Rikkyo University.

Ferrari, S. (2012). A longitudinal study of complexity, accuracy, and fluency variation in second language development. In A. Housen, F. Kuiken, and I. Vedder. (Eds), *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy, and fluency in SLA* (pp. 277-298). John Benjamins. https://doi.org/10.1075/lllt.32.12fer

Flores, J. F. F. (2015). Using gamification to enhance second language learning. *Digital Education Review, 27*, 32-54. https://doi.org/10.1344/der.2015.27.32-54

Garside, P. (2020). The effects of time pressure on fluency, complexity, and accuracy: A case study. *Journal of Foreign Language Education and Research, 1*(1), 31-45.

Gatbonton, E., & Segalowitz, N. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly, 22*, 473-492. https://doi.org/10.2307/3587290

Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing, 27*(3), 379-399. https://doi.org/10.1177/0265532210364407

Heil, C. R., Wu, J. S., Lee, J. J., & Schmidt, T. (2016). A review of mobile language learning applications: Trends, challenges, and opportunities. *The EUROCALL Review, 24*(2). https://doi.org/10.4995/eurocall.2016.6402

Huang, W. H. Y., & Soman, D. (2013). Gamification of education. *Research Report Series: Behavioural Economics in Action. Rotman School of Management*, University of Toronto.

Hurling, S. (2012). Introduction to EDC. *New Directions in Teaching and Learning English Discussion*, 1(1), 2-10.

Jaruka. (2022). Carrots Icon [format]. Canva. https://www.canva.com.

de Jong, N. H. (2018). Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly, 15*(3), 237–254. https://doi.org/10.1080/15434303.2018.1477780

de Jong, N., & Perfetti, C. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning, 61*(2), 533-568. https://doi.org/10.1111/j.1467-9922.2010.00620.x

Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning, 64*(4), 809-854. https://doi.org/10.1111/lang.12084

Kahng, J. (2018). The effect of pause location on perceived fluency. *Applied Psycholinguistics, 39*(3), 569-591. https://doi.org/10.1017/S0142716417000534

Kang, O., Rubin, D., & Kermad, A. (2019). The effects of training and rater differences on oral proficiency assessment. *Language Testing, 36*(4), 481-504. https://doi.org/10.1177/0265532219849522

Kerismaker. (2022). Spinach [format]. Canva. https://www.canva.com.

Kita, S., Sturges, J. G., Arthurson, D., & Truxel, D. (2022). *What's your opinion? Interactive skills for effective discussion.* Tokyo, Japan: Rikkyo University.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of

second language learners. System, 32(2), 145-164. https://doi.org/10.1016/j.system.2004.01.001

Kukulska-Hulme, A. (2016). *Personalization of language learning through mobile technologies.* Cambridge, UK: Cambridge University Press.

Larsen-Freeman, D. (2020). Complexity theory. In B. VanPatten, and J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 248-270). New York, NY: Routledge.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), Perspectives on fluency (pp. 25-42). Ann Arbor, MI: The University of Michigan Press.

Maurice, K. (1983). The fluency workshop. *TESOL Newsletter, 17*(4), 29.

Memon, M. (2019, February 5). *16 important UX design principles for newcomers*. Springboard. https://www.springboard.com/blog/design/ux-design-principles/

Michel, M. C. (2011). Effects of task complexity and interaction on L2 performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance*, (pp. 141-173). John Benjamins Publishing Company.

Michel, M. C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *IRAL-International Review of Applied Linguistics in Language Teaching, 45*(3), 241-259. https://doi.org/10.1515/iral.2007.011

Molina Barriga, M., & Briesmaster, M. (2017). The use of the 3/2/1 technique to foster students' speaking fluency, *i.e.: inquiry in education, 9*(2), 1-13.

Pickering, M. J., & Garrod, S. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8-11. https://doi.org/10.1016/j.tics.2003.10.016

Red-Hawk. (2022). Rabbit Line Logo [format]. Canva. https://www.canva.com.

Rego, I. (2015). Mobile language learning: How gamification improves the experience. *Handbook of Mobile Teaching and Learning?* https://doi.org/10.1007/978-3-642-54146-9_76

Saeki, M., Matsuura, R., Suzuki, S., Miyagi, K., Kobayashi, T., & Matsuyama, Y. (2021). InteLLA: A speaking proficiency assessment conversational agent with adaptive interview strategy). 第12回対話システムシンポジウム.

Salomatin, Y. (2019). *SpeechRate* (1.0.3) [Mobile app]. App Store. https://apps.apple.com/us/app/speechrate/id1450633756

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*(4), 357-385. https://doi.org/10.1017/S0272263100011189

Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching, 54*(2), 79-95. https://doi.org/10.1515/iral-2016-9991

Segalowitz, N. (2010). *Cognitive Bases of Second Language Fluency.* New York: Routledge. https://doi.org/10.4324/9780203851357

Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics, 14*(3), 369-385.

Shrosbree, M. (2020). The relationship between L1 fluency and L2 fluency among Japanese advanced early learners of English. *Journal of Foreign Language Education and Research, 1*(1), 111-117.

Skehan, P., Foster, P., & Shum, S. (2016). Ladders and Snakes in Second Language Fluency. *International Review of Applied Linguistics in Language Teaching, 54*(2). https://doi.org/10.1515/iral-2016-9992

Skehan, P. (2003). Task-based language instruction. *Language Teaching, 36,* 1-14. https://doi.org/10.1017/S026144480200188X

Suzuki, S., & Kormos, J. (2020). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition, 42*(1), 143-167. https://doi.org/10.1017/S0272263119000421

Suzuki, S., Kormos, J., & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal, 105*(2), 435-463. https://doi.org/10.1111/modl.12706

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal, 65*(1), 71-79. https://doi.org/10.1093/elt/ccq020

Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching, 54*(2), 133-150. https://doi.org/10.1515/iral-2016-9994

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 58*(2), 439-473. https://doi.org/10.1111/j.1467-9922.2008.00446.x

Tavakoli, P., & Hunter, A. M. (2018). Is fluency being 'neglected' in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research, 22*(3), 330-349. https://doi.org/10.1177/1362168817708462

Tavakoli, P., Nakatsuhara, F., & Hunter, A. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal, 104*(1), 169-191. https://doi.org/10.1111/modl.12620

Tavakoli, P., & Skehan, P. (2005). 9. Strategic planning, task structure and performance testing. *Language Learning and Language Teaching*, 239-273. https://doi.org/10.1075/lllt.11.15tav

Tavakoli, P., & Wright, C. (2020). *Second Language Speech Fluency: From Research to Practice.* London: Cambridge University Press.

Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly, 50*(2), 369-393. https://doi.org/10.1002/tesq.232

Tran, M. N., & Saito, K. (2021). Effects of the 4/3/2 activity revisited: Extending Boers (2014) and Thai & Boers (2016). *Language Teaching Research,* https://doi.org/10.1177/1362168821994136

Watanabe, H., & Long, R. (2019). Is there an increase in fluency over levels of proficiency? A study of English L2 usage of Japanese university students. *International Journal of Humanities, Arts and Social Sciences, 5*(6), 215-223. https://doi.org/10.20469/ijhss.5.20001-6

【Research Brief】

Drawing the Line: Integrating Kialo to Deepen Critical Thinking in Debate

Jon Mahoney

Abstract

This study reflects on using the online debate site Kialo as a supplementary tool to elicit students' opinions about various debate topics. In total, 118 students took part in the study. A mixed methods approach was utilized to collect both qualitative and quantitative data in the form of class notes and a Google Form, respectively. In general, students gave positive impressions about using Kialo, suggesting that it helped improve the content of their debates, their critical thinking skills, and their written English. Results from this study suggest that integrating the Kialo platform into the debate class syllabus increased students' participation and satisfaction from the course. In future, Kialo could be integrated into other English classes as a means of a) a reflective tool, b) a place for students to brainstorm ideas about projects and discussions, and c) a platform to help the students achieve their course goals.

Keywords: *Kialo, critical thinking, debate*

INTRODUCTION

The English debate module is a 14-week course taken by all first-year students at Rikkyo University in Tokyo, Japan. This module has been mandatory since 2020, with the primary goals being to (i) understand the nature and structures of debate in English, (ii) to develop students' critical thinking skills by analyzing and formulating arguments on issues from multiple perspectives, and (iii) to help students learn how to respond to questions through the development of research skills (Debate Committee, 2020). Students are expected to learn to practice and use the academic skills that they study in discussion, reading, writing, and presentation classes in a social setting. They are also required to engage in critical thinking and logical thinking and improve listening, research, and team building skills (Debate Committee, 2020).

English debate classes present students with a chance to communicate using academic English in a structured manner. In contrast to merely absorbing information, debate demands that students actively apply information in a meaningful way (Kennedy, 2007). All four of the English skills are practiced when debating in EFL (English as a Foreign Language), as well as providing an opportunity to practice language skills in an authentic situation (Alasmari & Ahmed, 2013). Good debate requires higher order critical thinking skills and offers an opportunity for students to move beyond the acquisition of primitive knowledge in a subject matter (Elliot, 1993). Due to the fluid nature of debate with the turn taking between teams, it is unfeasible to completely prepare for a debate. Therefore, spontaneous use of English by students is advantageous in that it aids in building in tandem both oral communication and critical thinking skills (Combs & Bourne, 1994). Classroom debates authorize students to cooperate with one another while searching for information, which enhances their interactive learning and their reasoning ability (Ebata, 2009; Zare & Othma, 2013), and the competitive nature of debate activities also helps develop teamwork skills and cooperation (Williams et al., 2001). Practicing academic debate in university classes could thus be seen as preparing students to speak in English in meaningful real-life situations, alongside improving their language

and presentation skills in conjunction with cooperative skills.

The aim of this paper is to investigate the impact of using the website Kialo to help boost students' critical thinking skills and practice for their speaking debates. By using the site, students would be able to see the other side's arguments and challenge them, while also honing their ideas and debate skills for the forthcoming speaking debates. Woodward and Padfield (2021) have indicated that using Kialo for debate classes had a positive impact on team collaboration, creative thought, and debate performance. Additionally, Mahoney (2021) found that providing asynchronous discussion boards for students to express their ideas prior to discussion classes gave students the opportunity to practice and refine their ideas and to think more deeply about the discussion topics, which led to the enrichment of in-class discussions. By using Kialo in a similar fashion, the author is aspiring for analogous results in debate class.

KIALO

Kialo is an online debating platform that helps people take part in thoughtful discussion, appreciate different points of view, and assist with collaborative decision making (Kialo, 2020). This portal increases students' critical thinking by making them face opposing views and re-examining their own. They must also ensure that their arguments are well researched and provide reliable evidence. The moderator (or teacher) creates a proposition that participants (students) can either make a support claim (pro) or an attack claim (con). These claims show visual reasoning through a tree-based structure. Participants can also ask follow-up questions to each claim for extra evidence or clarification. This enables a detailed exploration of some claims, in addition to the main thesis topic (Chaudoin et al., 2017).

Figure 1
Kialo argument tree

Note. Tree depiction of a Kialo debate with a single thesis. Green designates pro arguments and red designates con arguments.

PROCEDURE

Each class consisted of approximately 20 students, with each student placed into classes with other students of a similar English competence. One class was level 1, with students who all scored a combined TOEIC listening and reading score of over 680 (CEFR B2 and above), two were level 2 classes with average TOEIC scores of 480–679 (CEFR B1-B2), and three level 3 classes with average scores of 280–479 (CEFR A2-B1). Due to the ongoing COVID-19 pandemic, the semester began with online classes with each class 100 minutes in duration. From lesson five, classes were changed to a conventional face-to-face format. In the classroom, students were organized into groups of four or five, with each team color-coded and given an affirmative or negative side to argue from. Seating charts were created for each lesson beforehand to ensure that students sat next to different classmates in the next lesson.

In the first lesson of the term, students received a 40-minute explanation about Kialo, which

included a 5-minute YouTube video. After this, students registered their own accounts. Students were advised to use their real names, so that their posts and replies could be identified easily. Each Kialo group was named the same as the class code of the class (for example, FT505), and then students received a link that would allow them to join each Kialo group. In this way, only students who were members of the Kialo group, or those who were invited to join the Kialo debates via the link, could gain access and participate in the Kialo debates. Students were informed that participation in Kialo was a weekly homework, which would amount to 10% of their overall grade in the debate class, and that the more they participated and asked questions, the higher this score would be.

The first topic, "Junk food should be banned on all campuses" was chosen as it was a relatively simple proposition and something that students could easily relate to. The second topic "Eating less meat is a good way to be eco-friendly" was chosen as it was also food-related and followed the suggested themes of the textbook. New propositions were posited weekly throughout the semester on topics that were related to the themes of textbook: food, the environment, technology, gender, and the media. In the early debates, many students read debate arguments from their mobile phones, and some read directly from the Kialo debates during face-to-face discussions. Students were advised that this was not suitable and that they needed to write their claims on Post-it Notes or on folded paper, as well as to have good body language when making their arguments to the other team. From around lesson seven, the higher-level students were advised to avoid putting their strongest arguments on Kialo and save them for the speaking debate to surprise the other team with a powerful piece of evidence. The students immediately understood and followed this instruction until the end of term.

In later lessons, students were allowed to write their own propositions on Kialo. For the final debate test, the students were asked to propose topics of their choice on Kialo, and then they were allowed to vote for their favorite topics, with the most popular choice being selected as the final debate test propositions. The students produced a rich variety of propositions ranging from topics such as "The existence of aliens" to "Zoos and aquariums should be banned." It was found that giving the students the power to create and vote for their own topics was both motivating and rewarding.

DISCUSSION AND FINDINGS

Student participation

In total, the participating students (n= 118) made 1873 contributions (claims & replies), which is an average of 15.9 contributions per student. Six students (5%) made 0–5 contributions, while nineteen students (16%) made 60 contributions. Forty-five students (38%) made 11–15 contributions, and twenty-three students (19%) made 16–20 contributions. Sixteen students (14%) made 21–25 contributions, and five students (4%) made between 26–30 contributions. Four students (3%) made between 41–45 contributions, and one student made between 51–55 contributions. A total of 1328 (posted and completed) claims were made by the students. Nine students (8%) made 0–5 claims, while thirty-four students (29%) made 6–10 claims. Fifty-three students (45%) made 11–15 claims, and sixteen students (14%) made 16–20 claims. Six students (5%) made 21–25 claims. A total of twenty-six posts were deleted. All these deleted posts took place in the first three debates, suggesting that students were making mistakes more often earlier in the semester when they had just started using the platform.

Topics

The most popular topic that students participated in with 219 contributions was in week nine, with “Debate topics.” In this week, two different debate topics from the textbook were posited: “The benefits of social media are greater than the risks” and “Playing video games causes violent behaviour.” The fact that lesson nine was a test was perhaps a major factor for participation in this week to be the greatest, along with the two relatable topics for students. Week four (“Everyone should purchase an electric car”) was the second most popular topic with 213 contributions. Weeks two and three (“Junk food should be banned on all campuses” /“Eating less meat is a good way to be eco-friendly”) were popular topics with 193 contributions each. As car manufacturing is a main industry in Japan, it could have figured in the popularity of week four. Weeks two and three may have been popular as these took place early in the term so students’ motivation may have been high, and the topics are food-related, which is a topic that students tend to enjoy discussing. The least popular topics were in weeks 13 and 14 (“Final debate topics”/“Kialo is a useful tool for debate class”), with 84 and 96 contributions. respectively. This drop off may well be attributed to the fact that the final two classes took place after the winter holiday, and students’ motivation may have focused more on the final speaking debate, which took place in week 13.

Equal participation

Of the 87 female students, a total of 1358 contributions were made (16.6 per student), and 962 claims (11.1 per student) were made. Out of the 31 male students in the study, a total of 515 contributions were made (16.6 per student) and 366 claims were made (11.8 per student). These figures would suggest that participation between genders was almost on par.

Advantages

In the final class of the term, students were asked to complete a Google Form, in which the final question was open-ended asking students to leave any positive or negative comments about using Kialo. This same question was also posed as the final Kialo topic. In a combined total of the Google Form and Kialo debate, 20 different advantages were identified by the students. The most common was that Kialo was useful for the class (78 comments). The second most popular advantage given (43 comments) was that Kialo deepened the debate and was motivating. The third most common advantage (25 comments) was that the platform was easy to use and kept track of everyone’s opinions. Below are some of the positive comments made:

Using Kialo helped me cultivate a critical perspective.
Being able to see other people’s opinions helped change my inflexible mind.
I could improve my critical thinking skills by researching deeply and seeing different sides to an argument.
If I entered the Kialo debate late, it was difficult to come up with other ideas, so I had to research deeply.
It was useful to check anytime and helped me to think of rebuttals.
It was easy to use and a good place to practice before the class.

Students were also asked to agree or disagree on a Likert scale with a series of statements in the Google Form. For the statement "Using Kialo helped improve my debate and critical thinking skills," 82% (n= 84) of students agreed that it had helped improve them. Only 5% of students disagreed with this statement, clearly indicating that using Kialo had been very beneficial to most students' debate and critical thinking skills. For the statement "Seeing people's ideas on two opposing sides helped deepen my understanding of the debate topics," 87% (n=89) of students agreed with this. Having a clear line between pros and cons and therefore separating two sides of an argument helped students to carve their stance internally on each topic. 82% (n=84) of students agreed with the statement "It was useful to practice my debate arguments on Kialo before the speaking debates." The feedback suggests that it was especially useful for the lower students to practice, who often used the same ideas on Kialo in the speaking debates.

Disadvantages

A total of seven different disadvantages were indicated on the one Google Form open-ended question and final Kialo debate. The most common disadvantages given was the 500-word limit for each post (12 students). The second most common disadvantage was trouble with using the site or PC (seven students). This may partly be due to that fact that the Kialo platform added new features during the term, and these changes were not clearly explained. Six students from the level 1 class expressed that being able to see the other team's ideas was detrimental to the speaking debates. These views were also expressed in the final class when students discussed the question "What are the advantages and disadvantages of using Kialo?" Below are some of the negative comments made:

I think that seeing the other team's ideas before is cheating. It is good to prepare, but our opinion can be copied and stolen. This makes the debates superficial.
It is not good to know the opposite opinion before the debate. It should be a secret. We should keep the important data for the speaking debate.
I cannot use my Kialo opinion in the speaking debate because the other team can make rebuttals easily.
The 500-word limitation is not good. I wanted to express my opinions in more detail.

Although students were reminded to participate at the end of each lesson, and good examples of their arguments were highlighted in class, five students mentioned that it was easy to forget. Although the platform was embedded on the students' class homepage on BlackBoard and a direct link to each debate was shared for weekly feedback, class feedback and the BlackBoard class homepage may have been largely ignored by some students.

Differences in perception

There was a clear difference in how the higher-level students of this course perceived the usage of Kialo compared with those of a lower level. It was notable that students who were placed in the level 1 class (TOEIC scores of over 680) and high level 2 classes (TOEIC scores of 480–679) stressed that they wanted to keep their opinions confidential, to have more authentic debates. These high-level students also wanted extended word limits, so they were far more enthusiastic. Most of the level 3 students (TOEIC scores of 280–479) mentioned that it was hard to find new ideas if they did not

post an opinion early in the debate. The level 3 students also indicated that seeing the ideas before was more beneficial to them. The level 3 students tended to use the same ideas on Kialo in the speaking debate, whereas the higher levels did not. These findings would seem to suggest that in future usage of this platform, the unique needs of each class will need to be considered more carefully, and debates should be arranged according to the proficiency in English of each class.

CONCLUSION

This study has examined the use of Kialo in a debate class and has considered the possible permutations for its continued usage. To concur with Matsumoto (2021, p.171) who claims that "Technologies like smartphones seem to have the potential to expand L2 learners' choices and agency for learning," the implementation of the Kialo platform into the debate course has aided students in the attainment of these skills. By giving students a week to research about the forthcoming topic of the speaking debate, students were able to delve deeply into the topics and become more flexible in their way of thinking. The use of Kialo falls in line with what Guilloteaux and Dornyei (2008) refer to as "the motivation orientation of language teaching (MOLT)." MOLT identified 25 motivational practices used by teachers in 27 countries (Lightbown & Spada, 2013), which included primarily (i) teacher discourse to arouse student curiosity, which would involve choosing appropriate topics and introducing them (ii) participation structure, such as group work or pair work (iii) activity design, which would refer to team competition, and (iv) encouraging positive retrospective self-evaluation and activity design, which refers to the positive feedback given to students on Kialo or at the start of each class. Overall, student participation on Kialo was active, and most students did more than what was required of them to earn participation points in their score, suggesting that the platform pushed the students to participate in a variety of ways. Since all the participants owned a smartphone, they could comfortably access the site.

As argued in this paper, it is important that when using Kialo, the teacher should tailor the usage of the platform to the needs of each class and, depending on the English competence of each class should offer students pedagogical guidance. By augmenting Kialo as part of their course, teachers may be able to a) facilitate students' critical thinking skills, b) assist students in the planning and output of their face-to-face debates, and c) authorize improvement in students' motivation. From the findings of this study, it could be argued that Kialo was a useful tool and is a project that will continue to grow organically in the age of information. Further research could also be conducted in more general English classes, and not only in just debate classes, as the platform is malleable and could be used to boost critical thinking skills in various English classes, such as English presentation, discussion, or reading and writing classes. It may be potentially used as a reflective tool, as a platform for students to collectively brainstorm ideas for forthcoming classes, or as a means of providing peer-feedback.

REFERENCES

Alasmari, A., & Ahmed, S. S. (2013). Using debate in EFL classes. *English Language Teaching*, *6*(1), (pp. 147-152).

Chaudoin, S., Shapiro, J., & Tingley, D. (2017). Revolutionizing teaching and research with a structured debate platform. *Journal of Political Science*, *58*, (pp. 1064-1082).

Combs, H., & Bourne, S. (1994). The renaissance of educational debate: Results of a five-year study

of the use of debate in business education. *Journal on Excellence in College Teaching*, 5(1), (pp. 57-67).

Debate Committee. (2020). Debate instructor handbook fall 2020. *Center for Foreign Language & Research.* Rikkyo University.

Ebata, M. (2009). Content-based instruction. *Teaching English as a Second Language: A New Pedagogy for a New Century*, (pp. 129-135).

Elliot, L. B. (1993). Using debates to teach the psychology of women. *Teaching of Psychology*, *20*(1), (pp. 35-38).

Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly*, *42*(1), (pp. 55-77).

Kennedy, R. (2007) In-class debates: Fertile ground for active learning and the cultivation of critical thinking and oral communication skills. *International Journal of Teaching and Learning in Higher Education. 19*(2), (pp. 183-190).

Kialo. (2020). Kialo Edu—The tool to teach critical thinking and rational debate. Kialo. https://www.kialo-edu.com/

Lightbown, P., & Spada, N. (2013). How languages are learned (4th ed.). Oxford: Oxford University Press.

Matsumoto, Y. (2021). Student self-initiated use of smartphones in multilingual writing classrooms: Making learner agency and multiple involvements visible. *The Modern Language Journal*, *105*(S1), (pp. 142-174).

Mahoney, J. (2021) Using asynchronous discussion boards to complement online discussion classes. *Journal of Multilingual Pedagogy and Practice 1* (1), (pp. 56-64).

Williams, D. E., McGee, B. R., & Worth, D. S. (2001). University student perceptions of the efficacy of debate participation: An empirical investigation. *Argumentation and Advocacy*, *37*(4), (pp. 198-209).

Woodward, H., & Padfield, L. (2021). A blended approach to flipped learning for teaching debate. *Journal of Multilingual Pedagogy and Practice*, *1* (1), (pp. 44-55).

Zare, P., & Othman, M. (2013). The relationship between reading comprehension and reading strategy use among Malaysian ESL learners. *International Journal of Humanities and Social Science*, *3*(13), (pp. 187-193).

Appendix
Student Survey

1. Using Kialo helped improve my critical thinking skills
   Strongly Agree　Agree　Neutral　Disagree　Strongly Disagree

2. It was useful to practice my debate arguments on Kialo before the speaking debates
   Strongly Agree　Agree　Neutral　Disagree　Strongly Disagree

3. Seeing people's ideas on two opposing sides helped deepen my understanding of the debate topics
   Strongly Agree　Agree　Neutral　Disagree　Strongly Disagree

4. Using Kialo motivated me to research deeply for the team debates
   Strongly Agree　Agree　Neutral　Disagree　Strongly Disagree

5. Using Kialo was troublesome and not useful for debates
   Strongly Agree　Agree　Neutral　Disagree　Strongly Disagree

6. It was difficult to begin the Kialo debates
   Strongly Agree　Agree　Neutral　Disagree　Strongly Disagree

7. I felt unhappy if someone disagreed with me on Kialo
   Strongly Agree　Agree　Neutral　Disagree　Strongly Disagree

8. Using Kialo did not help my understanding of debate structures and flow
   Strongly Agree　Agree　Neutral　Disagree　Strongly Disagree

9. Please write any positive or negative comments about using Kialo here

【Research Brief】

Translanguaging in Discussion Class: Investigating the Viability of a Bilingual Pedagogy in a Japanese University EFL Context

Omar Shelesh

Abstract

This study originated from reflective teaching practices tentatively engaging with the bilingual-focused pedagogy of translanguaging in a Japanese university setting. While initial findings have been encouraging, there is a clear need for more empirical research into the viability of translanguaging in this context. Therefore, in light of its initial research findings, it is the purpose of this study to propose a suitable research setting, appropriate research questions, and practical suggestions for a research methodology.

Keywords: *translanguaging, English discussion, bilingualism, Japanese university*

Introduction

The linguistic segregation of learners' first language (L1) from the second "taught" language (L2) in educational settings has long been accepted as the norm. In English language teaching (ELT) contexts, in particular, practitioners have continued to maintain traditional methods and approaches in ELT, teaching the target language in immersive, English-only environments; any use of the learners' L1 is considered to be counterproductive "interference" (Ooi & Abdul Aziz, 2021). However, interest in the advantages of incorporating learners' L1 into the L2 teaching and learning process has generated a great deal of excitement among scholars in recent years (see Cummins, 2007; Lin, 2015), adding momentum to the progressive notion that considers all second language learners as "emergent bilinguals" (García, 2009). This ideological movement aims to establish bilingualism as a key objective of conventional foreign language education, through the normalization of the combined and deliberate use of multiple languages in the foreign language classroom (Canagarajah, 2013; García & Wei, 2014). This has consequently raised interest in the development of effective bilingual teaching and learning pedagogies, with the concept of *translanguaging* (Cenoz & Gorter, 2017, 2022; García & Wei, 2014) receiving much attention from educational scholars and practitioners. However, despite this extensive interest generating a substantial body of research, there is still a lack of context-specific studies documenting the viability of pedagogical translanguaging in Japanese tertiary EFL education, and English oral communication contexts, in particular. Therefore, this study proposes that a further examination of the viability of this emerging bilingual pedagogy be carried out. This study introduces literature pertaining to the theoretical context of translanguaging, referencing relevant prior research. This will be followed by an evaluation of the nominated EFL course as the potential subject of this study, a presentation of some pertinent research questions, along with practical suggestions for a research methodology. The study concludes with a brief discussion of the initial findings.

Background

The term *translanguaging* is a prefix-extended form of *languaging*, a concept introduced by Swain (1985) and later developed by scholars such as Becker (1988) who defines language not only as a code or system of rules but as a dynamic and potentially limitless process of context-sensitive communication and interaction. With this concept of languaging in mind, it was Welsh educationalist Cen Williams who, in the course of conducting research on bilingual secondary education in Wales in the 1980s, conceived of the Welsh term *trawsiethu* (Williams, 1994), as an all-encompassing label for the complex bilingual pedagogical practices that he was studying among Welsh- and English-speaking learners. Although Williams' idea at first only appeared in an unpublished doctoral thesis, the term was subsequently translated into English by scholar Colin Baker, adopting the word *translanguaging* (Baker, 2001), after which the concept would receive wide recognition from other scholars and academics. This initiated a paradigm shift in bi- and multilingual education exploration, which is still evolving to this day, as evidenced in a significant—and increasingly complex—body of contributory research from scholars, such as Baker (2001, 2011), Canagarajah (2011, 2013, 2018), Cenoz and Gorter (2017, 2022), García (2009), to name a few.

In order to gain a more practical and in-depth understanding of what translanguaging means in its original educational context, we should look to one of Baker's many contributions to the conceptual arena. It is here that he describes four important potential benefits of translanguaging, namely,(1) promoting a more complete understanding of the subject matter; (2) promoting the development of L2, or the "weaker" language; (3) enhancing links and co-operation between learners' homes and institutions; and (4) facilitating the integration of advanced speakers with beginners (Baker, 2001, as cited in García & Lin, 2017). Therefore, translanguaging can be described from a teaching and learning perspective (and relating to benefit points [1] and [2] specifically), as an applied process in which the subject matter or content of a lesson is taught through the use of two languages. More specifically, the subject matter is presented in one language, after which learners can demonstrate their understanding by producing it in another (Baker, 2011). As Baker explains in this often-quoted example, "To read and discuss a topic in one language, and then write about it in another, means that the subject matter has to be processed and 'digested'" (Baker, 2011, p. 289). This notion of *internalization* taking place within bilingual learners as they are translanguaging is a key tenet of the concept and should not be underestimated. Moreover, as seemingly innocuous and simple as some of these practices appear to be, translanguaging activities require that learners activate and engage with highly complex cognitive processes on both a conscious and subconscious level. This suggests that while translanguaging practices could be relatively simple to implement, they may induce profound and transformative changes within bilingual learners (García & Wei, 2014).

Extending the concept

While the meaning of the "trans" prefix of translanguaging mostly aligns with the characteristics and practices described thus far, that is meaning to go "across", or "between" languages as per Williams' (1994) original concept, it was academic Ofelia García, who extended the meaning to include the notion of "beyond", by which the concept transcends the classroom and pervades the wider world (Cenoz & Gorter, 2022). García, a North American-based academic responding to the complex sociolinguistic realities of bilingualism in the U.S., saw fit to redefine translanguaging as "multiple discursive practices in which bilinguals engage in order to make sense of their bilingual

worlds" (García, 2009, p. 45). The context of García's description is based on the experiential practices of bilingual English–Spanish speakers, as they navigate and negotiate all aspects of their lives through the prism of two different languages. Extending the translanguaging concept in this way would have far-reaching implications for the movement as a whole, promoting further discourse and delineation of its features and characteristics, giving rise to the notion of *weak* and *strong* translanguaging (García & Lin, 2017). These terms form a key part of the theoretical research for this brief, and can be defined as follows:

- *Weak translanguaging*, primarily used in educational contexts, represents Williams' (1994) original translanguaging concept that acknowledges that bilingual learners have two official, "separate" languages. However, figuratively speaking, engaging in translanguaging can cause the boundary between the two to become "soft" and "permeable", as observed by Williams (1994) in his study of the bilingual practices of English- and Welsh-speaking learners.
- *Strong translanguaging* is based on García's (2009) extended theory of the concept, which states that bilingual people do not speak separate languages, but possess a single language "repertoire", the features of which they deploy selectively, and at will, in response to the demands of a specific communicative context (García & Lin, 2017).

The notions of weak and strong translanguaging have subsequently become commonly used expressions when describing translanguaging practices in education and will be referenced in subsequent sections of this brief. Furthermore, the application of translanguaging practices in educational contexts will now be referred to as *pedagogical translanguaging*, elaborated in the following section.

Pedagogical translanguaging

Pedagogical translanguaging is a term denoting the specific application of translanguaging practices in the classroom. In their authoritatively written guide on the subject, Cenoz and Gorter (2022) reinforce the definition of pedagogical translanguaging as, "a theoretical and instructional approach that aims at improving language and content competences in school contexts by using resources from the learner's whole linguistic repertoire" (Cenoz & Gorter, 2022, p. 1). They also posit that "[t]ranslanguaging is learner-centered and endorses the support and development of all the languages used by learners. It fosters the development of metalinguistic awareness by softening boundaries between languages when learning languages and content" (Cenoz & Gorter, 2022, p. 1). It is clear from this description, which includes phrases such as "whole linguistic repertoire" and "softening boundaries", that Cenoz and Gorter readily accept both the strong and weak notions as a framework through which pedagogical translanguaging practices can be described. Going further, they explain that pedagogical translanguaging practices are designated as strong or weak, "… depending on the degree of pedagogical intervention that takes place in the process of learning…" (Cenoz & Gorter, 2022, p. 30). This definition is particularly useful, as the practical research proposed for this brief aims to utilize certain strong interventional practices which can raise learners' "metalinguistic awareness", that is using more than one language in a class to teach specific subject matter (Cenoz & Gorter, 2022, p. 31). Overall, what is understood from Cenoz and Gorter's contribution, is that pedagogical translanguaging, particularly in the context of EFL, can be a flexible pedagogical process, over which the teacher has a considerable degree of governance. Therefore,

through deliberate planning and activity design, the teacher can leverage learners' metalinguistic awareness to simultaneously achieve goals in both language acquisition and subject knowledge, aims that would be otherwise difficult to reach in a purely monolinguistic setting.

Translanguaging in a Japanese educational context

While this study provides an introduction and theoretical context to the concept of translanguaging, it still only represents a small fraction of the academic discourse available on the subject. However, as the author has discovered, there is still an apparent scarcity of concrete, context-specific studies of the practical application of pedagogical translanguaging in tertiary foreign language education, particularly in settings such as English education in Japanese university. There is one important exception to this, however. It is a case study conducted by Blake Turnbull (2019) who investigated the effects of weak and strong translanguaging practices in the planning of English academic and creative writing within a Japanese university's EFL program. Turnbull's study demonstrates the effective operationalization of pedagogical translanguaging in an as yet comparatively untested academic context and presents new insights and avenues of potential inquiry into its strategic application.

The methodology of Turnbull's study was to use a relatively small sample of two English writing classes consisting of first-year Japanese EFL students. The classes were sub-divided into three groups, each group being tasked with a discussion activity within which students were to plan for writing an essay on a given topic. Importantly, the first group was limited to preparatory discussions using monolingual English only (no translanguaging); the second group could use a weak form of translanguaging, and the third group could engage in strong translanguaging (Turnbull, 2019, p. 237). For empirical data collection, a mixed-methods approach was used involving qualitative audio conversation analysis of student discussions, as well as quantitative analysis of students' composition assessment scores (Turnbull, 2019, p. 242).

In addition to providing a strong basis upon which further studies could be modeled, Turnbull's research methodology yielded important findings consistent with notions advocated by the conceptual literature on multi-competence, as espoused by Cook (1991) and García and Wei's (2014) theories on translanguaging. For example, Turnbull was able to observe significant differences in writing scores between students using monolingual English (L2) only and those engaging in strong translanguaging practices (Turnbull, 2019, p. 245). He concludes that "The freedom and confidence that allowing [foreign language] learners to engage in translanguaging practices as the emergent bilinguals they are not only raises their ability to produce the [target language] in desired settings but also affords them the ability to express themselves, to make meaning, and to learn as whole individuals acting in their bi- or multilingual worlds" (Turnbull, 2019, p. 248).

Building upon the encouraging conclusions drawn in Turnbull's study, it is the intention of this brief to propose an expansion of research in translanguaging pedagogies to encompass practical English oral communication programs within a similar Japanese university's EFL setting. For the purposes of this research brief, a specific English discussion course at a Japanese university was proposed as the teaching context and subject of investigation.

Context: The Discussion class

The Discussion class (DC) at Rikkyo University is a compulsory course for all first-year

undergraduate students, consisting of practical, discussion-based English language classes delivered weekly, over a 14-week semester. The program's original concept and curriculum documents, authored by Hurling (2012), stipulate that a near-identical procedural and methodological communicative language teaching approach is followed by instructors, whatever the topical or language learning objectives may be. Teaching groups are divided up according to their faculty of major study, and designated one of four ability levels, based on the student's TOEIC test performance: Level I (TOEIC score band 680 or above) to Level IV (TOEIC score band below 280). A level-specific textbook is also issued to all students, chronologically covering all target language and discussion topics for the semester. All scheduled classes last for 100 minutes and are typically made up of between 9 and 10 students, which is intended to facilitate greater student-centered learning and student-to-student interaction time. The intimate group learning environment is an essential element of the course, as one of its primary aims is to develop fluency through maximizing opportunities for student interaction and enable students to better participate in the exchange of views by performing various oral functions commonly utilized in discussions (Hurling, 2012). These functions are introduced in the DC curriculum as Discussion Skills (e.g., giving and asking for opinions, reasons, and examples) and Communication Skills (e.g., appropriately reacting to others' ideas and checking understanding). In a typical lesson, students are presented with new Discussion Skills phrases (e.g., "In my opinion, ...") alongside a predefined topic on a contemporary issue, through which they practice and apply the new language. This is enacted through the DC lesson plan, which follows a set menu of practical and interactive stages, that is the Fluency stage; the Function Presentation stage; the Practice stage; and two main production stages, Discussion 1 and Discussion 2 (Hurling, 2012). In addition, students are more formally assessed on their ability to apply the discussion skills at regular intervals throughout the semester, by way of a summative discussion test.

In order to ascertain the suitability of the DC as a candidate for conducting pedagogical translanguaging research studies, the course was evaluated against three qualifying criteria:

(1) The course's inherent compatibility in the form of evidence of pre-existing bilingual or translingual concepts available in the theoretical framework of the course.
(2) Potential for the course to accommodate pedagogical translanguaging strategies on a practical level, without unduly disrupting standard lesson procedures and/or timing.
(3) Potential for the course to accommodate pedagogical translanguaging strategies without disrupting or undermining the functional, linguistic objectives of the course.

As for criterion point (1), Hurling's (2012) founding documentation on the course's curriculum design was examined for references to bilingualism or translanguaging that could be exploited, however, no evidence of this was found. The apparent absence of any bilingual considerations at the course's conception could be viewed as a disadvantage; however, it also indicates potential opportunities for the introduction of fresh teaching practices onto fertile new ground. Criterion (2) was evaluated positively, as DC lessons are principally based on uniformly consistent and repeatable teaching methods that would allow for strategic and systematic implementation of supplementary pedagogical translanguaging activities. Moreover, the course's well-resourced institutional context makes it possible to conduct research over multiple lessons and across multiple teaching groups. Criterion (3) was also given a positive assessment, as the linguistic scaffolding provided by certain functional Discussion Skills, that is organizational phrases, such as "Who would like to start?" and "What shall we discuss first?" could facilitate a bilingual framework within which English and the

learners' L1 (in this case Japanese) could be used interchangeably in any given translanguaging task, and without interfering with the target language and/or aims of the lesson.

It should also be mentioned that a research study conducted in this teaching context may benefit wider, institutional-level objectives, as the course is administered by the Center for Foreign Language Education and Research (FLER), which alludes to translanguaging in its core statement of philosophy, citing the concept as part of efforts to promote intercultural understanding through multilingual interaction and "translanguaging dialogues" (The philosophy of the FLER establishment, 2022).

Therefore, taking the above criteria into consideration, the Discussion course was determined to be a suitable testbed for conducting a pedagogical translanguaging research study.

Methodology

Based on the aforementioned previous research studies and proposed research context, the following research questions were formulated in order to provide focus for subsequent research activities. It should be acknowledged that this is not an exhaustive list and that these questions may require adaptation in response to variations in teaching context, and/or in order to optimize research outcomes. However, for the purposes of this brief, the following questions were adopted as a representative sample in this particular discussion class context, which informed suggestions for the research methodology. The research questions are as follows:

(1) *What are the effects on students' English (target language) output in terms of fluency and accuracy when implementing translanguaging pedagogies in an English discussion course?*
(2) *To what extent does the intervention of translingual pedagogies influence students' attitudes toward using L1 in the English discussion classroom?*
(3) *To what extent does the intervention of translingual pedagogies affect students' self-identity as "bilingual" speakers?*
(4) *What are the practical planning considerations when implementing translanguaging pedagogies on an English discussion course?*
(5) *To what extent do teachers' attitudes toward bilingualism and the use of L1 in English classes affect the implementation of translanguaging pedagogies?*

From a methodological perspective, these research questions encompass a wide range of potential data sources, variables, and other influential factors related to the use of translanguaging; therefore, research carried out for this study was assessed to be best conducted under the umbrella of *action research*, combined with a *mixed-methods* approach. Action research is known as "[a] powerful tool for change and improvement" (Cohen, et al., 2007, p. 297) as it encourages a more disciplined investigative approach from researchers, requiring that they "plan, act, observe and reflect more carefully, more systematically, and more rigorously than one usually does in everyday life" (Kemmis & McTaggart, 1992, p. 10, as cited in Cohen, et al., 2007, p. 297-8). Moreover, incorporating a mixed-methods approach into action research enhances the ability of practitioners to triangulate multiple sources by "making use of all available data (both qualitative and quantitative) in order to build a rigorous, cohesive set of conclusions" (James, et al., 2008, p. 81). Therefore, in the course of implementing both strong and weak pedagogical translanguaging strategies, the principles of action research dictate that researchers/practitioners record and respond effectively to what they are observing while students are translanguaging. It is therefore recommended that researchers

maintain documentation in the form of *retrospective field notes* (Murphy, 2014) written shortly after the lesson has finished, also documenting *reflections on* and *for action* (Murphy, 2014). Journaling is a form of qualitative, longitudinal data collection that could also be adapted to involve student participants. For example, students could be asked to make brief weekly journal entries regarding their experiences of how translanguaging is affecting them, answering questions to help researchers gain insights into changing attitudes toward L1 use in the EFL classroom (based on research question 2), as well as expose issues pertaining to bilingualism and identity (research question 3).

Other mixed qualitative and quantitative methods, such as pre- and post-discussion course surveys could be deployed for student participants to complete in order to assess the comparative impact of translanguaging on changing attitudes. Surveys could be extended to teaching staff to obtain further background data concerning attitudes and beliefs toward bilingualism and L1 use in the EFL classroom (research question 5), which could be cross-examined for correlations with corresponding student attitudes toward converging issues.

Preliminary Results and Discussion

As stated previously, this brief is based on preliminary research consisting of exploratory reflective teaching practices tentatively engaging with the bilingual-focused pedagogy of translanguaging. The author initially set out to gauge the practical effectiveness of the pedagogy, as per research question (1) *What are the effects on students' English (target language) output in terms of fluency and accuracy when implementing translanguaging pedagogies in an English discussion course?* To this end, casual experimentation was conducted by devising and implementing certain weak and strong translanguaging practices (García & Wei, 2014) at strategic stages of a discussion lesson, while observing the reactions of students as they were exposed to new bilingual activities, along with any changes in their fluency or accuracy when producing the target language. This yielded noticeable results, as witnessed when conducting preparatory hybrid L1 L2 (Japanese–English) translanguaging discussions prior to the main Discussion stages, in an average ability class. This consisted of deploying the functional target language, that is organizational English Discussion Skills as an L2 framework of phrases, around which students could conduct their discussions and add the L1 content of their ideas. Looking back reflectively at this intervention, students were afforded an opportunity to experience a strong form of deliberate, teacher-directed translanguaging (Jones, 2017). This particular hybrid discussion task demanded that students should not shift between English and Japanese arbitrarily, but do so in a controlled and intentional manner. Students clearly demonstrated the ability to cognitively engage with the task on multiple linguistic levels, that is simultaneously synthesizing their knowledge of the interactional Discussion Skills patterns in English (Hurling, 2012), while effectively rehearsing ideas in their L1, in interactions that resembled *strategic planning* and *rehearsal* (Ellis, 2005, 2009). As a consequence of this intervention, the monolingual English discussions started off with students visibly more confident and enthusiastic about the task they were to undertake. Furthermore, they were able to interact fluidly, with little or no hesitation when recalling and applying the target language phrases mid-discussion.

This limited intervention demonstrates that the effective implementation of pedagogical translanguaging practices is possible with relatively minor adjustments in the planning of regular classroom activities, while also producing tangible, practical outcomes in terms of target language output and fluency. At the time of writing, the author continues the reflective journaling process of documenting experiences of experimenting with pedagogical translanguaging practices in the

discussion class context, the findings of which will guide the development of a future research study.

Conclusion

This research brief set out to draw attention to the possibilities and advantages of conducting research into the application of pedagogical translanguaging in an English discussion course in a Japanese university setting. As the existing literature and preliminary research presented in this brief have demonstrated, there is both a clear need and opportunity for further empirical research in this area. There are still many questions to be resolved; therefore, the framing of relevant research questions may need further consideration. The design of surveys and other research instruments will also require careful attention in order to harvest the most relevant data. Most importantly, the development of effective pedagogical translanguaging activities to implement during this study will be paramount.

Finally, it is expected that this study will provide the best opportunity for both learners and teachers alike to have a collective translanguaging experience, while exploring the many tangible potential benefits that can be had from engaging in pedagogical translanguaging practices, not only in terms of increasing attainment levels in L2 acquisition but also on a sociolinguistic level, through reinforcing the identities of learners as *true* bilingual speakers.

References

Baker, C. (2001). *Foundations of bilingual education and bilingualism (3rd ed.)*. Clevendon: Multilingual Matters LTD.

Baker, C. (2011). *Foundations of bilingual education and bilingualism (5th ed.)*. Bristol, UK: Multilingual Matters.

Becker, A. L. (1988). 'Language in particular: a lecture' In D. Tannen & R. O. Freedle (Eds.). (1988). *Linguistics in context: Connecting observation an understanding: Lectures from the 1985 LSA/TESOL and NEH Institutes (Vol. 29)*. Praeger.

Canagarajah, S. (2011). Translanguaging in the classroom: Emerging issues for research and pedagogy. *Applied Linguistics Review*, *2*(2011), 1-28. https://doi.org/10.1515/9783110239331.1

Canagarajah, S. (2013). *Translingual practice: Global Englishes and cosmopolitan relations*. Routledge.

Canagarajah, S. (2018). Translingual practice as spatial repertoires: Expanding the paradigm beyond structuralist orientations. *Applied Linguistics*, *39*(1), 31-54. https://doi.org/10.1093/applin/amx041

Cenoz, J., & Gorter, D. (2017). Minority languages and sustainable translanguaging: Threat or opportunity? *Journal of Multilingual and Multicultural Development*, *38*(10), 901-912. https://doi.org/10.1080/01434632.2017.1284855

Cenoz, J., & Gorter, D. (2022). *Pedagogical translanguaging*. Cambridge: Cambridge University Press.

Cohen, L., Manion, L., & Morrison, K. R. B. (2007). *Research Methods in Education*. London: Routledge.

Cook, V. J. (1991). The poverty-of-the-stimulus argument and multicompetence. *Interlanguage Studies Bulletin (Utrecht)*, *7*(2), 103-117. https://doi.org/10.1177/026765839100700203

Cummins, J. (2007). Rethinking monolingual instructional strategies in multilingual classrooms. *Canadian Journal of Applied Linguistics/Revue Canadienne De Linguistique Appliquee, 10*(2),

221-240

Ellis, R. (Ed.). (2005). Planning and task performance in a second language. *Language Learning and Language Teaching*. https://doi.org/10.1075/lllt.11

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, *30*(4), 474-509. https://doi.org/10.1093/applin/amp042

García, O. (2009). *Bilingual Education in the 21st Century: A Global Perspective* (1st ed.). Oxford: Wiley-Blackwell.

García, O., & Lin, A. M. Y. (2017). Translanguaging in bilingual education. In O. García & A. M. Y. Lin (Eds.), *Bilingual and multilingual education (Encyclopedia of Language and Education,* Vol. 5), 117-130. Dordrecht, Netherlands: Springer International Publishing.

García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Basingstoke, UK: Palgrave Macmillan.

Hurling, S. (2012). Introduction to EDC. *New Directions in Teaching and Learning English Discussion, 1*(1), 1.2-1.10.

James, E. A., Milenkiewicz, M. T., & Bucknam, A. (2008). *Participatory action research for educational leadership: Using data-driven decision making to improve schools*. Thousand Oaks, California: SAGE Publications, Inc., https://dx.doi.org/10.4135/9781483329369

Jones, B. (2017). Translanguaging in Bilingual Schools in Wales. *Journal of Language, Identity & Education*, *16*(4), 199-215. https://doi.org/10.1080/15348458.2017.1328282

Kemmis, S., & McTaggart, R. (1992). *The Action Research Planner (3rd edition)*. Geelong, Vic.: Deakin University Press.

Lin, A. M. (2015). Conceptualising the potential role of L1 in CLIL. *Language, Culture and Curriculum*, *28*(1), 74-89. https://doi.org/10.1080/07908318.2014.1000926

Murphy, J. M. (2014). Reflective teaching: Principles and practices. In M. Celce-Murcia & D. M. Brinton (Eds.), Teaching English as a second or foreign language (4th ed.) (613-629). Boston, MA: Heinle Cengage.

Ooi, W. Z., & Abdul Aziz, A. (2021). Translanguaging pedagogy in the ESL classroom: A systematic review. *International Journal of Academic Research in Progressive Education and Development*, *10*(3). https://doi.org/10.6007/ijarped/v10-i3/10816

The philosophy of the FLER establishment. (2022). The Center for Foreign Language Education and Research (FLER). https://fler.rikkyo.ac.jp/message/index.html Accessed: 2022/6/15.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass, & C. Madden (Eds.), *Input in second language acquisition,* 235-253. Rowley, MA: Newbury House.

Turnbull, B. (2019). Translanguaging in the planning of academic and creative writing: A case of adult Japanese EFL learners. *Bilingual Research Journal*, *42*(2), 232-251. https://doi.org/10.1080/15235882.2019.1589603

Williams, C. (1994). Arfarniad o Ddulliau Dysgu ac Addysgu yng Nghyd destun Addysg Uwchradd Ddwyieithog [An evaluation of teaching and learning methods in the context of bilingual secondary education]. *Unpublished doctoral thesis*, University of Wales, Bangor. https://research.bangor.ac.uk /portal/files/20574535/null

【研究報告】

日本語の過去表現と中国語の助詞“了”の対応についての考察

白方直美

Abstract

The particle “了” (le), which expresses the completion of an action in the Chinese language, is frequently translated as “～た” (－ ta) in Japanese, which serves a similar purpose. For that reason, many learners of Chinese interpret “了” as indicating the past tense. However, the concept of “tense,” and by extension “past tense,” does not exist in the Chinese language to begin with. With that in mind, when translating past expressions in Japanese into Chinese, when should the particle “了” be used and when should that use be avoided ? In this paper, using “挪威的森林”(Nuowei de Senlin), the Chinese language translation of Haruki Murakami’s novel “ノルウェイの森” (Noruuei no Mori/Norwegian Wood) by Lin Shaohua, I verified situations in which past expressions in the Japanese language are translated into Chinese with use of the particle “了.”

Keywords: *過去形　助詞“了”　動作の完了　『ノルウェイの森』　中国語翻訳書*

はじめに

中国語の助詞“了”には、動詞の後に付く“了1”と文末に置く“了2”の二種類がある。前者“了1”は動態助詞やアスペクト助詞と呼ばれて「動作の完了」を表し、後者“了2”は語気助詞の一つで「変化」や「新状況の出現」、「肯定の語気」などを表す。例えば“了1”は“中午我吃了很多。(お昼はたくさん食べた。)”、“了2”は“他去中国了。(彼は中国に行った。)”のように、いずれも既に終った過去の動作を表すことができる。

すでに終った動作の表現を“了”を使って教える時、初級中国語では①（動詞に目的語がない場合）動詞＋“了”、②動詞＋“了”＋数量表現または修飾語＋目的語、③動詞＋目的語＋“了”の三パターンで教えるのが一般的である。実際の“了”の使用状況はかなり複雑だが、初級学習者が混乱しないように配慮され、多くの教科書でこうした教え方が用いられている[1]。

本稿では、村上春樹著の『ノルウェイの森』と、林少華が『ノルウェイの森』を中国語に訳した《挪威的森林》を対照し、すでに終った動作がどのように中国語に訳されているか、どのように“了”を用いて訳されているかを検証した。

1：中国語翻訳書《挪威的森林》にみられる過去の表現

先ずは林少華が中国語に翻訳した《挪威的森林》について、日本語の過去表現が中国語でどのように訳されているかをみた結果、表1に示す通り8パターンあった。

1　相原茂・陳淑梅・飯田敦子（2019）『日中いぶこみ交差点』第13課、第14課
清原文代・井手克子韓・艶玲・高恵敏（2012）『一歩一歩学漢語』下冊第1課
杉野元子・黄漢青（2013）『大学生のための初級中国語40回』第19課、第20課など

表 1
《挪威的森林》にみる過去の表現

	(形式)	日本語原文	中国語翻訳文
1	動詞＋助詞“了”	「どうしてこんなところに**来た**の？」第二章	“怎么到这儿**来了**？”
2	動詞＋助詞“着”	そんな一対の美しい瞳が長いあいだ僕の中を**のぞきこんでいた**。第一章	这对如此美丽动人的眸子久久地、定定地**注视着**我。
3	動詞＋助詞“过”	ちょうど僕がかつての僕自身が**立っていた**場所から確実に遠ざかりつつあるように。第一章	正如我逐渐远离自己一度**站过**的位置一样。
4	“是……的”構文	我々は**何かの目的があってここに来たわけではなかった**。第二章	我们也**不是**有什么目的才来这里**的**。
5	動詞単独	やれやれ、またドイツか、と僕は**思った**。第一章	罢了罢了，又是德国，我**想**。
6	動詞＋補語	直子は**立ちどまった**。僕も**立ちどまった**。第一章	直子**停住**脚，我也**停住**。
7	“一”＋動詞	飛行機が**着地を完了すると**禁煙のサインが消え、天井のスピーカーから小さな音でBGMが流れはじめた。第一章	飞机**一着陆**，禁烟显示牌倏然消失，天花板扩音器中低声流出背景音乐，
8	動詞＋“罢”	そして**食べ終ると**「ここの冬はいいですよ。この次は是非冬にいらっしゃい」と昨日と同じことを言って去っていった。第六章	**吃罢饭**，又重复了一句“这里的冬天不错哟，下次务必冬天里来看看”，这才离去。

2 《挪威的森林》にみられる助詞“了”を用いない過去の表現

2-1 動詞単独で過去事象を表現

中国語は時制の概念がなく、また動詞自体が変化や活用しない[2]ため、動詞の後ろに助詞を置いたり、動詞の前に副詞を置くことにより、動作の様々な様相を表現する。しかしこうした助詞や副詞がなく、動詞単独で過去事象を表現している例は非常に多い。特に直接話法や間接話法の前後の動詞「…と言った」や「…と尋ねた」、「…と答えた」は全て動詞単独で表現されている。例えば第一章「『大丈夫です。ありがとう』と僕は言った。(“不要紧的，谢谢。”我说。)」、「やがてドイツ人のスチュワーデスがやってきて、気分が悪いのかと英語で訊いた。(很快，一位德国空中小姐走来，用英语问我是不是不大舒服。)」、「『もちろん』と僕は答えた。(“那还用说。”我回答。)」の“我说”、“问我”、“我回答”のように動詞が単独で用いられている。劉月華らがその著書『現代中国語文法総覧』で、「直接話法または間接話法の前（または後）の動詞には“了1”を用いることができない」[3]と述べていることを証明している。

具体的な動作を表さない動詞も、動詞単独で過去事象を表現している。具体的な動作を表さない動詞には、「～は～である」という判断や認定を表す関係動詞の“是”、「知っている」、「感じる」のような知覚を表す“知道”や“觉得”、「好きだ」、「嫌だ」という感情を表す“喜欢”や“讨厌”、所有や存在を表す“有”、存在を表す“在”などがある。先ず“是”についてみてみると、表２の例「はじめて直子に会ったのは高校二年生の春だった。(第二章)」の「～は～だった」は、現在のことと同様に“…是…（～は～である)”の形で表現される（“第一次同直子见面，是高中二年级的春天。”)。「である」の過去「だった」を表すために“是”の後ろに“了”を置くことはできない[4]。“弟弟是大学生了。”のように文末に“了”を置く文があるが、これは「弟は大学生だった。」という意味ではなく、「弟は大学生になった。」という変化を表す。次に存在と所

2 郭春貴（2008）『中国語検定対策３級・４級文法編』p. 12
3 劉月華・潘文娱・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 298
4 劉月華・潘文娱・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 569

有を表す“有”の例をあげると、「永沢さんには大学に入ったときからつきあっているちゃんとした恋人がいた。（第三章）」の「いた」は、中国語では“从一入学开始，永泽就有一个地地道道的女朋友。”と動詞“有”のみで表されている。第六章に“婚后两年有了孩子。（結婚して二年後に子供が生まれて）”という表現があるが、この“有了”は「無から有になった」という変化を表すものであり、「子供がいた」のではなく、子供が「いない」状態から「できた」という変化を表現する。「いる」、「ある」という存在を表す動詞“在”も同様に、過去の「いた」、「あった」を表現する場合は“在了”ではなく“在”と表現する。

心理状態や知覚を表す状態動詞も、過去の事象を動詞単独で表現する。例えば第一章「もちろん直子は知っていたのだ。（直子当然知道，）」、第二章「僕はとても悪いことをしてしまったような気がした。（我觉得自己好像做了件十分对不起人的事。）の“知道（知っていた）”、“觉得（気がした）”など知覚を表す動詞や、第三章「冬が深まるにつれて彼女の目は前にも増して透明に感じられるようになった（随着冬日的延伸，我感到直子的眼睛比以前更加透明了）。」の“感到（感じられるようになった）”も動詞単独で過去の事象を表現している。また第三章「僕自身は知らない女の子と寝るのはそれほど好きではなかった。（我自己其实不大喜欢同萍水相逢的女孩同窗共衾。）」の“不大喜欢（それほど好きではなかった）”、第二章「それに父親が山登りが好きだったせいで、小さい頃から日曜日になると山登りをしてたの。（而且，由于父亲喜爱登山，我从小每到星期天就往山上爬。）」の“喜爱（好きだった）”など感情を表す動詞も単独で過去事象を表現する。

まとめると、①直接話法や間接話法の前後の動詞、②知覚や感情を表す状態動詞、③判断や存在を表す関係動詞は、動詞の後ろに完了を表す助詞“了”を置かず、動詞単独で過去事象を表現する。

表2

《挪威的森林》において動詞単独で過去事象を表現する例

		日本語原文	中国語翻訳文
動作動詞	直接話法、間接話法の前後	「大丈夫です。ありがとう」と**僕は言った**。第一章	“不要紧的，谢谢。”**我说**。
		やがてドイツ人のスチュワーデスがやってきて、気分が悪いのかと英語で**訊いた**。第一章	很快，一位德国空中小姐走来，用英语**问我**是不是不大舒服。
状態動詞	知覚	もちろん直子は**知っていたのだ**。第一章	直子当然**知道**，
		僕はとても悪いことをしてしまったような**気がした**。第二章	我**觉得**自己好像做了件十分对不起人的事。
	心理	僕自身は知らない女の子と寝るのは**それほど好きではなかった**。第三章	我自己其实**不大喜欢**同萍水相逢的女孩同窗共衾。
		それに父親が山登りが**好きだった**せいで、小さい頃から日曜日になると山登りをしてたの。第二章	而且，由于父亲**喜爱**登山，我从小每到星期天就往山上爬。
関係動詞	“是”（判断）	国旗を掲揚するのは東棟（僕の入っている寮だ）の寮長の役目**だった**。第二章	升国旗**是**东楼（我所住的楼）楼长的任务。
		はじめて直子に会ったのは高校二年生の春**だった**。第二章	第一次同直子见面，**是**高中二年级的春天。
	“有”（存在）	永沢さんには大学に入ったときからつきあっているちゃんとした恋人が**いた**。第三章	从一入学开始，永泽就**有**一个地地道道的女朋友。
		東京について寮に入り新しい生活を始めたとき、僕のやるべきことは**ひとつしかなかった**。第二章	到得东京，住进寄宿舍开始新生活时，我要做的**仅有一件事**，
	“在”（存在）	彼が上段で僕が下段**だった**。第二章	结果他睡上床，我**在**下床。
		彼女が先に立ち、僕がその少し後ろを歩いた。第三章	**她在前边**，我离开一点跟在后头。

2-2　動詞の後ろに補語を置いて過去事象を表現

“了”を使わないで過去事象を表現しているものの中で、動詞単独に次いで、動詞の後ろに補語を置いて過去を表現しているものが多かった。

動詞の後ろに結果補語または方向補語がある場合、“了”は省くことができる[5]。また様態補語はすべて動作や状態の結果または程度を表すもので[6]、ある動作がどのような様態でなされたか、または、どのような状態に達しているかについて補足説明するときに用いる[7]、すなわち様態補語で表現される事自体にすでに終っている意味が含まれていることがあるために、“了”を使わなくても過去事象の説明ができる[8]。第三章には「坂を上り、川を渡り、線路を越え、どこまでも歩きつづけた（上坡，过河，穿铁道口，只管走个没完）。」のような“个”に導かれる様態補語もみられた。この文は補語が否定形式なので「止まらない」意味を表し[9]、「歩きつづけた」ことを表現している。可能補語は「できる」「できない」を表現するものである。例えば“找不到”は「（捜しても）見つからない」という意味を表すが、後ろに“了”を置いて“找不到了”にすると、「それは見つけられなかった」ではなく「見つからなくなった」という変化を表す[10]。表３の例は“触不到”という可能補語の否定形を用いて「触れなかった」を表現している。「触れなかった」を中国語に訳せば“没触到”となるが、この文の情景では、「手を伸ばして光の粒子をつかもうとしたのに（どうしても）それに触る事ができなかった」ことを表現したいために、可能補語の否定形が使われていると考えられる。

表３

《挪威的森林》において動詞の後ろに補語を置いて過去事象を表現する例

動詞＋補語	日本語原文	中国語翻訳文
動詞＋結果補語	直子はポケットから左手を出して僕の手を**握った**。第一章	直子从衣袋里抽出左手**握住**我的手。
	電灯がうぶ毛をきれいな黄金色に**染めた**。第三章	电灯光把她细细的汗毛**染成**美丽的金黄色。
動詞＋方向補語	僕はときどき空中に漂う光の粒子に向けて**手を伸ばしてみた**が、その指先には何も触れなかった。第三章	我时不时向空中漂浮的光粒子**伸出手去**，但指尖什么也触不到。
	彼女はそのまま何も言わずに先に立って**歩きはじめた**。第一章	她便没再开口，开始在我前边**走起来**。
動詞＋様態補語	「申しわけないけれど僕の方は**かなりくたくただよ**」第二章	“这么说你别见怪，我可是**累得够呛**。
	いつもキズキが一座の中心にいたし、彼はそういうのが**上手かった**。第二章	木月总是节目的中心，而他又**干得的确得心应手**。
	坂を上り、川を渡り、線路を越え、**どこまでも歩きつづけた**。第三章	上坡，过河，穿铁道口，只管**走个没完**。
動詞＋可能補語	僕はときどき空中に漂う光の粒子に向けて手を伸ばしてみたが、その指先には何も**触れなかった**。第三章	我时不时向空中漂浮的光粒子伸出手去，但指尖什么也**触不到**。
	「すごく。だから**我慢できなかった**んです」第十一章	“好极了，所以才**忍不住**。”

5　劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 291
6　劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 495
7　瀬戸口律子（2006）『完全マスター中国語の文法』p. 146
8　例えば“今天我们玩儿得很高兴（今日は楽しく遊んだ）。”郭春貴（2014）『誤用から学ぶ中国語』p. 129
9　劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 504
10　韓志剛・張文賢（2004）《HSK 语法指要与训练》p. 127

2-3　助詞“了”を用いないで過去事象を表現（動詞単独、動詞＋補語を除く）

続いて助詞“了”を用いず、動詞単独と動詞の後ろに補語を置いて過去事象を表現する以外の例として、表４に示すような形式がみられた。

表4

《挪威的森林》にみる過去の表現（助詞“了”、動詞単独、動詞＋補語を除く）

中国語表現形式	日本語原文	中国語翻訳文
“一”＋動詞	**飛行機が着地を完了すると**禁煙のサインが消え、天井のスピーカーから小さな音でBGMが流れはじめた。第一章	**飞机一着陆**，禁烟显示牌倏然消失，天花板扩音器中低声流出背景音乐，
	直子は立ちどまって**にっこりと笑い**、僕の腕をそっとつかんだ。第一章	直子停下，**动情地一笑**，轻轻抓住我的胳臂，
“是…的”構文	我々は**何かの目的があってここに来たわけではなかった**。第二章	我们也**不是有什么目的才来这里的**。
	そのときの僕は**それを言葉としてではなく、ひとつの空気のかたまりとして身のうちに感じたのだ**。第二章	但当时的我**并不是将其作为语言，而是作为一团薄雾样的东西来用整个身心感受的**。
動詞＋助詞“着”	そんな一対の美しい瞳が長いあいだ**僕の中をのぞきこんでいた**。第一章	这对如此美丽动人的眸子久久地、定定地**注视着我**。
	と言いかけて直子は口をつぐみ、そのまま**歩きつづけた**。第一章	说到这里，直子蓦地合拢嘴唇，**继续往前走着**。
動詞＋助詞“过”	ちょうど僕がかつての僕自身が**立っていた場所**から確実に遠ざかりつつあるように。第一章	正如我逐渐远离自己一度**站过的位置**一样。
	私が**存在し、こうしてあなたのとなりにいたこと**をずっと覚えていてくれる？」第一章	记住我这样**活过、这样在你身边呆过**。可能一直记住？”
“来着”	彼女はそのとき**何の話をしていたんだっけ**？第一章	她那时究竟**说什么来着**？

先ず動詞の前に副詞“一”を置く「“一”＋動詞」は、ある短い動作から何らかの結果や結論が得られることを表す[11]。例えば“他一听，果然不高兴了。(彼はそれを聞くと案の定不機嫌になった。)[12]”のように使われる。この例からもわかるように、この文型は“一”の後ろの動作が完了したうえで次の動作がおこることを表しているので、動詞の後ろに“了”を置く必要はない。

また中国語は、“是…的”構文を用いてすでに起ったことを表現するが、“是…的”構文は、ただ単に過去の事象を描写するのではなく、あることが起ったという事実があり、それがいつ、どこで、どんなふうに起ったか、すなわち状況語を尋ねたり説明する時に用い、“是…的”構文では“了”を用いることはできない。

中国語には動詞の後ろに置く動態助詞が三つあり、完了を表す“了”以外に、動作の持続を表す“着”と過去の経験を表す“过”がある。“着”は時制の制限を受けないために、命令文のような未来のことにも、現在の状態の持続にも、過去の状態の持続にも用いられる。このため、表４の例１「僕の中をのぞきこんでいた（注视着我）」、例２「歩きつづけた（往前走着）」のような動作が続いていたことを表すには「動詞＋“着”」が使われる。さらに“着”の後ろには他のアスペクト助詞を置くことができないため[13]、“着了”の形はとれない。動詞の後ろの“过”は「かつてあることをしたことがある」ことを表し、過去についてしか用いられない[14]。このため「動詞＋“过”」の形は必ず過去のことであり、今はもうそうではないことを表す。表４の

11 『中国語文法用例辞典』東方書店（2003）
12 『中日辞典』講談社（2002）
13 劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 320
14 劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 320

例１「かつて僕自身が立っていた場所」は“自己一度站过的位置”と訳されているが、“了”が使われていないのは、「かつて立っていたことがある場所」という経験を表しているからである。例２の「私が存在し、こうしてあなたのとなりにいたこと」は“我这样活过、这样在你身边呆过”と“过”を用いて訳されているが、これも「かつて存在していた」、「かつてとなりにいた」という経験を表しているので、完了の“了”ではなく“过”を用いて訳している。“来着”は文末に置いて過去に起きたことを回想する気持を表す助詞であるが、口語にしかみられない[15]。

3 《挪威的森林》にみる過去の表現“了“を使うもの

ここでは、助詞“了”を用いて過去を表現しているものを詳しくみていく。分析の結果、大きく分けて表5に示す五つのパターンがあることがわかった。

表5

《挪威的森林》にみる過去の表現“了“を使うもの

	日本語原文	中国語翻訳文
① 動詞＋“了”（目的語なし）	「どうしてこんなところに**来た**の？」第二章 三日間、僕は**我慢した**。第二章	“怎么到这儿**来了**？” 头三天，我都**忍了**。
② 動詞＋“了”＋単独の目的語	緑は行くところがあるからと言って四ツ谷駅から**電車に乗った**。第四章 僕と直子は四ツ谷駅で**降りて**、線路わきの土手を市ヶ谷の方に向けて歩いていた。第二章	绿子说她有个地方要去，从四谷站**上了电车**。 我和直子在四谷站**下了电车**，沿铁路边上的土堰往市谷方面走去。
③ 動詞＋“了”＋数量/修飾語（＋目的語）	彼女は「ディア・ハート」のメロディーをもう一度**何小節か軽く弾いてからワインをすすった**。第十一章 ただなんとなくそれ以来**玉撞きをする機会がなかったんです**。それだけのことですよ。」第八章	她又轻轻**弹了几小节《宝贝儿》的旋律，呷了口葡萄酒**。 “只是不知为什么，从那以后就**失去了打桌球的机会**一就这么回事。”
④ 動詞＋目的語＋“了”	それでも僕らはしばらく四人で親しく話をしていたのだが、待ちあわせの相手が来ると二人は**そちらへ行ってしまった**。第四章	但我们四人还是亲热地聊了一会，约会的男朋友一来，两个女孩便**去那边了**。
⑤ 動詞以外＋“了”	その他には**どんな物音もなかった**。第一章 俺の人生っていったい**何だったんだ**？この女のおもりをするだけのことなのかって。第一章	此外便**万籁俱寂了**。 你会想：这辈子到底是**怎么了**，只落得给这女人当护身符不成?

3-1 動詞の後ろに目的語があるもの

動詞の後ろに目的語がなければ、「動詞＋“了”」で動作が完了したことを表現できるが、動詞に目的語がある場合は、いくつかある程度規則的なパターンがある。「はじめに」で述べたように、目的語の前に数量表現や修飾語がある場合には「動詞＋“了”＋数量表現または修飾語＋目的語」、目的語の前に数量表現も修飾語もない場合は「動詞＋目的語＋“了”」になる。

3-1-1 「動詞＋“了”＋数量表現＋目的語」

目的語の前にある数量表現を、その性質別に整理したものを表６に示す。

15 劉月華・潘文娯・胡斡（1996）『現代中国語文法総覧』（相原茂監訳）p. 322

表6

目的語の前に数量表現がある例

	日本語原文	中国語翻訳文
① 動詞＋“了”＋数量詞＋目的語	花屋が一軒店を開けていたので、**僕はそこで水仙の花を何本か買った**。第四章	花店倒有一家开了门，**我买了几枝水仙花。**
② 動詞＋“了”＋動量詞＋目的語	「そうかもしれないな」と彼は言って二、**三度小さく肯いた**。第九章	“或许。”他轻轻**点了两三下头，**
③ 動詞＋“了”＋時間量＋目的語	緑は**しばらくまっすぐ立ちのぼる煙草の煙を眺めていた**。第七章	绿子**望了半天烟头上笔直升起的烟**。

《挪威的森林》では、目的語の前に数量表現がある過去事象は、全て動詞の後ろに“了”を置いて表現されている。

3-1-2 「動詞＋“了”＋修飾語（数量以外）＋名詞の目的語」

次に、《挪威的森林》にみられる数量を含まない修飾語がある目的語の例をみていく。

表7

「動詞＋“了”＋修飾語（数量以外）＋名詞の目的語」

	日本語原文	中国語翻訳文
① 動詞＋“了”＋指示代詞＋目的語	**そういう映画**をどこかで見たんじゃないかしら。第四章 「君に会ったおかげで少し**この世界**に馴染んだような気がするな」と僕は言った。第七章	说不定她在哪里看了**这类电影。** “见到你，我觉得多少适应了**这个世界**。”我说。
② 動詞＋“了”＋疑問代詞＋目的語	だから僕には**何がどうなっているのか**さっぱりわからなかった。第九章	根本搞不清发生了**什么事**。
③ 動詞＋“了”＋人称代詞＋“的”＋目的語	伊東は泊っていけよと言ったが、僕はちょっと用事があるからと言って断り、ウィスキーの礼を言って九時前に**彼のアパート**を出た。第十章	伊东叫我住下，我说还有点事，谢过他招待的威士忌，九点前离开了**他的住所**。
④ 動詞＋“了”＋名詞＋“的”＋目的語	僕はそのとき本当に**眠りの中枢**に達していたのだ。第九章	当时确实达到了**睡眠状态的极限**，
⑤ 動詞＋“了”＋形容詞＋（“的”）＋目的語	名前を呼ばれても僕が黙っていると、教室の中に**居心地のわるい空気**が流れた。第四章	当点名我也不应时，教室里便出现了**尴尬的气氛**。
動詞＋“了”＋形容詞フレーズ＋目的語	**いろんなところ**に行って、**いろんな話**をして。第六章	去了**很多地方**，说了**很多话**。
⑥ 動詞＋“了”＋動詞フレーズ＋“的”＋目的語	彼女の声を聞いて、僕は自分が**何か間違ったこと**を口にしたらしいなと思った。第一章 僕は礼を言って、男からもらった五千円冊で**東京までの切符**を買った。第十一章	听她这么说，我察觉自己大概说了**不该说的话**。 用渔夫送给的五千钞票买了**去东京的车票**。
⑦ 動詞＋“了”＋主述フレーズ＋“的”＋目的語	そして僕は緑に、**彼女のいないあいだに起った出来事**をひとつひとつ報告した。第七章 そして五十一曲目に**いつものバッハのフーガ**を弾いた。第十一章	我向绿子一五一十汇报了**她不在时发生的事：** 玲子第五十一首弹了**她经常弹的巴赫赋格曲**。
⑧ 動詞＋“了”＋前置詞フレーズ＋“的”＋目的語	私は**私に対するあなたの好意を感じる**し、それを嬉しく思うし、その気持を正直にあなたに伝えているだけです。第五章	我感受出了**你对我的好意**，并为此感到高兴 一只是想把这种心情如实地告诉你。

⑨ 動詞＋“了”＋固定フレーズ（成語）＋“的”＋目的語	台所の日だまりの中でTシャツ一枚になってドイツ語の文法表を方端から暗記していると、何だかふと**不思議な気持になった**。第六章	我拣了一块暖洋洋的向阳处，这时我不由产生了**不可思议的感觉**：

表７より、動詞の後ろに名詞の目的語があり、その目的語に修飾語がある場合、修飾語には様々なタイプがある事がわかる。ほぼ全ての定語[16]（指示代詞、疑問代詞、人称代詞、名詞、形容詞（フレーズ）、動詞（フレーズ）、主述フレーズ、前置詞フレーズ）について、「動詞＋“了”＋修飾語＋目的語」が成立している。

3-1-3 「動詞＋“了”＋動詞フレーズまたは主述フレーズの目的語」

思考や発言、知覚を表す動詞は主述フレーズを目的語にとる事が出来るが[17]、目的語が動詞や動詞フレーズ、主述フレーズなどの述詞性のことばである場合、述語動詞の後にはふつう“了1”を用いることはできない[18]。表８の①“认为（～と考える）”や②“以为（～と考える）”の例をみても“认为了”、“以为了”となっていないことがわかる。ただし、“忘（～を忘れる）”については、肯定文では名詞、節を客語にともなうとき、必ず“了”を“忘”の後ろに付けて[19]“忘了”の形をとる（表８③）。また述語が“进行”、“作”などの動詞であれば、その目的語が述詞性のものであっても、後に“了1”を用いることができる[20]（表８④）。表８の⑤については、目的語が“种菜”、“做意大利面条”と動詞フレーズであるが、動詞の後に“上”と“会”という結果補語があり、さらに目的語は日本語で「野菜作り」、「おいしいスパゲティーの作り方」と名詞句になっており、中国語でも名詞同様の働きをしていると思われる。また表８の⑥動詞“谈”（発言）と⑦動詞“目睹”（知覚）も、後ろに助詞“了”を置いて主述フレーズの目的語をとっている[21]。

表８

目的語が主述フレーズ、動詞フレーズの場合[22]

①“认为”＋動詞フレーズ	今まで殆んど誰にも話したことはないのだけれど、**あなたには話しておいた方がいいと思う**から話すのだと彼女は言った。第六章	她说，这话还几乎没向任何人讲过，但**认为还是向我讲了为好**。
①“以为”＋主述フレーズ	**寝ちゃったのかしらと思ってね**。第六章	本来我**以为她睡着了**呢，
②“忘”＋“了”＋主述フレーズ	あのレストランに電話しようかと思ったんだけど**店の名前も覚えていない**し、あなたの家の電話だって知らないし。第四章	本想给饭店打个电话，但**忘了那店叫什么名**，又不晓得你家的电话，。
③“进行”＋“了”＋動詞フレーズ	彼の家は四国の田舎の旧家でね、両親が**私のことを徹底的に調べて**、入院歴が二回あることがわかっちゃったのよ。第六章	他家是四国乡下有些来历的家族，父母对我**进行了彻底调查**，知道我住过两次院，
④ 動詞＋結果補語＋“了”＋動詞フレーズ	私も**野菜作りは大好きになりました**。第５章 今イタリア料理店でアルバイトをしていて、コックから**おいしいスパゲティーの作り方を習った**。第十章	我也**爱上了种菜**， 眼下我在一家意大利餐馆打工，从厨师那里**学会了做意大利面条**，

16 定語は連体修飾語のこと。
17 守屋宏則（1997）『やさしくくわしい中国語文法の基礎』p. 21
18 劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 299
19 『中国語文法用例辞典』東方書店（2003）
20 劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 299 の例
21 他にも“他目睹了爷爷头发变白。”などの例がある。微文语录 https://www.wwyulu.com/read/11205.html
22 主述フレーズを目的語に従える動詞については、守屋宏則（1997）『やさしくくわしい中国語文法の基礎』p. 21 に従った。

⑤“谈”＋“了”＋主述フレーズ	そして**私と二人でどこに住もうだの、どんなことしようだのといったようなことを話したの**。第十一章	接下去**谈了我俩住在哪里、做什么工作等等**。
⑦“目睹”＋“了”＋主述フレーズ	僕は**永沢さんが酔払ってある女の子に対しておそろしく意地わるくあたるのを目にして**以来、この男にだけは何があっても心を許すまいと決心したのだ。第三章	自从我**目睹了永泽酩酊大醉后想方设法捉弄女孩子**以后，我就决意万万不可向他交心。

3-1-4 「動詞＋“了”＋（修飾語なし）目的語」

これまでは助詞“了”を用いた過去事象の表現のうち、“了”の後ろに数量表現や修飾語がある目的語をもつ例をみてきた。次に、目的語の前に数量表現も修飾語もない例をみていく。「動詞＋“了”＋目的語単独」の場合、ふつう文が終らないで、「～したら、～」、「～してから～」のように文が続く。劉月華らも、「“了1”を含み、さらに目的語を含む文では、ふつう目的語には定語が必要とされる」が、目的語の前に定語がなくてもかまわない場合として「二つ（またはそれ以上）の、緊密につながっている動詞フレーズ・分句または対句中に目的語が現れる場合」を挙げている[23]。

《挪威的森林》にみられる目的語が単独で、さらに文が続く例の一部を表9に示す。こうした例は非常に多くみられた。

表9

「動詞＋“了”＋目的語，～」

日本語原文	中国語翻訳文
僕と直子は四ツ谷駅で**降りて、線路わきの土手を**市ヶ谷の方に向けて歩いていた。第二章	我和直子在四谷站**下了电车，沿铁路边上的土堰**往市谷方向走去。
僕は新宿で簡単な**昼食を食べ、それから**電話ボックスに入って小林緑に電話をかけてみた。第四章	我在新宿简单**吃了早餐，进电话亭**给小林绿子打电话。
僕は切符売り場で**切符を買い、それから**近くの書店に入って地図を買い、待合室のベンチに座って「阿美寮」の正確な位置を調べてみた。第六章	我在售票处**买了车票，然后走入**近处一家书店，买张地图，坐在候车亭的凳子上查找“阿美寮”的准确位置。
十年も毎日毎日やってるからさ、**やり始めると、む、無意識に全部やっちゃうんだ**。第二章	我是十年如一日做过来的。**一旦开了头，就、就下意识地一做到底**。

3-1-5 「動詞＋“了”＋（修飾語なし）目的語。」（文が終る）次に、単独の目的語で文が終る例をみていく。これは「はじめに」で述べた三パターンにはあてはまらないので、初級学習者には教えないことが多いが、実際こうした例は非常にたくさんあった。《挪威的森林》でみられた「動詞＋“了”＋単独の目的語」で文が終っている例を、述語や目的語の性質に基づいて分類したものの例を表10に示す。

表10

「動詞＋“了”＋目的語」

目的語	日本語原文	中国語翻訳文
①動詞＋結果補語＋“了”＋場所 / 時間	そして穴の中には暗黒が一世の中のあらゆる種類の暗黒を煮つめたような暗黒が一つまっている。第一章	井筒非常之黑，黑得如同把世间所有种类的黑一古脑儿**煮在了里边**。
	私たちは支払うべきときに代価を支払わなかったから、そのつけが**今まわってきてるのよ**。第六章	我们在应该支付代价的时候没有支付，那笔账便**转到了今天**。

23 劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 293

② 動詞＋“了”＋動詞	それは一瞬**胸がつまってしまう**くらいあたたかくて素敵な仕草だった。第一章 そんな手紙を読みかえしてみると、僕自身が**慰められた**。第十章	一瞬间，我觉得一股暖流穿过全身，心脏都好像**停止了跳动**。 写罢反复阅读之间，我本身竟也**得到了慰藉，**
③ 動詞“成”＋“了”＋名詞	そして我々は**友だちになった**。第三章 もう**病気みたいなもの**よね、これ。第四章	我们果真**成了朋友**。 简直**成了病态**。
④ 離合詞	降りましょうよと直子が言って、我々は**電車を降りた**。第二章 病院に入って少ししてから小指は動くようになったから、音大に復学してなんとか**卒業することはできた**わよ。第六章	直子说声下车吧，我们就**下了车**。 住院不久，小手指可以动了，便去音乐大学复学，总算**毕了业**。
⑤ 動詞＋“了”＋疑問詞（不定を表す）	**何を言っているの**か聞きとれないということもあった。第二章 そして**どんなだったか**あとで教えてほしいの。第七章	也有时听不清楚她**说了什么**。 事后把结果告诉我，例如都**做了哪些。**”
⑥ 動詞＋“了”＋固有名詞	そのようにして**僕と直子は出会った**のだ。第二章 それから彼女は「**ノ―ホエア・マン」を弾き、「ジュリア」を弾いた**。第六章 まあそれで東京に戻ってきて二、三日ぶらぶらして、それから今度は一人で気楽に旅行しようと思って**青森に行ったの**。第九章	就这样我**遇到了直子**。 接着，她**弹了《没有归宿的人》，弹了《朱莉娅》。** 返回东京后，游逛了两三天，这回想一个人无拘无束地旅行一趟，就**去了青森。**
⑦ 動詞＋“了”＋人称代詞	それで私、**彼のことがすごく好きになったの**。第六章 レズビアンのようなことをしかけられたんだ、それで**打ったんだ**って。第六章	这一来，我不折不扣地**喜欢上了他**。 说那女孩动手动脚要搞什么同性恋那样的鬼名堂，所以才**打了她**。
⑧ 動詞＋“了”＋数量詞	そのようにして僕は十八から**十九になった**。第三章	这么着，我从十八岁**进入了十九岁。**
⑨ 動詞＋“了”＋複合方向補語	十五分ほど経ったところで教室のドアが開いて緑が入ってきた。第四章 レイコさんが台所からまねき猫の形をした貯金箱を持ってきて、直子が財布から百円玉を出してそこに入れた。第六章	大约过了十五分钟，教室的门开了，绿子**闪了进来**。 玲子从厨房拿出一个招手猫形的贮币盒，直子从钱包里找出一枚百元硬币，**投了进去**。
⑩ 動詞＋“了”＋一般名詞	約束どおり僕が**ゲーム代を払った**。第二章 永沢さんは食事に行くところだったので、僕も一緒に食堂に行って**夕食を食べた**。第四章	我按事先讲好的**付了费用**。 永泽正要去食堂吃饭，我也一起跟去**吃了晚饭**。

表 10 の④離合詞の場合、動詞部分の後ろに“了 1”が置かれる事が多い。劉らは「動詞と目的語が比較的固定した組み合わせの動目フレーズであり、ふつう動詞の前にさらに時間に関わる状語があれば目的語の前に定語がなくてもかまわない」と述べている[24]。さらに、目的語が固有名詞の場合（表 10 の⑥）も目的語の前に定語がなくてもかまわないとある。[25]

3-2　文末（または節の終り）の“了 2”

「はじめに」で述べたように、動詞の目的語に数量表現や修飾語がない場合、動詞の後ろではなく、ふつ

24　劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 293
25　劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 294

うは文末に“了2”を置くことにより文が終結する[26]。《挪威的森林》でみられた文末や節の終りに置かれた“了2”の例を表11に示す。

表11

「動詞＋（目的語＋）“了2”。」

目的語	日本語原文	中国語翻訳文
① 動詞＋“了”目的語なし	僕が質問すると、彼はああほんとうだよ、それ、と言った。第三章	我一问，他就痛快**承认了**，
② 動詞＋結果補語＋場所＋“了”	そうするとこのへんの人は言うの、あれは野井戸に落っこちたんだって」第一章	于是这一带的人就说：**准保掉进那荒草地的井里了**。”
③ 動詞＋結果補語“成”＋目的語＋“了”	「大人になったからだよ」第十章	放下伞不**淋成落汤鸡了**？”
動詞“成”＋目的語＋“了”	傘ささないとずぶ濡れになっちゃうよ」第十章	“**成大人了嘛**。”
④ 離合詞＋“了”	「病院を出てしばらくしてから主人と知りあって結婚したの。第六章	出院后不久和丈夫**结婚了**。
	「そうそう、このあいだ試験の**発表あったよ**。受かってたよ」と永沢さんが言った。第八章	“对了对了，上次那场考试**发榜了**，中了。”永泽说。
⑤ 疑問詞＋“了”	俺の人生っていったい**何だったんだ**？　第一章	你会想：这辈子到底是**怎么了**，
	「それでそのクラブはどうしたの？」第七章	“那民歌小组**怎么办了**？”
⑥ 動詞＋固有名詞＋“了”	「スペイン語始めたんですか？」第八章	开始**学西班牙语了**？”
	「いいよ。それで奈良に行ったんだ」 第九章	“没关系。所以就**去奈良了**？”
⑦ 動詞＋人称代詞＋“了”	「ねえ、このあいだの日曜日あなた**私にキスした**でしょう」と緑は言った。第七章	“咦，上次那个星期日你**吻我了吧**？”绿子说，
	「病院よ。お父さんが入院していて、今日いちにち私がつきそってなくちゃいけないの。**私の番なの**」第七章	“医院呗。我爸爸住院，今天该我陪伴一天，**轮到我了**。”
⑧ 数量詞＋“了”	寮に戻ったのは**十一時半だった**。第八章	回到宿舍，已经**十一点半了**。
	俺はもう**二十歳になった**んだよ。第十章	我已经**二十岁了**！
⑨ 動詞＋複合方向補語＋“了”	しかし二日めの朝になると彼はむっくりと起きあがり、何事もなかったように**体操を始めた**。第三章	但第二天清早，他居然“咕噜”一声翻身下床，若无其事地**做起广播体操来了。**
	「どうしてそんなこと急に**思い出した**の？」第六章	“怎么突然**想起这个来了**？”
⑩ 動詞＋一般名詞＋“了”	「**退寮した**」と寮長は言った。第四章	“**退宿舍了**。”主任说。
	「あなたもだんだん**世の中のしくみがわかってきた**みたいじゃない」第六章	“你也慢慢开窍，**懂得社会结构了嘛！**”
⑪ 連動文	レイコさんも**煙草をくわえたまま笑った**。第六章	玲子也**叼着烟笑了**：
	夕方近くになると直子はお母さんにもう帰っていいわよ、あとは大丈夫だからって言って、それで**お母さんはタクシーを呼んでもらって帰っていったの**。第十一章	傍晚时，直子对她母亲说往下不要紧了，请母亲回去。她母亲**叫一辆出租车回去了**。

26 沈国威监修　氷野善寛・小嶋美由紀・海暁芳・紅粉芳恵・阿部慎太郎（2020）『中国語でコミュニケーション』教授用資料 pp.80-81 に「もし、目的語に数量詞や修飾語がない場合は、文末に更に“了”を加えると文が終結します（例：我昨天买了衣服了）。これは、文末の“了”には文を終結させる機能があるからです。」とある。また楊寄洲（2004）《汉语教程》第二册上の第32課にも「如果宾语前没有数量词或其他宾语时，句末要有语气助词“了”才能成句。」と、同様の記述がある。

⑫ 固定フレーズ（成語）＋“了”	そして我々は**別れた**。第二章 「でもそれ以来、誰も俺に対して**何も言えなくなったよ**。第三章	随后我们就**分道扬镳了**。 “不过，从那以来，谁对我都**无可挑剔了**。
⑬ 動詞＋動量詞＋“了”	**よく洗いこまれた**ものらしく、ずいぶん感じよく色が褪せていた。第二章 **通して読むのは三度め**だが読みかえせば読みかえすほど面白いと感じる部分がふえてくると僕は答えた。第三章	看上去**洗过好多遍了**，颜色褪得恰到好处。 我答已经**通读三遍了**，读的次数越是多，越觉得有趣的部分层出不穷。
⑭時間量＋“没”＋動詞＋“了”	直子と会ったのは殆んど**一年ぶり**だった。第二章 僕は**山のぼりなんてしばらくしていない**せいで息が切れた。第六章 「だって私、**鍋ものなんて何年も何年も食べてない**んだもの。第十一章	我差不多**有一年没见到直子了**。 倒是我因**好久没跟山打交道了**，不免气喘吁吁。 “我有**好些年好些年没吃火锅了**，
⑮ 動詞＋様態補語＋“了”	僕は**歩きつかれて**いささかぐったりとしていたし、彼女はテーブルの上に両手を置いてまた何かを考え込んでいた。第二章 だから自然に**足腰が丈夫になっちゃった**の」第二章	我**走得累了**，有点打不起精神，她两手放在桌面上沉思什么。 所以，腿脚就自然而然**变得结实了**。”
⑯ 助動詞＋動詞＋“了”	担当医は私がそろそろ**外部の人と接触を持ち始める時期だ**と言います。第五章 わりに大事なコンクールがあって、私ずうっとそのための練習してたんだけど、突然左の小指が**動かなくなっちゃった**の。第六章 **またピアノ弾けるんだわ**って思ってね。第六章	主治医生说我现阶段已经**可以慢慢开始同外界的人接触了**。 有个比较重要的音乐会，我为此练习了很长时间。不料小指突然**不会动了**， 毕竟又**能弹钢琴了**。
⑰ 動詞＋可能補語＋“了”	その男が本当に医者なのかどうか僕はますます**わからなくなってしまった**。第六章 「すみません。**我慢できなかった**んです」と僕は言った。第十一章	于是我愈发**弄不清**他是否真是医生**了**。 “对不起，我**忍不住了**。”我说。

表 11 のうち、⑧数量詞と⑯助動詞がある文、⑰可能補語の後ろの“了”については、動作が終ったことではなく、変化を表すと考えられる[27]。また本稿 2-2 で、様態補語はある動作がどのような様態でなされたかを表すので、“了”がなくても過去の事象を説明できると述べたが、劉月華らは、様態補語が述詞性の場合、後ろに“了2”を置くことができるとし[28]、特に動作が原因になって至った結果の状態を表す様態補語には“了”が使われることがある。⑪連動文では前の動詞の後ろに“了”は置けない[29]。⑭「時量＋“没”＋動詞＋“了”」は、ある状況が発話までにどのくらいの時間持続したかを表す[30]。すなわち発話までに、提示された期間その動作や状況が起らなかったことを表している。

27 韓志剛・張文賢（2004）《HSK 语法指要与训练》p. 127 に基づく。
28 劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 309
29 韓志剛・張文賢（2004）《HSK 语法指要与训练》p. 124
30 劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 304

3-3 “了1”を使うか“了2”を使うのか

表10と表11で示した“了1”を用いた表現と“了2”を用いた表現の数を比較したものを表12に示す。

表12

述語や目的語に基づいた《挪威的森林》にみられる“了”の表現例の数の比較[31]

	表現形式	“了1”	表現形式	“了2”
			動詞＋“了”（目的語なし）[32]	15
①	動詞＋結果補語＋“了”＋場所 / 時間	6	動詞＋結果補語＋場所＋“了”	11
②	動詞＋“了”＋動詞（句）	2		
③	動詞＋結果補語“成”＋“了”＋目的語 動詞“成”＋目的語＋“了”“了” 動詞＋結果補語“成”＋目的語＋“了”	3	動詞＋結果補語“成”＋目的語＋助詞“了” 動詞“成”＋目的語＋“了”“了” 動詞＋結果補語“成”＋目的語＋助詞“了”	3 2
	動詞“成”＋“了”＋目的語 “了1”＋目的語	7	動詞“成”＋目的語＋“了” “了1”＋目的語	2
④	離合詞　間に“了”を置く	11	離合詞　後ろに“了”を置く	13
⑤	動詞＋“了”＋疑問詞	3	動詞＋疑問詞＋“了”	13
⑥	動詞＋“了”＋固有名詞	13以上	動詞＋固有名詞＋“了”	3
⑦	動詞＋“了”＋人称代詞	11	動詞＋人称代詞＋“了”	5
⑧	動詞＋“了”＋数量詞	1	動詞＋数量詞＋“了”	8
⑨	動詞＋“了”＋複合方向補語	11	動詞＋複合方向補語＋“了”	17
⑪			連動文	18
⑫			固定フレーズ（成語）＋“了”	18
⑬			動詞＋回数＋“了”	3
⑭			時量＋没＋動詞＋“了”	12
⑮			様態補語＋“了”	18
⑯			助動詞＋動詞＋“了”	10
			“该……了”	5
⑰			可能補語＋“了”	14

述語や目的語の性質が同じ場合でも、助詞“了”の位置は必ず一致するとは限らないことが、表12からわかる。

表12の動詞の後ろの“了1”と文末の“了2”について、①、③、④、⑨の例を比較してみる。

31 ⑩一般名詞については、数が多いため省略する。

32 文末に“了”が完了の意味を表していれば、それはふつう“了1＋了2”である。劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 312。

表 13 ①

動詞＋結果補語＋“了 1”＋場所 / 時間	動詞＋結果補語＋場所 / 時間＋“了 2”
井筒非常之黑，黑得如同把世间所有种类的黑一古脑儿**煮在了里边**。 唱到这里时，旗升到旗杆中间，“遍覆青苔”一音刚落，国旗便**爬到了顶尖**。 其实他也没说什么绘声绘色的话，但他一开口，女孩大多听得入神，一副痴迷的样子，不觉之间便喝得昏头昏脑，结果和他**睡到了一起。** 我们在应该支付代价的时候没有支付，那笔账便**转到了今天**。 我也因此陡然**冲到了顶峰**。 但谁也没搭腔，如同小石子**掉进了无底洞**。	于是这一带的人就说：准保**掉进那荒草地的井里了**。” 暑假结束，新学期一开始，直子便十分自然地，水到渠成地**走在我身旁了**。 而是突然**消失到什么地方了**。 少顷，我突然想起水仙花**忘在楼下了**。 不知不觉就**沉浸在这岑寂之中了**。 突然睁眼时，直子已经**站在那里了**。 你怕是把直子**忘在脑后了**吧？” 月亮也被**掩到后面去了**。 我想劝阻，但他头也不回地倏忽**隐没在黑暗中了**。 喂，木月，你终于把直子**弄到手了**！ 事情算是已经**落实到该落实的地方了**？”

表 13 ①については、“把”構文では 1 例を除き、文末に“了”が置かれている。これは次の表 13 ③も同様である。

表 13 ③

動詞＋結果補語“成”／動詞“成”

動詞＋結果補語“成”＋“了 1”＋目的語	動詞＋結果補語“成”＋目的語＋“了 2”
这一年里，直子**瘦成了另一个人**。 全被**染成了深色**。 他给我**养成了卫生习惯**。	你都**湿成这样了**。 或许把我**错看成另外某个人了**， 放下伞不**淋成落汤鸡了**？”
動詞“成”＋“了 1”＋目的語	動詞“成”＋目的語＋“了 2”
我们果真**成了朋友**。 由于写这本书的人是关西人，我做的菜也就跟着**成了关西风味**。 而在我们这个小天地中，反常恰恰**成了前提条件**。 简直**成了病态**。 后来规模逐渐扩大，**成了法人**。 但和昨晚不同，**成了毛毛雨**。 我简直**成了活人自动唱机**。	“**成大人了**嘛。” 我在亲戚中岂不**成块笑料了**！”
動詞“成”＋“了 1”＋修飾語＋目的語	
三人在一起，便俨然**成了电视中的专题采访节目：** 结果赶到直子房间时，蛋糕已经土崩瓦解，活脱脱**成了古罗马的圆形剧场，** 可如今却**成了精神病院的归来者**。 直子来到后，我就没完没了地弹甲壳虫，活活**成了可怜的音乐奴隶**。” 星期天就**成了洗东西熨东西的日子**。 恐怕就**成了我的责任**， 结果**成了他得的最后一分**。	

③“成”は「(…に) なる」、「(…に) 変わる」という意味を表し、うしろには必ず名詞の目的語が必要である [33]。例の数は少ないが、“把”構文や反語文では文末の“了”が用いられている。

33 『中国語文法用例辞典』東方書店（2003）

表 13④

離合詞

離合詞（動詞）＋“了”＋離合詞（目的語）	離合詞＋“了”
直子说声下车吧，我们就**下了车**。	就我本人来说，本打算租间公寓，一个人落得逍遥自在，但想到私立大学的入学费和学费以及每月的生活费，也就不好意思**开口了**。
随后一个人掏腰包**付了账**。	说完，便和她**分手了**。
就在绿子家附近发生了一场火灾，我们爬上三楼的晾衣台看热闹，而且不知不觉地**接了吻**。	医院？莫非她受伤或**患病了**不成？
也不知为什么，我总是觉得似乎自己在茂密的森林中**迷了路**。	出院后不久和丈夫**结婚了**。
住院不久，小手指可以动了，便去音乐大学复学，总算**毕了业**。	我跑上二楼，敲姐姐房间的门，喊声**吃饭了。**
可当时就好像**着了魔**似的。	“**说话了**？”玲子问直子。
侍者把信用卡拿来，永泽确认一下款额，用圆珠笔**签了名**。	她那嘴巴也实在灵巧，听着听着就**入迷了**。
随即**付了账**。	“我们就**离婚了**，
大概这个那个弄得太紧张了，以致月经也**乱了套**。	“怎样，这几天可又**失火了**？”
过一会儿，他问我**吃了饭**没有。	绿子说想爬树，不巧新宿没有可爬的树，御苑已经**关门了**。
结果见面一看，我马上一颗心**落了地**－	“对了对了，上次那场考试**发榜了**，中了。”永泽说。
	发型变了嘛”，结果还是**落空了**。
	“喂喂，学校早已**开学了**。”

④離合詞の“了”の位置については、3-1-5 でみたように、「動詞と目的語が比較的固定した組み合わせの動目フレーズであり、ふつう動詞の前にさらに時間に関わる状語があれば目的語の前に定語がなくてもかまわない」とあるが（脚注 23）、実際には前に時間に関する状語がなく離合詞の動詞部分の後ろに“了 1”が置かれる例が多かった。また韓志剛[34]らも目的語が単独の場合、“了”の位置は文中でも文末でもよいとして、離合詞“回家”の例をあげ、“他回了家。”、“他回家了。”いずれも可としている。表 13④に示した《挪威的森林》中の両者も、特に際立った違いは見出だせなかった。

表 13⑨

複合方向補語

動詞＋“了”＋複合方向補語	動詞＋複合方向補語＋“了”
大约过了十五分钟，教室的门开了，绿子**闪了进来**。	若无其事地**做起广播体操来了**。
所以这些民房才得以原样**保留了下来**。	“怎么突然**想起这个来了**？”
之后，我们就势**睡了过去**。	“这里怎么**冒出黄瓜来了**？”
玲子从厨房拿出一个招手猫形的贮币盒，直子从钱包里找出一枚百元硬币，**投了进去**。	“书店不想再**办下去了**？”我试着问。
就这样，皱纹一条条**爬了上来**。”	但不管怎样，总算把毛衣按时**赶出来了**。
就反对这门婚事，**吵了起来**。	行李都**寄过去了**。”
	猫在玲子腿上伸伸要，换了姿势，又**睡过去了**。

劉らによると、“了 2”は派生的意味を表すいくつかの複合方向補語の前に置くことができ、“動詞＋‘了’＋複合方向補語”と“動詞＋複合方向補語＋‘了’”の意味は全く同じとある[35]。ただし守屋は動詞の後ろに複合方向補語があり、さらに目的語がある場合、“了”は文末に置かれる[36]と述べており、《挪威的森林》においても、目的語をもつ複合方向補語は、すべて文末に“了”が置かれていた。

4　おわりに

本稿では日本語の過去の表現が、中国語でどのように訳されるかについて『ノルウェイの森』の中国語翻

34　韓志剛・張文賢（2004）《HSK 语法指要与训练》p. 128
35　劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）p. 307
36　守屋宏則（1997）『やさしくくわしい中国語文法の基礎』p. 191

訳書《挪威的森林》を分析して検討した。その結果日本語の過去事象は、完了を表す助詞“了”を用いて表現するばかりではなく、動詞単独で過去を表現したり、様々な表現形式がある事がわかった。中国語には時制がないこと、そして日本語の過去事象を中国語で表現する時に完了を表す助詞“了”は必ずしも必要ではないことを改めて証明することができたと思う。しかし過去を表現する時に、“了”を動詞の後ろに置くか、文末に置くかに関しては、絶対的な法則を見出すことができなかった。これは“了”が動作の完了を強調するだけでなく、すでに起こった事実を確認したり、様々な語気を表すこともあるため、過去事象の表現に用いることもあれば、用いないこともあるからだと思われる。ひとつの考えとして、目的語が初めて話題にのぼった時は“了”は目的語の前、目的語がすでにわかっているものである場合は“了”は目的語の後ろに置くという記述がある[37]。これは文末に置く“了”が確認の語気をもつ語気助詞の一種であるため、目的語がすでにわかっているものの場合は、文末の“了”を用いると考えられる[38]。本稿の表13③で動詞“成”または結果補語“成”の後ろに置く“了”の位置について比較したが、“了”を動詞“成”の後ろにおいて“成了”の形をとるものが、文末に置く“了”より多いことがわかる。これは“成”には新しいものに「なる」、「変わる」という意味があるため、その目的語は当然「新しいもの」、「未知のもの」である可能性が高いために、助詞“了”は“成”の後ろ、目的語の前に置かれることが多いと考えられる。助詞“了”の習得は容易ではないが、より正しい中国語の運用のためには、避けて通れないものである。

＜参考文献＞

村上春樹（2004）『ノルウェイの森』（上）（下）株式会社講談社
林少華訳（2005）《挪威的森林》上海訳文出版社
相原茂・陳淑梅・飯田敦子（2019）『日中いぶこみ交差点』朝日出版社
清原文代・井手克子・韓艶玲・高恵敏（2012）『一歩一歩学漢語』下　白帝社
杉野元子・黄漢青（2013）『大学生のための初級中国語40回』白帝社
沈国威監修　氷野善寛・小嶋美由紀・海暁芳・紅粉芳恵・阿部慎太郎（2020）『中国語でコミュニケーション』教授用資料　朝日出版社
郭春貴（2008）『中国語検定対策3級・4級文法編』白帝社
郭春貴（2014）『誤用から学ぶ中国語』白帝社
韓志剛・張文賢（2004）《HSK语法指要与训练》
瀬戸口律子（2006）『完全マスター中国語の文法』語研
孫徳金（2003）《汉语语法教程》北京語言大学出版社
守屋宏則（1997）『やさしくくわしい　中国語文法の基礎』東方書店
楊寄洲（2004）《汉语教程》第二册 上　北京語言大学出版社
劉月華・潘文娯・胡韡（1996）『現代中国語文法総覧』（相原茂監訳）くろしお出版
『中国語文法用例辞典』東方書店（2003）
『中日辞典』講談社第二版（2002）
微文语录 https://www.wwyulu.com/read/11205.html

37　韓志剛・張文賢（2004）《HSK 语法指要与训练》p. 128　以下は原文である。
根据什么决定“了”放在前边还是后边呢？（1）根据宾语的性质。如果宾语是第一次谈到，“了“应该在宾语前边：如果动词宾语表示的事情是已经想到、知道的事，“了”应该在宾语后边。

38　孫徳金（2003）《汉语语法教程》p. 254 に次のような記述がある。
b 句（他去上海了。）的“了”表达的是确认“他去上海”已成为事实。

外国語教育研究ジャーナル投稿規定

概要：言語教育に関する研究論文、研究報告、及び書評を年 1 回出版

Ⅰ．投稿資格

1. 執筆者は（共著の場合はファーストオーサーのみ）、原則として現職の外国語教育研究センター所属教員とする。特別号、特集号などの場合はこの限りではない。
2. 投稿論文はそのカテゴリーを問わず 1 教員につき 1 本までとする（共著による論文も 1 本と数える）。ただし、投稿論文数によっては、Journal & Research 委員会による決定を経て掲載号・論文枚数等の調整を受ける場合がある。
3. 投稿論文は過去に出版されておらず、他のジャーナルに現在投稿されているものではないこと。

Ⅱ．使用言語

原稿執筆にあたっては、センターに所属する教員が広く互いの研究及び知識の共有を活性化するにあたり、原則として日本語あるいは英語での執筆を基本とする。朝鮮語、中国語、フランス語、ドイツ語、スペイン語による原稿執筆も可。

Ⅲ．執筆要項

投稿原稿は未公版のものに限る。

1. 原稿の種類は、次のいずれかの区分に属するものとする。区分は執筆者が判断するが、ジャーナル & リサーチ委員会が確認し、必要に応じて、変更を促すこともある。
   （1）研究論文：当該分野で一定の結論を出し、その分野に貢献すべき内容が分析、議論されているもの。従って、執筆者は少なくとも、セクションとして導入と結論を明記し、先行研究を踏まえた当該論文の位置づけ、分析や議論（またはそれに代わるもの）にあたる記述を含めること。
   （2）研究報告：当該分野における研究の進捗状況を公表し、将来その分野において貢献すべき可能性のある研究経過をまとめたもの。
   （3）書評：書籍や文献を紹介しまとめたもの。
2. 書式は以下の項目全て可能な限り厳守すること。
   （1）原稿サイズ
   A4 版を使用し、上下各 25mm、左右各 25mm ずつあける。
   横書きでフォントは日本語（MS 明朝）、英語は（Times New Roman）12 ポイントを使用。その他の言語のフォントは同一論文内で齟齬のないよう当該言語における学術論文で一般的に使用されるフォントを選択し使用すること。行間は 1 行とし、本文は左揃えとすること。
   （2）字数
   研究論文：20000 字程度
   研究報告：6000 字程度
   書　　評：2500 字程度
   図表、参考資料、参考文献、注、Appendix など全て含める。なお、図表については、明瞭なものを当該個所へ貼りつける。貼り付けられない場合は、別ファイルを用意し、挿入個所を明示する。
   （3）原稿タイトル
   18 ポイントでセンタリングし、各文字を太字にしたスタイルに従う。フォントは上記書式に従う。
   （4）氏名
   右寄せ、ゴシック体で 12 ポイントとする。タイトルとの間は 1 行あける。

(5) 要旨

全ての原稿に、要旨と 3 ～ 5 項目のキーワードをつけること。要旨は 500 字程度で執筆すること。書式は、左右 15mm ずつ全行インデントし、フォントは MS 明朝、11 ポイントを使用する。尚要旨の言語は執筆言語に関わらず英語あるいは日本語とする。

(6) 本文への註釈は、対応する註記を各ページの下に 9 ポイントで表記する。

Ⅳ. 原稿の提出

原稿の提出は、指定の Google フォームに必要事項を入力のうえ、投稿内容を収めた電子媒体をアップロードし、送信すること。

Ⅴ. 脚注および参考文献の形式

原則的に APA スタイル（第 7 版）を用いること。執筆者の責任において同一論文の中で齟齬のないよう確認すること。英語以外の言語での執筆の際は、APA（第 7 版）の形式に可能な限り沿うよう配慮すること。剽窃を行わないよう十分に注意すること。

Ⅵ. 投稿開始及び締切日

投稿の受付は毎年 4 月春学期の開始から行う。初稿の提出締め切り日は 6 月初週最終日とする。

Ⅶ. 査読に関して

研究論文カテゴリーへの投稿は、匿名の査読員 1 名による査読が行われる。査読は規定のルーブリックを用いて行い、査読員は任意のコメント、修正案等を追記することができる。査読終了後 Journal & Research 委員会より、春学期終了前までに執筆者へ査読結果を報告し校正等の打診を行う。研究報告及び書評への投稿は原則として査読は行わない。ただし基本的な内容及び体裁のチェックを Journal & Research 委員会が行い、掲載可否を判断する。

Ⅷ. 校正及び再提出

査読を受けた執筆者は、論文の校正を行い、秋学期第 1 週目末日までに再提出を行う。校正後の原稿は Journal & Research 委員会による最終確認を経て掲載可否の判断を行うものとし、執筆者に結果を通知する。なお出版社より体裁等の修正依頼があった場合は、再度の校正を執筆者に依頼する場合がある。

Ⅸ. 出版

当該ジャーナルは毎年 12 月に出版される。

Ⅹ. CiNii 及び立教リポジトリへの登録

掲載された論文は、外国語教育研究センターへ掲載され、CiNii（国立情報学研究所論文情報ナビゲーター［サイニィ］）及び立教大学学術リポジトリに登録される。

その他の要件

1. 原稿料は支払わない。
2. 掲載された論文の著作権は、原則として立教大学外国語教育研究センターに帰属する。ただし、著者が著者自身の研究・教育活動に使用する際は、許可なく使用することができるものとする。
3. 万が一出版後、剽窃等の不正が発覚した場合は当該論文をジャーナルから削除する。

Journal of Foreign Language Education & Research
Manuscript Submission Guidelines

Scope: The journal annually publishes research articles, research briefs, and book/article reviews.

I. Eligibility

1. Contributions to the journal are primarily limited to individuals affiliated with Center for Foreign Language Education & Research (FLER). In the case of co-authored papers, this requirement applies only to the first author. Exceptions may be made for special editions.
2. A maximum of one contribution per issue is accepted (co-authored papers are also counted as one contribution). Due to space limitations, your submission may be considered for publication in a later issue, or you may be asked to reduce the length of the submitted article.
3. Work submitted to the journal should not have been previously published and should not be under consideration for potential publication by other journals.

II. Language

In order to effectively share knowledge and research activity amongst FLER-affiliated instructors, we accept, in principal, manuscripts written in either English or Japanese. Manuscripts written in Korean, Chinese, French, German, or Spanish may be accepted.

III. Content and Formatting Guidelines

1. Contributions are limited to previously unpublished work.
2. Upon submission, authors are required to select one of the categories listed below for their contribution. Appropriateness of the selected category will be reviewed by the Journal & Research Committee. You may be requested to change the submission category, if necessary.
   (1) Research Articles: An article which can be considered as making substantial contribution to the relevant field through explicit analysis of data or theoretical discussion which leads to a clear and sound conclusion. The author is expected to include an introduction section in order to situate the research in reference to previous empirical research and/or theoretical discussion and a clear conclusion section well-connected to the research findings or discussion presented.
   (2) Research Briefs: An article which reports an aspect of a research project underway at the time of writing and demonstrates the potential of the topic for generating further questions of interest or prospective contributions in the relevant field.
   (3) Book/Article Reviews: A critical review of a published book or article relevant to the field.
3. Please ensure to follow all formatting guidelines listed below. Authors who intend to write in Japanese should refer to the Japanese version of this document.
   (1) Size: Use A4-sized paper, leaving margins of 25mm at the top and bottom and of 25mm on both sides of the text. The letters in the text should be Times New Roman 12 point and single-spaced. For Japanese Use (MS *Mincho)*, for any other languages, use the font type considered standard in the selected language. Use the same font consistently

throughout the manuscript unless otherwise noted.

(2) Length: Research articles should be approximately 8000 words in length, including graphs, charts, the reference list, and appendices. Research briefs should be approximately 3000 words, and book/article reviews should be approximately 1000 words. Graphs and charts should be embedded in the text. However, if it is difficult to do so, please submit as a separate file, but leave space and indicate where they should go in the text.

(3) Title: The title should be in 18 point and centered following the capitalization rules—font as above.

(4) Author's name: The name of the author/s should be indented to the right side and written in Gothic 12 point. Leave one line between the title and the name of the author/s.

(5) Abstract: Manuscripts written in English should be accompanied by a 150-250 word abstract in either English or Japanese, which includes 3 to 5 keywords for the article at the bottom. For the abstract, the entire text should be indented 5 spaces from the left and right and written in Times New Roman 11 point.

(6) Footnotes: Footnotes should be placed at the bottom of each page, in 9 point.

IV. Manuscript Submission

To submit a manuscript, fill in the necessary details in the designated Google form, upload the electronic file containing the article content, and submit.

V. Footnotes and Referencing

The author is responsible for consistently adhering to APA (7th edition). Make sure to avoid committing plagiarism.

VI. Call and Deadline for Submission

Submissions begin at the beginning of Spring semester. The deadline for submission is the last day of the first week in June. Call for submission is circulated via SPIRIT email.

VII. Peer Review

All submissions to the research article category will be reviewed by one anonymous reviewer. The review is conducted with a designated rubric and the reviewer may provide you with comments and suggestions for further improvement of your work. After the review is completed, review results will be sent to the author from the Journal & Research Committee before the end of Spring semester. No peer review will be conducted for research briefs or book/article reviews. However, the Journal & Research Committee will check if the submitted work meets the journal submission guidelines and determine if the work is publishable.

VIII. Revision and Resubmission

After receiving the results of peer review, the author is expected to revise the manuscript for resubmission. The revised manuscript must be submitted by the end of the first week in Fall semester. A final review will be conducted by the Journal & Research Committee to determine if the work is publishable. The author will be notified of the decision once the final review is completed. The author may be asked to further revise the manuscript if there is any stylistic/format issue.

IX. Journal Publication

The journal is published annually in December.

X. Registration on CiNii and Rikkyo Repository

Contributions to the FLER Journal will be published on the FLER webpage and registered on the national CiNii database and the Rikkyo University Academic Repository.

XI. Other conditions

1. No remuneration is offered to the author(s).
2. The copyright of articles published in the FLER Journal resides with Center for Foreign Language Education & Research, Rikkyo University. However, the author(s) retains the right to use his/her work for future research and/or educational purposes without permission.
3. If any plagiarism or misconduct is discovered after the work is published, the published work will be removed from the journal.

【執筆者・Authors】

Daisuke Baba

Daisuke Baba is an adjunct lecturer of German language at Rikkyo University, from which he received his Ph.D. in March 2019. Baba's research focuses on the academic exchange between Germany and Japan in the late 19th and early 20th centuries. He is the author of A Hybrid Origin of Modern Historical Writing of Japanese Literature. Karl Florenz's A History of Japanese Literature and the German–Japanese Contact of Academic Cultures (Tokyo: Sangensha 2020).

Deborah Maxfield

Deborah Maxfield is currently serving as a Lecturer in English Education at Rikkyo University. Her main research interest is applying concepts from her MSc in Psychology to second-language education, particularly in how group dynamics combine with cognitive and affective states to affect learning outcomes, student motivation, and L2 anxiety. Her recent research included development and psychometric testing of the first questionnaire to measure L2 group cohesion, and its applications for improving student learning experience.

Ian Hart

Ian Hart is a lecturer in English education at Rikkyo University's Center for Foreign Language Education and Research. He holds a master's degree in Applied Linguistics and TESOL, a bachelor's degree with honors in Design, and has taught at universities such as Waseda University and Toyo University. His research interests include technology-enhanced language learning, instructional design theory, and learner motivation.

Jeffrey Mok

Jeffrey Mok is currently a specially appointed Associate Professor at FLER, Rikkyo University. His research interests include but not limited to learning and teaching pedagogies, second language acquisition, technology enhanced language learning, and communication skills. Jeffrey was a senior faculty developer at National Technological University and taught at the National University of Singapore. He was also a senior language specialist and trained English teachers in Reading and Writing at the Regional Language Centre, Singapore.

Joshua Rappeneker

Joshua Rappeneker is an adjunct lecturer at the Center for Foreign Language Education and Research. He has developed and taught higher education courses in Japan for over 10 years and is currently researching CALL & the development of spoken fluency at Rikkyo University. He is also interested in the use of technology in the classroom and developing software to assist in classroom management.

Satchie Haga

Satchie Haga is a Lecturer at Rikkyo University in their Center for Foreign Language and Education Research. Her work examines how technology intersects with instructional processes within the context of foreign language learning and English Medium Instruction. Areas of interest include Activity Theory, computer-supported collaborative learning, virtual exchange, and English for Specific Purposes viewed from a sociocultural perspective.

James Carpenter

James Carpenter has an MA in TESL and an M Ed in educational technology from Northern Arizona University. He has taught students from all over the world. His research focuses on how people learn in unique situations. He is currently an adjunct lecturer at Rikkyo University, and is pursuing a PhD in Education, Concentration in Applied Linguistics at Temple University.

Daniel Beck

Daniel Beck is an Adjunct Lecturer in the English Language Program of the Center for Foreign Language Education and Research at Rikkyo University. He has served actively on the Faculty Development Committee and the Digital Communications Committee. His research interests include technology in education, teacher education, and communities of practice.

Masakazu Mishima

Masakazu Mishima, Ph.D is an associate professor at Center for Foreign Language Education & Research at Rikkyo University. His research interests extend to a number of areas in the field of TESOL/Applied Linguistics including SLA, language teacher education, NNEST issues, World Englishes, Language Assessment, CALL, and Second Language Writing and Pedagogy.

Yuka Yamamoto

Yuka (Maruyama) Yamamoto, Ph.D., Applied Linguistics, is a professor at Aoyama Gakuin University. She has extensive experience in designing curricula and creating textbooks. Her current research interest is in CLIL curriculum development.

Tanya L. Erdelyi

Tanya Erdelyi is an adjunct lecturer at Rikkyo University's Center for Foreign Language Education and Research. She has extensive experience teaching academic writing and research to undergrad and graduate students, and developing curricula and teaching materials for various second language courses. Tanya's current research interests include academic writing for second language learners, content and language integrated learning (CLIL), and language immersion.

李 菲（リ フェイ）

李菲（リ　フェイ）。立教大学外国語教育研究センター教育講師（中国語）。研究分野は、中国語学と認知文法。現代中国語（北京語）の様々な文法現象について、認知文法の理論と観点からとらえ、記述することを中心に研究を行う。これまで、中国語の結果補語と離合詞に関する論文を発表している。近年、言語習得の観点から見た文法記述のあり方、構文から見た文法体系の再構築といった研究課題にも関心をもつ。

岡村　佳奈

日本学術振興会特別研究員、帝京大学外国語学部非常勤講師などを経て、立教大学外国語教育研究センター教育講師。博士（学術）。専門は社会言語学、語用論、韓国語学。これまでは、日本語と韓国語におけるあいさつの使用とその意識、あいさつの日韓差に関する原因、日本人学習者による韓国語の使用様相および語用論的失敗などについて取り組んできた。

Alex Blumenstock

Alex Blumenstock is an adjunct lecturer at the Center for Foreign Language Education and Research at Rikkyo University. He has also taught in the United States and Osaka. He holds an M.A. in English and a TESOL certificate, both from East Tennessee State University. His current research interests are instructional technology, peer feedback, collaborative learning, and flipped classrooms.

Devon Arthurson

Devon Arthurson works as an adjunct lecturer at the Center for Foreign Language Education and Research. She has been teaching English in Japan since 2008. Devon earned her Bachelor of Social Work from the University of Manitoba and completed her Master of Arts in Integrated Studies from Athabasca University. Her current teaching and research interests include fostering learner autonomy and providing a platform for students' voices. Her volunteer activities include social justice issues such as poverty alleviation and awareness-raising about human trafficking.

Heather Woodward

Heather Woodward is a lecturer in English education at the Center for Foreign Language Education and Research. She has taught higher education courses since 2019 and is currently conducting research on spoken fluency and mobile application development. She is also interested in the material development of lessons for learners of English as a second language.

Charu Gupta

Charu Gupta is currently an English lecturer at Rikkyo University, Japan. She received her M.S.Ed. in TESOL from the University of Pennsylvania in 2018. Her research interests include reflective practices, educational technology and task-based language teaching.

Jon Mahoney

Jon Mahoney is a lecturer in English education at the Center for Foreign Language Education and Research. He has taught higher education courses in the Tokyo area for over 5 years, and is currently conducting research on how to enhance critical thinking skills in Japanese university students. Other areas of research he is interested in include CLIL, English as a lingua franca, and the usage of virtual exchanges.

Omar Shelesh

Omar Shelesh is a lecturer in English education at Rikkyo University's Center for Foreign Language Education and Research. He holds a Postgraduate Certificate in Education from the University of Cambridge and has developed and taught higher education courses in Japan for over 7 years. He is currently conducting research on bilingual pedagogical practices.

白方直美

白方直美（シラカタ　ナオミ）立教大学教育講師

北京の大学に留学した後、中国広州市の華南師範大学に日本語教師として 2 年間勤務した。この間、広東地方で使用されている広東方言特有の音訳外来語に興味を持ち、研究を始める。帰国後大学院に進学し、標準中国語に影響を与え始めていた広東方言について修士論文を執筆した。現在は中国語の授業実践において生じた様々な文法上の疑問について研究している。

外国語教育研究ジャーナル　第３巻
(JOURNAL OF FOREIGN LANGUAGE EDUCATION AND RESEARCH, Vol. 3)

発行日　2022年12月８日

発行者　立教大学 外国語教育研究センター
(Center for Foreign Language Education and Research, Rikkyo University)
〒171-8501　東京都豊島区西池袋３－34 －１

製　作　株式会社 遊文舎
〒101-0061　東京都千代田区神田三崎町２－20 －７　水道橋西口会館 8F

ISSN 2436-0325

JOURNAL OF FOREIGN LANGUAGE EDUCATION AND RESEARCH

Vol. 3

December 2022

FLER

RIKKYO UNIVERSITY
Center for Foreign Language
Education and Research